

**A realist evaluation of the role of external  
peer review programmes in improving the  
quality of mental health services**

---

Sarah Jasim

00549986

Faculty of Medicine, Imperial College London

Thesis submitted for the degree of Doctor of Philosophy

## **Abstract**

**Background:** Available evidence suggests that membership of external peer review programmes can help improve the quality of health care services. However, little is known about how this is achieved and what key mechanisms and contexts are essential for quality improvement.

**Methods:** I undertook a mixed methods realist evaluation of peer review networks and accreditation schemes in inpatient and community-based mental health services provided by the Royal College of Psychiatrists' Centre for Quality Improvement (CCQI). Informed by a systematic literature review, I collected qualitative data from coordinators (four focus groups) and participants (122 interviews) of external peer review programmes. I also collected quantitative data from 178 community-based memory clinics and 33 inpatient mental health services to examine whether organisational readiness for change influenced service quality.

**Results:** Causal mechanisms including sharing and learning, consultation, and engagement of senior management and junior staff were essential for sustained quality improvement. The most salient contexts were type of external peer review and length of membership. In accreditation schemes, most changes occurred before or during self-review, and following written feedback for peer review networks. A two-level linear model signalled services with higher baseline readiness for change achieved greater quality improvement through membership of a peer review network, however findings were not statistically significant. Qualitative findings echoed the importance of readiness for change constructs.

**Conclusions:** Differences in when change occurs between peer review networks and accreditation schemes should be considered by organisations that provide external peer review programmes. Sharing and learning was the main essential causal mechanism of external peer review programmes. To maximise the benefit of participation, this mechanism should be further supported and

enhanced. A future of increased competition in healthcare could reduce sharing and learning opportunities; indicating a need to further develop the evidence base for external peer review.

## **Acknowledgements**

Primarily, I would like to acknowledge and show my utmost gratitude to all the study participants from the CCQI, forensic mental health services, psychiatric intensive care units and memory clinics. Their efforts, time and reflections deeply enriched the quality of the data I could collect. I would also like to thank those who provided me with access to the participants (Samantha Holder, Holly Robinson and Sophie Hodge). As I mostly collected data from locked inpatient environments, I felt privileged to be able to gain this experience. The initial visits collecting data for the evaluation of low secure units (eLSU) study provided a valuable introduction into the 'world' of quality improvement in forensic mental health and it was the start of a very rewarding journey.

I cannot thank my supervisor enough, Professor Mike Crawford, for his guidance and mentoring throughout my PhD. His support and constant encouragement from the onset, have definitely made the PhD experience my most memorable academic degree at Imperial College London. His attention to detail, pastoral care, inspiring and thoughtful supervisions, and inspiring leadership style have made the world of difference.

I would also like to express my sincere gratitude to the mentors that I have had the pleasure of working alongside at the CCQI. Dr Alan Quirk provided unparalleled qualitative guidance and support throughout my PhD, and I am truly grateful and honoured for the numerous hours of discussion, thinking space and mentoring he has offered. Dr Lina Aimola and I worked on the eLSU study together, and I would like to thank her for the encouragement and great support she provided me with during the data collection and data entry phases of the eLSU study. I am so grateful for her support, guidance and encouragement throughout. I would also like to thank and recognise Dr Adrian Worrall for his guidance from the very beginning. His early input in helping navigate key concepts in the field of quality improvement was pivotal to shaping my research questions and structuring my systematic literature review. I would like to

additionally thank Dr Alan Quirk and Dr Rob Baskind for looking over my qualitative topic guides.

I would like to thank and recognise Joseph Eliahoo from the Statistical Advisory Service at Imperial College London for checking over my statistics plan. I would like to express heartfelt gratitude to my dear family friend Miss Misbah Ahmed for the evenings and weekends of help and support she dedicated to increase my confidence in advanced statistics. I am truly grateful for her guidance and help with familiarisation of the software package SAS. I would like to recognise early academic conversations with Dr Jonathan Benn, Dr Sankalp Chaturvedi and Mr Christopher Nicolay for their valuable insights. I would like to thank the team at CLAHRC North West London for their helpful inputs and inclusivity, and for the opportunities to discuss my research at CLAHRC collaborative learning events.

Throughout my PhD, various academic staff have contributed to my academic development. I would like to express sincere gratitude to my early and late stage assessment examiners, Professor Thomas Barnes and Dr Christine Franey, for their constructive thoughts and for contributing to the development of my research focus. I would like to extend my thanks to Imperial College Library, especially the librarians Natasha Suri and Geraldine Farrell, for assisting me with numerous document delivery requests. I would like to show my appreciation to the Imperial College Graduate School Staff, especially Paul Seldon and Caroline Hargreaves for their comprehensive professional skills training that has been invaluable to both my academic and personal development. The support of both the Centre for Psychiatry administrators, Nicole Hickey and Sheila McKenzie, has been invaluable throughout my PhD. I would also like to show deep gratitude towards all my colleagues at the Centre for Mental Health for all their guidance and support. Their constant motivation, encouragement and advice has been immeasurable.

I would like to express my heartfelt gratitude towards my older sisters Sharon and Simone, and their families, for their life-long inspiration, love and early guidance. I am most grateful to have a wide circle of friends who have offered great emotional support and much needed moments of laughter and joy throughout my PhD.

Finally, and most importantly, I would like to dedicate this thesis to my parents: Mohammad and Rehana Jasimuddin. My parents provided my initial inspiration to improve quality and reduce inequity in healthcare, through visits to developing countries and rural areas of the world in my formative years. This had a long-lasting impression on my future decisions and academic choices, and I am so appreciative for these experiences. Their unwavering love and endless support throughout all my years at Imperial College London have been remarkable, and words cannot express how grateful I am for their compassion, kindness, generosity and encouragement. They have truly surpassed the role of parents and have been a lifeline throughout my PhD, for which I will be eternally grateful.

## **Declaration**

I declare that I am the sole author of this thesis, and all the work within is my own, except where referenced or carried out in collaboration with others who are appropriately credited.

Signed: Miss Sarah Jasim

Date: 5<sup>th</sup> December 2016

Final word count: 75,9966

**The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.**

## Statement of Work

I collected and analysed most of the data for this PhD. As my thesis ran parallel to the Evaluation of Low Secure Units (eLSU) study, a statement of my input to the data and analyses presented in this thesis is required to highlight my contribution. I recruited and facilitated four focus groups and 121 interviews for the qualitative component of the study, and completed the analysis myself. I selected the readiness for change assessment tool ([Bobiak et al., 2009](#)) for the quantitative component. I made the decision to embed this tool in both the National Memory Services Audit 2013 to acquire cross-sectional data, and then the manager's checklist component of the eLSU to obtain longitudinal data. Data collection for the manager's checklist was carried out by myself, and two colleagues from the CCQI (Dr Lina Aimola and Dr Neeraj Tripathi) who were collecting data for the wider eLSU study. I then completed the quantitative analysis by myself, having checked the plan for statistical analysis with a statistician.

Therefore, I undertook the majority of data collection for this thesis, and all the analyses, and it presents original work which is distinct from the parallel eLSU study.



## Table of Contents

Statement of Work .....	8
Glossary of terms, abbreviations and acronyms .....	19
Chapter 1 Introduction.....	22
1.1 Overview of quality in healthcare .....	23
1.2 Evolution of quality improvement.....	26
1.3 External quality improvement approaches .....	35
1.4 External peer review .....	38
1.5 Statement of the problem .....	49
Chapter 2 Realist Evaluation .....	51
2.1 Mechanisms .....	54
2.2 Outcomes (change).....	55
2.3 Context.....	57
2.4 Context-mechanism-outcome configurations .....	61
Chapter 3 Systematic Literature Review.....	64
3.1 Aims .....	64
3.2 Design.....	65
3.3 Search Strategy.....	66

3.4	Data extraction, quality assessment and data synthesis.....	74
3.5	Overview of Results.....	76
3.6	Results: Aim 1: Summarise what is already known about how change occurs in external peer review to generate improvements in service quality	87
3.7	Results: Aim 2: Investigate which programme factors are important for generating improvements in service quality in external peer review.....	95
3.8	Results: Aim 3: Identify the non-programme (contextual) factors which facilitate or impede their success.....	105
3.9	Discussion.....	116
3.10	Conclusion.....	120
Chapter 4	Research Context.....	122
4.1	Royal College of Psychiatrists – CCQI .....	123
4.2	Forensic Inpatient Mental Health.....	124
4.3	Memory Clinics.....	128
4.4	eLSU study.....	129
Chapter 5	Aims.....	130
Chapter 6	Research Methodology.....	132
6.1	Introduction.....	132
6.2	Critical Realism .....	133

6.3	Qualitative Methodology.....	137
6.4	Quantitative Methodology.....	141
6.5	Mixed Methods Research.....	146
6.6	Reflexivity: role as an external researcher.....	152
6.7	Initial Middle Range Theory (MRT) .....	153
Chapter 7 Qualitative Methods .....		154
7.1	Focus Groups.....	155
7.2	Semi-structured Interviews.....	159
7.3	Reliability, Validity, Rigour and Authenticity.....	170
Chapter 8 Quantitative Methods .....		173
8.1	Cross-sectional Memory Clinics Audit .....	173
8.2	Readiness for Change Longitudinal Study.....	178
Chapter 9 Results of Focus Groups .....		186
9.1	Practitioner Demographic Characteristics.....	186
9.2	Overview of Themes .....	188
9.3	When Change Happens .....	189
9.4	Mechanisms within Programme Factors .....	192
9.5	Outer Contexts.....	196

9.6	Inner Contexts .....	200
9.7	Readiness for Change .....	203
9.8	Conclusions.....	208
Chapter 10 Results of Semi-Structured Interviews .....		210
10.1	Service Characteristics.....	210
10.2	Overview of Themes .....	215
10.3	Programme Factors: When Change Happens .....	216
10.4	Mechanisms .....	225
10.5	Changes.....	236
10.6	Contexts.....	241
10.7	Organisational readiness to change .....	249
10.8	Conclusion.....	257
Chapter 11 Results of Cross-sectional Study.....		261
11.1	Response Rate .....	261
11.2	Service Characteristics.....	261
11.3	Summary of Readiness for Change Data .....	262
11.4	The relationship between readiness to change constructs and the primary and secondary outcomes of the National Audit.....	265

11.5	The differences between accredited and non-accredited memory clinics	266
11.6	Multivariate Model .....	279
11.7	Conclusion.....	281
Chapter 12	Results of Longitudinal Study .....	282
12.1	Response Rate .....	284
12.2	Relationship between baseline readiness for change, and baseline QELS checklist.....	284
12.3	Relationship between baseline readiness for change and change from baseline in QELS checklist score .....	285
12.4	Relationship between baseline readiness for change and change from baseline in QELS checklist score (accounting for the same respondent).....	287
12.5	Conclusions.....	289
Chapter 13	Discussion.....	290
13.1	Summary of Main Findings .....	291
13.2	Refined CMOCs .....	293
13.3	Aim 1: To explore stakeholder views of the causal mechanisms of external peer review programmes.....	296
13.4	Aim 2: To identify contexts which influence the ability of services to make use of participation in external peer review .....	302

13.5	Aim 3: To assess whether ‘readiness for change’ influences the ability of services to make effective use of an external peer review programme.....	304
13.6	Methodological Considerations .....	307
13.7	Recommendations .....	318
13.8	Final Remarks .....	325
	References .....	326
	Appendix A: Permission sought from authors .....	348
	Appendix B: Focus group participant demographic checklist .....	351
	Appendix C: CCQI focus group topic guide .....	352
	Appendix D: Draft e-mail to member services .....	355
	Appendix E: Participant information sheet.....	356
	Appendix F: Participant consent form.....	360
	Appendix G: Interview participant demographic checklist.....	362
	Appendix H: Interview topic guide .....	363
	Appendix I: Readiness for change tool ( <a href="#">Bobiak et al., 2009</a> ) .....	365
	Appendix J: Permission to republish tool ( <a href="#">Bobiak et al., 2009</a> ).....	366

## List of Figures

Figure 2.1 The realist evaluation cycle (adapted from ( <a href="#">Pawson and Tilley, 1997</a> )) .....	53
Figure 3.1 Preferred reported items for systematic reviews (PRISMA) flow chart .....	73
Figure 6.1 Convergent parallel study design (adapted from ( <a href="#">Creswell, 2014</a> )) .....	148
Figure 11.3 Differences in leadership between accredited and non-accredited memory clinics.....	267
Figure 11.4 Differences in adapting to new standards between accredited and non-accredited memory clinics .....	269
Figure 11.5 Differences in flexibility between accredited and non-accredited memory clinics.....	271
Figure 11.6 Differences in trying new things between accredited and non- accredited memory clinics.....	273
Figure 11.7 Differences in changing so fast between accredited and non- accredited memory clinics.....	275
Figure 11.8 Differences in the provision of services and treatments between accredited and non-accredited memory clinics.....	277
Figure 12.1 CONSORT flow-chart showing participation in the eLSU trial among eligible services ( <a href="#">Schulz et al., 2010</a> ).....	283
Figure 12.2 Relationship between baseline readiness for change (CC) and baseline in QELS checklist score (EC).....	284

Figure 12.3 Relationship between baseline readiness for change (CC) and change from baseline in QELS checklist score (EC).....	285
Figure 12.4 Relationship between baseline readiness for change (CC) and change from baseline (EC) in QELS checklist score (accounting for the same respondent).....	287



## List of Tables

Table 1.1 Distinctive features of networks (adapted from ( <a href="#">Malby and Mervyn, 2012</a> )) .....	42
Table 3.1 Five key reference papers.....	67
Table 3.2 Inclusion criteria, with definitions and rationales.....	68
Table 3.3 Final search strategy used in OVID for the review.....	70
Table 3.4 Electronic databases .....	71
Table 3.5 Included literature characteristics.....	77
Table 3.6 Themes from systematic literature review .....	86
Table 6.2 External peer review initial programme theory .....	153
Table 7.1 Different external peer review programmes.....	162
Table 9.1 Practitioner characteristics.....	187
Table 9.2 Focus group themes overview .....	188
Table 10.1 Purposive sampling of services .....	211
Table 10.2 SSI broad themes overview.....	215
Table 10.3 Number of subjects who had peer reviewed other services.....	220
Table 10.4 All mechanisms described by subjects .....	226
Table 10.5 All changes described by subjects .....	236

Table 10.6 All contexts described by subjects .....	241
Table 10.7 All organisational readiness to change constructs described by subjects.....	249
Table 11.1 Descriptive statistics for five readiness for change constructs .....	262
Table 11.2 Summary of outcomes of national memory clinics audit.....	263
Table 11.3 Relationship between readiness to change and the provision of services and treatments in memory clinics.....	265
Table 11.4 Relationship between readiness to change and waiting times in memory clinics.....	265
Table 11.5 Observed significance levels for waiting time.....	279
Table 11.6 Observed significance levels for maximum new service users.....	280
Table 12.1 Covariance parameter estimates between control and intervention LSUs .....	286
Table 12.2 Covariance parameter estimates between control and intervention groups accounting for the same respondent .....	288
Table 12.3 Observed significance levels from two-level organisational linear model predicting quality improvement (QELS checklist) scores based on readiness for change (n=33 services).....	288
Table 13.1 Realist matrix of context-mechanism-outcome configurations .....	293

## **Glossary of terms, abbreviations and acronyms**

AIMS –	Accreditation of Inpatient Mental Health Services
ALPHA –	Agenda for Leadership in Programmes for Healthcare Accreditation
CAQDAS –	Computerised Assisted Qualitative Data Analysis
CCQI –	College Centre for Quality Improvement, Royal College of Psychiatrists
CMOCs –	Context-mechanism-outcome configurations
CONSORT –	Consolidated Standards of Reporting Trials Statement
COPD –	Coronary Obstructive Pulmonary Disease
CPD –	Continuing Professional Development
CQC –	Care Quality Commission
CQI –	Continuous quality improvement
CQUIN –	Commissioning for Quality and Innovation
eLSU –	Evaluation of Low Secure Units
ExPeRT –	External Peer Review Technique Project
HCA –	Healthcare Assistant
IHI –	Institute of Healthcare Improvement
IOM –	Institute of Medicine

ISO –	International Organisation for Standardization
ISQua –	International Society for Quality in Health Care
LSU –	Low secure unit
MeSH –	Medical subject headings
MDT –	Multidisciplinary team
MHA 1983 –	Mental Health Act, 1983
MRT –	Middle-range theory
MSNAP -	Memory Services National Accreditation Programme
MSU –	Medium secure unit
NAPICU –	National association of psychiatric intensive care & low secure units
NCROP –	National COPD Resources and Outcomes Project
NHS –	National Health Service
NICE –	National Institute for Health and Care Excellence
PICU –	Psychiatric Intensive Care Unit
POMH-UK –	The UK Prescribing Observatory for Mental Health
PRISMA –	Preferred Reported Items for Systematic Reviews
PROMs –	Patient-reported outcome measures

QELS -	Quality of Environment in Low Secure Services
QICs -	Quality Improvement Collaboratives
QNFMHs -	Quality Network for Forensic Mental Health Services
TQM -	Total Quality Management
TSHAS -	Trent Small Hospital Accreditation Scheme
UK -	United Kingdom
USA -	United States of America
WHO -	World Health Organisation

## **Chapter 1 Introduction**

In chapters 1 – 5, I provide a foundation for conducting a realist evaluation of external peer review programmes in improving the quality of mental health services. I focus on the requirement for stakeholder views of causal mechanisms and contexts necessary to effect change in external peer review programmes, and the rationale for considering the context of readiness for change.

In chapter 1, I address the background of quality improvement, measures and different types of external peer review. I conclude by providing the rationale of studying peer review networks and accreditation schemes. In chapter 2, I introduce the concept of realist evaluation as a type of programme evaluation that I will use to structure this research. In chapter 3, I position the study within the existing literature, to identify gaps to explain the rationale for the research context and question. To provide context for the research, I offer reasons why I have chosen to study external peer review programmes in the area of inpatient mental health services in chapter 4. I frame the research question and its associated aims in chapter 5.

## 1.1 Overview of quality in healthcare

The pursuit of quality remains at the forefront of healthcare services; especially when faced with increasing financial pressures, competing priorities and concerns about patient safety. Given the suggestion that better quality services will result in improved clinical outcomes, quality improvement in health care has gained considerable attention ([Institute of Medicine, 2001](#)). The ability to measure and report quality improvement in meaningful and useful terms is also growing rapidly ([Leatherman and Sutherland, 2003](#)). There is also an increasing demand to hold specialties, managers and governments accountable.

As health care organisations around the world are recognising the increasing need to demonstrate the quality of their services, landmark articles from the Institute of Medicine's (IOM's) Quality Chasm Series such as *To Err is Human or Crossing the Quality Chasm* report ([Institute of Medicine, 2001](#), [Institute of Medicine, 2000](#), [Institute of Medicine, 2003](#), [Institute of Medicine, 2011](#)) are being cited more frequently in healthcare literature, and scholarship on the topic is greatly advancing. Different countries are making different decisions about which quality improvement processes best suit their needs ([Heidemann, 2000](#)). In response to increasing concerns about quality, many countries are carrying out large scale programmes which include national quality strategies, hospital-wide programmes and external peer review programmes ([Øvretveit and Gustafson, 2004](#)). The United Kingdom (UK) faces an immense challenge of bringing about improvements in quality at a time of growing financial and workload pressures ([Ham et al., 2016](#)).

The case for quality improvement and change management is broadly accepted by authorities, policy makers, and professionals in health care ([Grol et al., 2004](#)). However, while the management of service users is increasingly expected to be evidence-based, measures and programmes designed to change practice have more often been implemented based on firm beliefs.

### 1.1.1 Defining healthcare quality

Developments in healthcare quality have been professionally led which reflects different traditions and ways of working. As a consequence, a variety of healthcare definitions and dimensions of quality have been developed ([Boaden et al., 2008](#)). Healthcare researchers and clinicians both agree that high quality healthcare is fundamental to the protection and effective treatment of service users ([Mason et al., 2015](#)). Despite this, both groups have struggled to agree on a definition and extensive debate surrounds the true meaning of quality in healthcare. In 1990, Avedis Donabedian, whose name is synonymous with quality, outlined the ‘seven pillars of quality’ ([Donabedian, 1990](#)). This was distilled 11 years later in the Institute of Medicine (IOM) report *Crossing the Quality Chasm*, which highlighted six core aims for the 21<sup>st</sup> century health care system: to deliver care that is safe, effective, patient-centred, timely, efficient and equitable. In this landmark report, the IOM defined quality as:

‘the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge’ ([Lohr, 1990](#)).

Defining healthcare quality in broad terms and measuring multiple dimensions are particularly important, given the legitimate variability in perceptions of quality ([Leatherman and Sutherland, 2003](#)). According to Øvretveit, a quality health service is one which organises resources in the most effective way to meet the health needs of those most in need, safely, without waste and within higher-level requirements ([Øvretveit, 2009a](#)). The Darzi NHS (National Health Service) Next Stage Review characterised high quality simply as incorporating the effectiveness and safety of treatment and care alongside a positive experience for those who use services ([Darzi, 2008](#)). Recent government policy in the UK ([Department of Health, 2011](#)) adopted Lord Darzi’s definition of quality in a balanced NHS Outcomes Framework.

The technical field of quality improvement has an emerging shared perspective in which key domains of quality are important to measure, as well as an



increasingly common view of essential data elements, definitions of measures and reporting conventions. Similar frameworks for assessing and reporting quality of care are currently used in countries such as Australia, Canada, England and the United States of America (USA) ([McLoughlin and Leatherman, 2003](#)). These frameworks, although containing slight difference in terminology, generally include measurement in at least five domains: access; effectiveness and appropriateness; responsiveness; safety; and equity ([Leatherman and Sutherland, 2003](#)).

These different definitions may in part relate to the diversity of stakeholders involved in healthcare quality. Service users, commissioners, regulatory bodies, government and healthcare professionals have a stake in the quality of care provided by services. However, their views differ on what is important and how this should be measured ([Mason et al., 2015](#)).

Several authors have described the variety of definitions of quality, and the implications of this, and it is not necessarily the case that some are right and others wrong:

‘several formulations are both possible and legitimate, depending on where we are located in the system of care and on what the nature and extent of our responsibilities are’ ([Donabedian, 1988](#)).

It is argued that variation in the understanding of quality in healthcare and approaches for improving quality, have hampered the progress of quality improvement initiatives ([Øvretveit, 1997a](#)). This prevents different professions and departments from putting their quality improvement efforts into practice using a systems approach ([Boaden et al., 2008](#)).

## **1.2 Evolution of quality improvement**

Quality can be traced back to the advent of medicine itself. Since then, many different approaches to quality improvement have been developed and deployed in healthcare. As far back as Ancient Greece, the challenge in defining quality of care resulted in using lists of attributes or features to assist with conceptualisation. The ancient civilisations of Egypt and Babylon recognised that poor quality of care can lead to harm, and good quality to the absence of harm ([Reerink, 1990](#)).

### **1.2.1 Measuring quality improvement**

Quality improvement in healthcare is a complex, multidimensional concept that extends beyond the delivery of specific evidence-based treatments and interventions, making it difficult to operationalise and measure.

Two measurement approaches dominate the literature: assessing appropriateness of care, and adherence to professional standards ([Institute of Medicine, 2001](#)). This assessment involves creating a list of quality indicators that describe a process of care that should occur for a particular type of service user or clinical circumstance and then evaluating whether care is consistent with indicators. Quality indicators are based on standards of care, which are either found in the research literature and statements of professional medical organisations or determined by an expert panel. Current performance can be compared with reference to a benchmark that establishes a goal ([Institute of Medicine, 2001](#)). Standards address minimal requirements for a healthcare organisation to operate and care for service users ([Shaw, 2004](#)). As they are in the public domain, they can therefore be difficult to update. Adherence to standards can be checked using clinical audits. These are quality improvement cycles which measure the effectiveness of healthcare against agreed standards for high quality ([Jamtvedt et al., 2010](#)). Actions are taken to bring practice in line with standards to eventually improve the quality of care ([Brain et al., 2011](#)). For a long time, clinical audit with feedback to those who provided services was the leading approach to improving quality in many countries, although the

benefits from expenditure on this approach have never been adequately evaluated ([Grol et al., 2004](#)).

Current government policy in the UK seeks, through clinical governance, to integrate the management approach to quality assurance, promoted in the mid-1980s, with the professional approach to clinical audit that was promoted in the early 1990s. Clinical governance implies the comprehensiveness of total quality management but not the dynamic emphasis of quality improvement; although governance originally meant steering, it now implies administration and direction, an echo of quality control ([WHO, 2003](#)).

UK Regulators such as the Care Quality Commission (CQC) and Monitor have been established. The CQC is a regulator with responsibility for both health and social care settings ([Davies and Killaspy, 2015](#)). Despite relatively limited evidence, increased demands have been placed on health services to deliver data on service activity and performance to external bodies (such as the CQC and Monitor). Many of these external bodies require regular 'outcome' reports. Local commissioning bodies also request data on care quality indicators such as commissioning for quality and innovation payment frameworks (CQUINs) to justify continued investment in services ([Lewis and Killaspy, 2014](#)).

Historically, health care has focused on quality assurance, through a system for evaluating the delivery of services or the quality of products, and quality control, a system for verifying and maintaining a desired level of quality. Nevertheless, these methods used alone are not adequate to enhance outcomes. More recently, there has been a growing demand away from quality assurance, towards improvement, and for mechanisms to promote and maintain it.

The first champions of routine clinical outcome measurement were Florence Nightingale (circa 1860) and Ernest Codman (circa 1900). Nightingale pioneered the systematic and rigorous collection of hospital outcomes data to understand and improve performance. Codman advocated the 'end results idea', essentially following every service user treated for long enough to determine

whether treatment was successful, and if not to understand and learn from failures. However, political and practical barriers prevented these ideas from complete adoption until the last 25 years.

Systematic measurements of quality typically follow Donabedian's ([1980](#)) or Maxwell's ([1984](#)) broader approaches. Donabedian's ([1988](#)) classic paradigm for assessing quality of care is based on structure, process and outcomes, with each component having an effect on the next. He argued that:

‘good structure increases the likelihood of good process, and good process increases the likelihood of good outcome’ ([Donabedian, 1988](#)).

Structural quality evaluates health system capacities, process quality assesses interactions between clinicians and service users, and outcomes offer evidence about changes in service users' health status ([Institute of Medicine, 2001](#)). It is argued that the best process measures are those founded on research evidence that better processes lead to better outcomes. Similarly, the best outcome measures are those tied to processes of care. All three dimensions can provide valuable information for measuring quality, but the majority of the literature focuses on measuring processes of care ([Institute of Medicine, 2001](#)).

These three types of measures have their unique strengths, but each is also associated with conceptual, methodological and practical problems ([Raleigh and Foot, 2010](#)).

It should also be noted that good performance on indicators does not necessarily equate to good care at the individual service user level ([Raleigh and Foot, 2010](#)). Measurement can often miss areas where an evidence base or data are not available, or perhaps aspects of quality are not readily quantifiable. Therefore, it is important that quality measurement is used as one tool in a broader quality improvement strategy ([Werner and Asch, 2007](#)). It is also important to note that suboptimal performance on quality measures does not

necessarily mean deficient performance on the part of the health care provider or health care system ([Leatherman and Sutherland, 2003](#)).

Most studies to date have used their own definitions and measures of effects of quality improvement programmes. This limits comparisons between different studies, which are necessary to build a body of knowledge.

Although clinical outcome measures are often considered the gold standard for measuring effectiveness in health care, their use can be problematic. For example, outcomes cannot always be assessed in a timely or feasible fashion, and it can be difficult trying to understand the contribution of health services to health outcomes. Thus, measures of process are often used instead of, or in addition to, measures of outcome. Process measures have certain advantages, for example, they are quicker to measure, and easier to attribute directly to health service efforts ([Brook et al., 1996](#)). In addition, they can be considered by some as a better measure of quality as they examine compliance with what is perceived as best practice. However, process measures could be considered less value for service users unless related to outcomes. Moreover, they may ultimately ignore the effectiveness or appropriateness of the intervention and pre-judge the nature of the response to a health problem, which may not be identical in all settings, such as for patients who have multiple morbidities ([Klazinga et al., 2011](#)).

#### **1.2.1.1 Structure measures**

Structure can be defined as the relatively stable characteristics of the providers of care, of the tools and resources they have at their disposal and of the physical and organisational settings in which they work. Structure embraces the number, distribution, and qualifications of staff, as well as the number, size and equipment, and geographic disposition of the hospitals and other facilities. However, the concept also extends beyond production to include the ways in which the financing and delivery of health services are organised, both formally and informally. The way healthcare professionals conduct their work, in individual practice or in groups is an aspect of structure. Structure includes the

organisation of the staff in healthcare services, and the presence or absence of quality assurance ([Leatherman and Sutherland, 2003](#)).

Structural characteristics of the settings in which care takes place have a propensity to influence the process of care so that its quality is diminished or enhanced ([Donabedian, 1980](#)).

#### **1.2.1.2 Process measures**

Processes are a measurable set of activities that go on within and between healthcare professionals and service users, these activities can be referred to as the 'process' of care ([Donabedian, 1980](#)).

Process measures, are felt by some to be more responsive than outcome measures as they avoid the effects of time lags, reflect the quality of health care more accurately, and measure the care that service users receive ([Raleigh and Foot, 2010](#)). Some feel they are less contentious in terms of attribution, less susceptible to risk-adjustment issues, less prone to small-number problems, and are more actionable as they are direct measures of processes of care. However, compared with outcome measures, they may be less easy to produce in terms of data availability and more liable to gaming and the extent to which they link with outcomes is also variable ([Raleigh and Foot, 2010](#)).

Process measures may also focus on the overutilisation and underutilisation of health care services, and variation in care that is unjustified by the medical needs of service users. In England, the National Service Frameworks provide useful templates for defining standards of care that could serve as process measures.

#### **1.2.1.3 Outcome measures**

An outcome can be defined as a result or a visible effect ([Lewis and Killaspy, 2014](#)). Outcome measures include changes in current and future health status (such as mortality), functional status, health-related quality of life and satisfaction with care ([Leatherman and Sutherland, 2003](#)). The most typical

outcome measures in healthcare are Patient Reported Outcome Measures (PROMs), survival rates, complication rates, measures of clinical improvement, and performance indicators ([Raleigh and Foot, 2010](#)) ([Department of Health, 2008](#)).

Outcome indicators are intuitively attractive for some, as they are direct measures of outcomes for service users, hence they have face validity, are easy to understand, and data are more readily available. However, their interpretation is much more likely to be confounded by the effects of time lags, co-morbidities, and issues of causation and attribution. They are also less informative as pointers for remedial action, and may be subject to small-number problems. They should, therefore, form part of a balanced set of measures reflecting on quality ([Donabedian, 1966](#)).

### 1.2.2 Defining quality improvement

According to Øvretveit ([2009a](#)) ([2009b](#)), quality improvement can be defined as any change that improves quality, or more narrowly a change made using a specific approach:

1. **an improvement to quality:** in service user experience or clinical outcome
2. **a quality improvement change:** a change which results in an improvement to quality
3. **any method for making a quality improvement change:** such as training, merging two services, increasing staff, setting standards and inspecting or giving feedback on performance
4. **a specific method for making a quality improvement change:** such as a continuous quality improvement tool
5. **an approach for improving quality:** using a collection of ideas and tools, and which aims to develop organisations' capacity to use these tools, such as patient pathway flow, re-engineering, Six Sigma, Lean or Total Quality Management
6. **a change that raises quality and lowers costs:** value improvement

A quality improvement programme consists of planned activities carried out by an organisation, for teams from many organisations to improve quality ([Øvretveit and Gustafson, 2004](#)). Quality improvement is made up of two components: change and method ([Atkinson et al., 2010](#)). Quality improvement

principles, tools and techniques are largely based on organisational and change approaches developed in industry, such as Total Quality Management (TQM) or Continuous Quality Improvement (CQI) ([O'Neill et al., 2011](#)). Common tools include Plan-Do-Study-Act (PDSA) cycles, Lean, Six Sigma, Root Cause Analysis, trigger tool and scientific evaluative measures such as Statistical Process Control (SPC) or time series. Quality improvement programmes cover a range of activities which are more complex than a single quality team improvement project or the quality activities in one department ([Øvretveit and Gustafson, 2004](#)). Many programmes are evolving, and involve a different number of activities which start and finish at different times ([Øvretveit and Gustafson, 2002](#)). These activities may be mutually reinforcing and can have a synergistic effect if properly implemented. Some programmes are implemented over a long period of time; many cannot be standardised and need to be changed to suit the situation in ways which are different from when a treatment is changed to suit a service user. As the targets of programmes are not directly service users, but whole organisations or social groups which vary more than the physiology of an individual patient: they can be considered as complex adaptive social systems ([Plesk and Greenhalgh, 2001](#)).

Programmes aimed at monitoring and improving the quality and safety of patient care are in place in many countries ([Grol et al., 2004](#)). These approaches differ in perspective. Some focus on changing professionals' behaviour, others on changing organisations or on changing the interactions between participants in the system. Some emphasise the importance on self-regulation in changing care, others believe in the power of external control and incentives. Some prefer bottom-up, others top-down methods for changing practice.

Quality improvement becomes a continuous journey where standards may be progressively raised, even radically reformulated, as people increase their expectations about what might be possible ([Davies and Killaspy, 2015](#)). Raising standards through quality improvement programmes depends critically on staff being prepared to share their ideas and experiences, including those that might be characterised as 'failures'. This culture has been established in the air



transport industry where there is a general recognition of the value of detailed analysis of crashes or 'near-misses' using an analytical framework (root cause analysis) that aims to reveal both human error and system failure ([Nicolay et al., 2012](#)). As noted by many commentators recently, this is not the culture in health services at present. Many staff are frightened of discussing incidents openly because they assume that when it comes to apportioning 'blame' the organisation will err on the side of attributing human error, rather than system failure ([Davies and Killaspy, 2015](#)).

### **1.2.3 Internal quality improvement**

Internal quality improvement programmes such as implicit or case review are used to evaluate the quality and appropriateness of inpatient care by reviewing performance or medical records ([Hayward et al., 1993](#)). The term external implies that the quality improvement is led by an organisation separate from that which actually provides care ([Wareham, 1993](#)). External standards have been developed by a body external to the organisation being inspected, compared to standards that are used in audit and feedback, which are often set by the group whom they are applied ([Flodgren et al., 2011](#)). Sometimes this can be facilitated by having access to a respected figure, from outside the organisation, who can provide expert opinion and a degree of objectivity. For example: in the case of Professor Tim Briggs, the pioneer of the 'Getting It Right First Time' project, aiming to improve quality by reducing variation and cost in orthopaedic services ([Briggs, 2015](#)). Teams can be involved in external reviews of organisations in an area, to change practice across many organisations, or for a national or regional quality strategy. Often it is sufficient simply to give staff the time to discuss the problem and come up with their own solutions, or arrange for them to meet another staff team who have been grappling with the same problems ([Davies and Killaspy, 2015](#)). Perhaps the most potent element of a quality improvement programme is the reflection by clinical teams on their performance ([Barnes and Paton, 2012](#)). The major distinction is with internal approaches that are undertaken directly by the organisation that provides care.

Based on current limited evidence, none of the approaches to quality improvement can be regarded as superior; we might need all of them to be successful in achieving quality in health care. Furthermore, it is not yet clear which approach works best for what type of service in a given context ([Grol et al., 2004](#)).

### 1.3 External quality improvement approaches

External approaches are widely used in many countries' health care systems, and have a long and varied history ([Brennan and Berwick, 1996](#)). External approaches are widely used in the UK as a mechanism for quality improvement ([Walshe et al., 2001](#)). External review systems are used in health care to promote improvements in the quality of care, promoting changes in organisational structures, processes or healthcare provider behaviour ([Flodgren et al., 2011](#)). These review systems assume that externally promoted adherence to evidence-based standards through inspection and assessment will result in a higher quality of care. Review systems are popular among healthcare funders, who are more likely to make funding available (or less likely to withdraw funding) if standards are met and healthcare professionals and the public can have confidence in the standards of care provided ([Flodgren et al., 2011](#)). There are numerous external review systems such as: audit, regulation and statutory inspection, quality circles, quality improvement collaboratives, and different types of external peer review ([Shaw, 2004](#)). External quality improvement programmes play an important role in assessing and improving health care quality ([Kilsdonk et al., 2015](#)). They can be used to regulate, improve and market healthcare organisations. External approaches can also be used to respond to the increasing demands for public accountability, clinical effectiveness and improvement of quality and patient-safety ([Shaw, 2004](#)). The uptake and success of external quality systems in individual countries is closely connected to the social, political and economic climate that determines incentives for participation ([Shaw, 2000](#)). But despite their widespread use, their impact on the organisations that use them is not well understood ([Walshe, 1999](#)).

A number of different and overlapping terms are used to describe different forms of external approaches to quality improvement ([Shaw, 2000](#)). Just as there is great diversity in the terminology employed, there are also great differences between the agencies which undertake these activities, and between the approaches and methods used ([Klazinga, 2000](#)). Some external approaches

have a statutory basis in law and are effectively mandatory, while others are voluntary. Some approaches are undertaken by independent or professional organisations while others are led by government agencies; some are confidential while others are entirely open to public scrutiny. Some measure systems for quality improvement while others measure the quality of care directly; some make use of formal and explicit standards and measurements while others rely heavily on subjective reviewer judgements; and some result in little or no formal action while others rely heavily on subjective reviewer judgements or are linked to important financial or non-financial incentives and sanctions ([Walshe et al., 2001](#)).

There is considerable variation in standards and external assessment processes for different disciplines and countries. Many national programmes, especially within Europe, have agreed in principle to voluntary convergence of standards and assessment processes according to the ALPHA (Agenda for Leadership in Programmes for Healthcare Accreditation) Principles of the International Society for Quality in Health Care (ISQua). ISQua has developed general principles for standards based on analysis of existing accreditation programmes ([Heidemann, 2000](#)) to make them more reliable, valid and compatible within and between countries.

Quality circles and Breakthrough Series Quality Improvement Collaboratives (QICs) are considered by some as external peer review programmes. Commencing in The Netherlands (1979) and in other European countries in the 1980s and early 1990s, quality circles have become an important method of quality improvement ([Beyer et al., 2003](#)). Initially in ambulatory care in Germany at the end of the 1980s ([Grol, 1994](#)), they consist of a group of seven to ten physicians or interdisciplinary groups with other health professionals from a unit (or across units) who have volunteered to meet together regularly to analyse and make proposals about quality ([French and Bell, 1998](#)). They undertake activities aimed at assessing and continuously improving the quality of care. However, as quality circles are used more predominantly in continual

medical education or professional self-awareness, I did not consider them as an approach to external peer review in this thesis.

A great deal of evaluation research has been undertaken into QICs. They bring together groups of healthcare practitioners from different healthcare organisations to work in a structured manner to improve one aspect of the quality of their service through external review practices. QICs consist of a series of meetings to learn about best practice in a specific area. The approach is used to share experiences of making changes in local settings ([Duckers et al., 2009b](#)), ([Øvretveit, 2002b](#)). I did not consider QICs as an external peer review in this thesis as they did not systematically use external peer reviewers or conduct site visits.

External reviewers can be anyone from outside of an organisation; but using external peer reviewers involve peers from a similar-sized hospital of a similar type from outside an organisation. Some argue that this has the potential to remove or reduce bias ([Taubman, 1989](#)). Berwick ([1990](#))echoed this and stated that without the wisdom of peers in judging the quality of care and service, quality improvement programmes would be ‘severely retarded’. Multidisciplinary teams can offer diverse perspectives and inputs that can produce a more comprehensive view than a single reviewer. A multidisciplinary review team comprising peers from an outside unit or organisation who can conduct an independent, objective, valid, and unbiased review is suggested as essential to identify and mitigate risks ([Hudson et al., 2012](#)). Moreover, it has been argued by Walshe and Shortell ([2004](#)) that external peer reviewers can objectively see a culture that may be invisible to internal reviewers.

## 1.4 External peer review

In many health care systems, approaches to quality improvement are frequently promoted using external peer review ([Scrivens, 1997](#)). According to Berwick ([1990](#)), a clear universal definition for external peer review is difficult to find. He proposed that it be defined as ‘inspection and evaluation of health care structures, practices or results’. The key features of external peer review are involvement of professionals, an agenda for inspection and judgement of care and its components. External peer review of clinical departments in Europe can be defined as: standards based on on-site surveys conducted by health care professionals to assess the organisation of the care processes and its results aimed at improving the quality of care ([van Weert, 2000](#)).

A peer is a person who is equal in any stated respect ([Irvine and Irvine, 1991](#)). Usually in the same branch of health care provision, with comparable experience or training, peers may be colleagues of the same or different disciplines working together in a practice or hospital unit or care providers working together in local or regional unidisciplinary or multidisciplinary groups ([Grol, 1994](#)). Much of the success of the review depends on the expertise and attitude of the reviewers. It has been argued that they should have an up-to-date knowledge of the professional standards, be open-minded, have a broad view and have a constructive attitude ([van Weert, 2000](#)).

The External Peer Review Techniques (ExPeRT) project, funded by the European Union ([Shaw, 2000](#)), identified four systematic approaches of external peer review techniques that link national and international standards to local practice and have been applied to private or public health care ([Klazinga, 2000](#)), ([Heaton, 2000](#)). These approaches have been systematically compared in a number of studies of standards and methods used by programmes based on industry and health care ([Klazinga, 2000](#)), ([Bohigas and Heaton, 2000](#)), ([Donahue and van Ostenberg, 2000](#)). Standards and working methods were compared during a further analysis of these models ([Bohigas and Heaton, 2000](#)), ([Heaton, 2000](#)).

The main types of external peer review are European Foundation for Quality Management (EFQM), International Organisation for Standardization (ISO), *visitatie* (referred to as peer review networks in this thesis) and accreditation programmes ([Heaton, 2000](#)), ([Klazinga, 2000](#)).

The Balridge Awards were developed in the USA for improvement of quality in production industries ([Nabitz et al., 2000](#)). They inspired the European Foundation for Quality Management (EFQM) and the 'business excellence' model. Healthcare providers who seek voluntary development of a European Quality Award are assessed against performance standards for service industries in specific areas such as clinical results, service user satisfaction, administration and staff management ([Shaw, 2000](#)). The Balridge criteria for management systems have evolved into national and international assessment programmes such as those in Australia and Europe ([Nabitz et al., 2000](#)). Health care providers could voluntarily assess themselves or be externally assessed against performance standards. The EFQM model also provides a transparent framework to map organisational standards. Several countries, particularly those in Scandinavia, have introduced their own national quality awards based on the European framework ([WHO, 2003](#)).

The ISO developed a series of standards (ISO 9000) originally designed for the manufacturing industry (e.g. medicines and medical devices) ([WHO, 2003](#)). These are now used to assess quality systems in specific aspects of health services, whole hospitals and clinics. Most commonly, parts of hospitals are assessed by independent auditors (themselves regulated by a national 'accreditation' agency). Performance is measured in terms of compliance with ISO standards. Certification is widely available from independent certified auditors and is recognised across national borders ([Shaw, 2004](#)). ISO 9000 standards initially related mostly to administrative procedures rather than to hospital performance, and terminology was often difficult to relate to healthcare, leading to varying interpretations between national agencies ([Sweeney and Heaton, 2000](#)). However, when adapted in 2000, the ISO 9000

series of standards became more easily applied to healthcare and now include the assessment of outcomes and consumer satisfaction ([Shaw, 2004](#)).

Peer review networks are synonymous with the *visitatie* model that originated in The Netherlands. *Visitatie* is a Dutch word meaning 'to visit' ([Lombarts and B., 2003](#)), ([Klazinga et al., 1998](#)). Undertaken by two or more healthcare providers, *visitatie* involves collaboration amongst professional peers whom provide evaluation and support to each other. Although widely used in the selection and monitoring of specialty medical training, it has also been developed to focus on clinical practice, continuing professional development (CPD) and service quality. This includes programmes aimed at the recognition of specialty training (e.g. certification) or assessment of clinical practice.

Peer review networks, to a varying degree, are independent and use explicit standards to combine internal self-review with visits, surveys and assessments ([Shaw, 2001](#)), driven by professional and often multidisciplinary organisations ([Shaw, 2004](#)). Standards tend to be derived implicitly from practice guidelines and personal experience ([Shaw, 2000](#)). They are often specialty-based, not covering whole hospitals, and results are confidential and not publicly available ([Shaw, 2004](#)); ([van Weert, 2000](#)). They focus on improvement of care and on exchanging ideas. Several specialties are now introducing peer review visits, enabling staff to share and exchange ideas on best clinical and organisational practice ([Page and Harrison, 1997](#)). To acquire staff support and active participation, the reports and recommendations resulting from a peer review visit are confidential to the host organisations. The peer review visit takes place periodically and is a formalised event: it is planned and organised in line with explicit predetermined procedures, questionnaires and standards for reports. Information is obtained from documentation, observation, staff, and patient information. Standards of good quality care or best practice are used when available. Where bodies have not yet produced standards, relevant guidelines are developed during the course of the programme. The scope of the review is usually the care process and its organisational aspects ([van Weert, 2000](#)).



When this approach leads to a pass or fail score which indicates the degree of compliance with standards, it is referred to as accreditation ([Scrivens, 1998](#)). This is distinct from appraisal, a confidential process in which individuals' professional and performance development and job progress are reviewed against agreed objectives at regular intervals by an educational supervisor or clinical manager.

External peer review standards increasingly emphasise the clinician-management interface, evidence-based medicine and the continuum of care as seen by service users rather than managers ([Shaw, 2000](#)). Existing users, most of whom are not related to healthcare, drive the development of ISO and EFQM. ISO certification and the EFQM excellence model focus primarily on the managerial and organisational conditions under which care processes are executed. Peer review networks and accreditation are the closest to the actual deliverance of healthcare ([Klazinga, 2000](#)). Accreditation and peer review networks both use health care professionals as assessors, whereas ISO and EFQM assessors have more global backgrounds and expertise ([Heidemann, 2000](#)). Peer review networks and accreditation schemes are the two most commonly used approaches in healthcare in the UK ([Shah et al., 2010](#)). Thus, for the purpose of this thesis I am just going to focus on two approaches to external peer review: accreditation and peer review networks. There is a clear distinction in the literature between peer review networks and accreditation, where the latter is distinguished by a pass or fail award or certificate ([Kilsdonk et al., 2015](#)).

A preliminary observation was that in countries where both accreditation and peer review networks exist there is convergence rather than divergence ([Klazinga, 2000](#)). The UK Accreditation Forum was set up in 1998 to support external peer review programmes. The Academy of Medical Royal Colleges is also currently working towards more coherent procedures, but neither body has the resources or the authority to standardise standards or regulate peer reviewers across the country ([Shaw, 2001](#)).

### 1.4.1 Peer review networks

A network can be defined as: a cooperative structure where interconnected groups or individuals coalesce around a shared purpose on the basis of trust and reciprocity ([Randall, 2013](#)). Table 1.1 illustrates the distinctive features of networks. Evidence from a recent systematic review ([Cunningham et al., 2011](#)) demonstrates that creative, cohesive, collaborative networks can improve coordination of care whilst attending to quality and safety issues and agendas.

**Table 1.1 Distinctive features of networks (adapted from ([Malby and Mervyn, 2012](#)))**

<b>Distinctive features</b>	<b>Explanations</b>
<b>Diversity</b>	Network membership is diverse and is collectively able to innovate and be creative
<b>Distributed leadership</b>	Power and leadership is distributed across network members
<b>Reciprocity</b>	Relationships between network members are defined by reciprocity and exchange
<b>Common purpose</b>	Network members have a mutual interest in a common purpose
<b>Instability</b>	Members' commitment, engagement and impact fluctuates
<b>Adaptability</b>	Networks are able to adapt to survive and thrive
<b>Knowledge</b>	The knowledge function is central to the networks identity and mission

Peer review networks provide inbuilt mechanisms to spread successful change quickly, leveraging the power of social and professional connections, rather than relying on the formal chain of command of a hierarchical organisation ([The Health Foundation, 2014](#)). Some drive change across organisations; others simply unite individuals with common interests. Bradley et al. ([2009](#)) describe organisations that demonstrate consistently exceptional behaviour as 'positive deviants'. This enables other organisations in the peer review network to inspire change through their sustained and supported connection to one

another, through shared learning and consequent enhanced capacities ([Casebeer et al., 2009](#)). Peer review networks are growing in number and importance in UK healthcare. A learning report from the Health Foundation ([2014](#)) has suggested that networks contribute to healthcare improvement by providing a forum for experimentation and creating knowledge, exchanging information and spreading good practice. Some may argue that they are well positioned to tackle systemic and complex problems faced by commissioners, providers, regulators, frontline staff and service users as they provide an opportunity to focus directly and exclusively on quality improvement ([The Health Foundation, 2014](#)). Networks can provide a neutral environment where individuals from different organisations, disciplines and constituencies can collaborate on an equal footing, freed from the constraints and competition created by more hierarchical structures.

A recent review identified five core features of effective networks: common purpose, cooperative structure, critical mass, collective intelligence and community building ([The Health Foundation, 2014](#)). These features are interdependent, and interact to give the network energy and momentum. They ensure a clear direction, credibility and increased scale and reach, while enhancing knowledge, encouraging innovation and creating meaningful relationships. All five core features are mutually reinforcing, and their combined effect enables quality improvement, learning and change to happen ([The Health Foundation, 2014](#)).

Peer review networks (*visitatie*) have a common methodology, but have shown less consistency in other countries compared to the Netherlands ([Shaw, 2000](#)). Prior to the visit, a questionnaire with questions about practice organisation and care processes is completed. During the visit, which typically lasts one full day, interviews are held with staff from specialties involved in service user care ([Schulpen and Lombarts, 2007](#)). Casebeer et al. ([2009](#)) suggests collaboration through networks can emerge when such democratic and shared action is different from the prevailing organisational culture. An organisation can be categorised by two defining elements: its structure and culture. Culture refers to

those norms and values that characterise the operations of the organisation. A key aspect of organisational culture is readiness to change. This involves being willing to challenge the status quo, to examine the environment for new knowledge, and to accept the need for continuous improvement ([Owen, 2005b](#)).

Findings from visits are documented in a confidential report that concludes with recommendations for improvement ([Lombarts and B., 2003](#)). Implementation of the recommendations is left to the healthcare staff who underwent the review. Although there are no formal sanctions for non-compliance with the recommendations, specialty societies expect that their members will act upon the recommendations and implement the suggested changes.

Site visits are an important feature of a peer review network. The basis of assessment is primarily clinical, confidential and less explicit than the standards and processes of other models ([WHO, 2003](#)), ([van Weert, 2000](#)). A number of specialty societies have made peer review networks a mandatory activity for their members; others have opted for voluntary participation, hoping all colleagues will be surveyed once every five years through 'peer pressure' ([Lombarts and B., 2003](#)). Over the past decade, all Dutch specialist groups have undergone one or more collegial quality surveys through *visitatie*. Survey teams consist of three practising specialists from out-of-region hospitals, who have been trained in *visitatie* procedures, quality improvement approaches and interviewing skills ([Schulpen and Lombarts, 2007](#)). In most of the Dutch peer review networks (*visitatie*), the selection criteria for the reviewers are more elaborate than the UK. In addition, the written report is first discussed in a plenary committee to ensure that the conclusions and recommendations stated by the reviewers are in line with other reports ([van Weert, 2000](#)).

#### **1.4.2 Accreditation**

Accreditation originated in the United States in 1917 as a mechanism for recognition of surgical training posts ([Klazinga, 2000](#)) ([Braithwaite et al., 2006](#)); ([WHO, 2003](#)), ([Viswanathan and Salmon, 2000](#)). It can be defined as 'a means of

publicly recognising that a healthcare organisation meets predetermined national standards of operation' ([Pomey et al., 2005](#)), ([Scrivens et al., 1995](#)). As with quality improvement generally, accreditation traditionally developed in hospitals and then moved inwards towards clinical specialities ([Shaw, 2003](#)) and outwards towards community services and then to networks of preventative and curative services ([Shaw, 2004](#)).

Accreditation is a formal process by which a recognised body assesses and recognises that a health care organisation meets applicable pre-determined and published standards. Programmes are available in many countries by independent agencies that may be themselves accredited under international standards (such as ALPHA principles and standards). Through their activities and programmes accrediting agencies: set or monitor standards, using external peer reviewers that require health care organisations to gather, review, and analyse organisational and clinical performance data ([Braithwaite et al., 2010](#)). They assess and benchmark continuous quality improvement performance of health care organisations against industry standards and peer organisations in cycles ([Braithwaite et al., 2010](#)). Accrediting bodies encourage and support the spread of best practices; provide candid assessment reports to healthcare organisations with recommendations for improvements. They note positive and outstanding practices; recruit, train, and manage reviewer workforces comprising multidisciplinary external peers; and coordinate external peer reviewers to conduct on-site assessments of healthcare organisations using observations, interviews ([Greenfield et al., 2012](#)). In most models, organisations can be accredited, or be granted time to improve following remedial recommendations, or if performance falls below stipulated standards, they can lose their accreditation status ([Braithwaite et al., 2010](#)).

Reviews must be nonpunitive and separate from a regulatory body that could impose penalties to facilitate the open sharing of information in a transparent process without fear of reprimand, sanction, personal disparagement, or financial risk ([Hudson et al., 2012](#)). Both internal and external parties must mutually benefit from the review by learning from each other and cross-sharing

practices ([Hudson et al., 2012](#)). Names of accredited hospitals are generally published on individual programme websites but details of survey results are not publicly available, except for governmental programmes.

Accreditation is usually a voluntary process in which organisations choose to participate rather than one required by law and regulation. However, the 'voluntary' nature of accreditation has been changing over the last decade, and a number of countries now have mandatory accreditation ([Purvis et al., 2010](#)). Because criteria and assessment processes vary between programmes, accreditation status cannot be assumed to be directly comparable between other countries ([Shaw, 2004](#)). Half of the programmes, especially during the past ten years, have been funded or managed directly by national governments which use them more as a means of regulation and public accountability, rather than voluntary self-development ([Shaw, 2004](#)). Hudson et al. ([2012](#)) propose that the unit or organisation should voluntarily request a review to ensure that the organisation is motivated and engaged in the assessments.

Accreditation standards are regarded as optimal and achievable. The traditional accreditation model is when a variety of structural, process and outcome measures are developed for the health care facilities, both at the departmental level as well as with several quality assurance processes. The focus under this model is on inputs and processes ([Purvis et al., 2010](#)). The measurement process includes testing of internal systems to improve service user orientation, clinical process, outcome and organisational performance ([Shaw, 2004](#)). Improvement gradients are embedded in the process as standards are revised, and raised, over time ([ACHS, 2002](#)).

There is ample evidence in the literature about how hospitals undergoing accreditation rapidly increase their compliance with published standards in the months prior to external assessment, and improve organisational processes as a 'one-off focussed activity'. But there is less evidence that this brings long-lasting impact in terms of continuous quality improvement ([Keeler et al., 1992](#)).

A key study that evaluated accreditation was the Quality Assurance Programme randomised controlled trial. The Council for Health Service Accreditation of Southern Africa (COHSASA) offers a programme of graded recognition and facilitated accreditation which is designed to narrow the wide gap between rich and poor facilities ([Salmon et al., 2003](#)). 20 randomly selected public hospitals, stratified by size were selected. Randomisation of this important trial controlled for important contexts known to affect a hospital's ability to comply with and achieve accreditation standards for indicators such as hospital size, staffing levels, staff qualifications and budget levels. Apart from nurse perceptions of clinical quality, the independent research team observed little or no effect of accreditation on seven of the other indicators of quality, including patients' satisfaction, which did not improve among accredited hospitals compared to non-accredited hospitals.

Some studies suggest that hospitals might adopt opportunistic behaviours solely with the aim of gaining accreditation particularly when governments link accreditation with other objectives such as payment mechanisms, resource allocation, and so forth ([Pomey et al., 2005](#)). Scrivens ([1995](#)) argues that surveys permit snapshots in time; they do not monitor the continuous delivery of health care. In the long period between surveys, compliance with standards may deteriorate ([El-Jardali, 2007](#)). However, according to the study by McGurrin and Hadley ([1991](#)), hospitals were seen to strive to maintain or improve their accreditation status.

Furthermore, concerns have been raised on whether accreditation can only result in organisational changes regarding standardising the organisation and decision-making processes for care rather than improving quality of care as aimed ([Suñol et al., 2009](#)). This may be because the relationships between structure, process and outcome can be complex and ambiguous, and compliance to the accreditation standards of structure and process may not necessarily improve the outcomes. Although many countries are embarking on accreditation programmes with the aims of improving quality and ensuring the best use of resources, there is a shortage of evidence to show the effectiveness

of these programmes and identify the contexts that affect successful implementation ([WHO, 2003](#), [Øvretveit and Gustafson, 2003](#), [El-Jardali, 2007](#), [Shaw, 2001](#)).



## 1.5 Statement of the problem

For some quality improvement programmes, especially hospital programmes, there is no shortage of theories about implementation and the conditions needed for success, but few are empirically based ([Øvretveit and Gustafson, 2004](#)). For external peer review, there is very little theory which highlights a clear need for more evaluations, and focus on identifying how or why a quality improvement programme works ([Krein et al., 2010](#)), ([Bate, 2014](#)). In recent years, a dramatic international growth in external peer review has been observed ([Heidemann, 2000](#)) ([Kilsdonk et al., 2015](#)), despite scarce and inconsistent evidence which informs implementation ([Shaw, 2001](#)). Ample evidence exists demonstrating services rapidly increase compliance with published standards following membership to external peer review, but there is not enough evidence on clinical processes and outcomes ([Salmon et al., 2003](#)). Multiple studies have been performed internationally, but findings have revealed complicated relationships, leading to hesitation by authors to make strong claims about effects due to study limitations ([Greenfield and Braithwaite, 2008](#)) ([Hinchcliff et al., 2012](#)) ([van Harten et al., 2000](#)) ([Øvretveit and Gustafson, 2002](#)) ([Øvretveit and Gustafson, 2003](#)) ([Øvretveit, 2002a](#)). There are relatively few experimental studies ([Perneger, 2006](#)), and those which do exist often show weak or moderate effects at best ([Walshe, 2007](#)). Randomised controlled trials can only provide partial answers, while ignoring basic questions about critical success factors in change processes ([Grol et al., 2004](#)). External peer review programmes are complex adaptive social systems ([Plesk and Greenhalgh, 2001](#)) that cannot be easily controlled, randomised or matched in the same way as can patients ([Øvretveit, 2002a](#)). The evidence for the impact of external peer review is still inconclusive ([Kilsdonk et al., 2015](#)); with more evidence on accreditation programmes ([Alkhenizan and Shaw, 2011](#)) than peer review networks.

External peer review programmes with the same name can be implemented differently at different rates, coverage and depth ([Øvretveit and Gustafson, 2004](#)). Thus, implementation should not be assumed and more data is needed to

assess implementation to help explain outcomes and provide evidence of exactly what changes are made and when. Outcomes should be viewed in relation to how deeply and broadly programmes are implemented, and their stage or 'maturity' ([Hackman and Wageman, 1995](#)).

There is limited understanding of why some external peer review programmes are successful and others are not; or why a programme works in one context but not in another ([Robert et al., 2008](#)), ([Krein et al., 2010](#)). There is evidence that certain contexts appear to be necessary to motivate and sustain implementation, and create conditions likely to produce results ([Øvretveit and Gustafson, 2002](#)). Better understanding of context could help the transferability of quality approaches and methods ([Øvretveit, 1997b](#)). The need for mixed methods is suggested ([Kilsdonk et al., 2015](#)), especially longitudinal studies. Qualitative research can also help understanding at the level of the system and the team ([Caldwell and Mays, 2012](#), [Pope et al., 2006](#)). It can also identify potential enabling contexts, and inform the development of customised change programmes ([Barnes and Paton, 2012](#)).

## Chapter 2 Realist Evaluation

In this chapter, I discuss the structure of realist evaluation as used in this thesis, and I provide the definitions of mechanisms, outcomes (change), context and context-mechanism-outcome (CMOC) configurations. I offer a specific focus on readiness and resistance to change, and conclude by explaining why this research is a 'black box' of evaluation.

Programme evaluation is central to health services research as it aims to record what changes occur, but also what led to those changes. An applied inquiry involves the synthesis of evidence that culminates conclusions about the state of affairs, value, merit, worth or significance of a programme. Conclusions can encompass both an empirical aspect (that something is the case) and a nominative aspect (judgement about value). This value judgement distinguishes evaluation from other types of scientific inquiry ([Fournier, 2005](#)). Scriven ([1967](#)) made a crucial contribution to programme evaluation by distinguishing between formative and summative. Formative evaluation fosters improvement of ongoing activities; whereby summative evaluation is used to assess whether results have met the stated goals. I undertook formative evaluation, as it involves the collection of data while the programme is still active, with the aim of developing or improving it ([Bowling, 2014](#)).

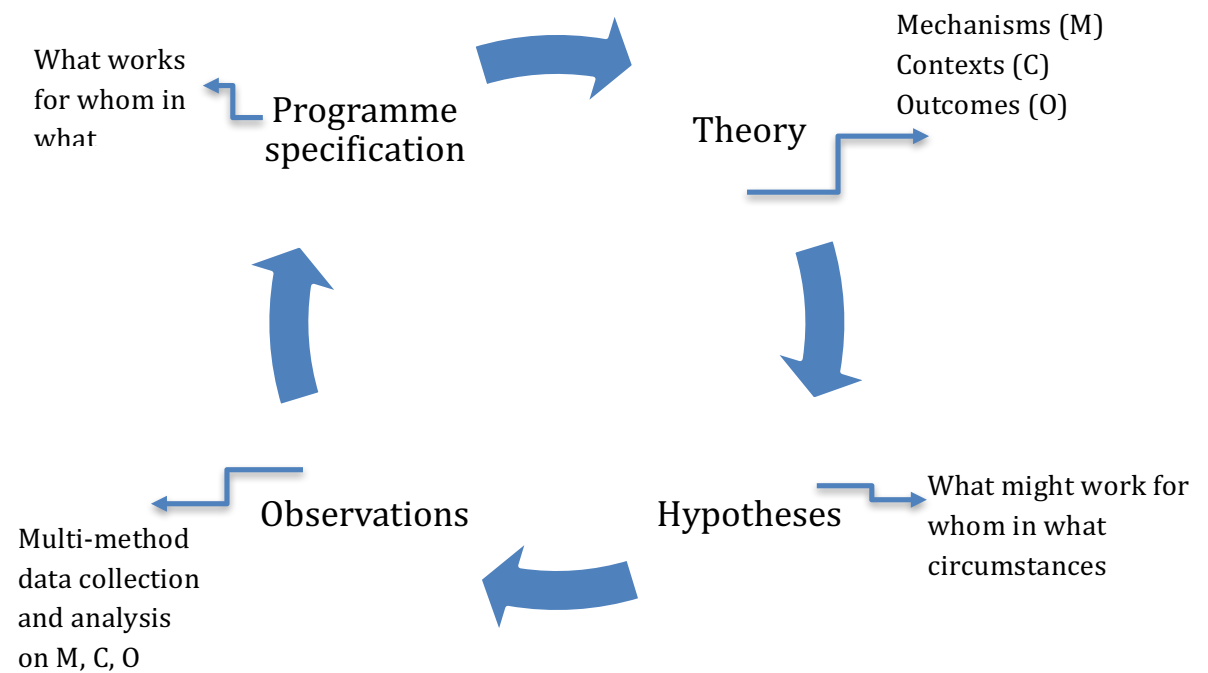
Reasons for the current lack of evaluation research in external peer review include the methodological challenges of measuring outcomes and attributing causality to complex, changing, long term social programmes ([Øvretveit and Gustafson, 2002](#)) that are context specific ([Craig et al., 2008](#)). Their effectiveness appears to vary across different clinical problems, contexts and organisations, presumably due to the presence of different barriers and facilitators ([Guo et al., 2015](#)). I used realist evaluation principles to overcome these challenges.

Realist evaluation is a type of theory-driven evaluation of social programmes ([Pawson and Tilley, 1997](#)). It is an interpretive orientation, based on critical

realism towards research methods, situated between positivism and relativism ([Bowling, 2014](#)). This approach was developed in response to recent interest in understanding how social programmes work rather than investigating their effectiveness ([Pawson and Tilley, 1997](#)); ([McEvoy and Richards, 2003](#)); ([Pawson, 2006](#)). Although Pawson and Tilley ([1997](#)) initially used the term realistic evaluation, the term 'realist evaluation' is now widely used by other authors and has been switched ([Pawson, 2006](#)). Pawson and Tilley ([1997](#)) argued that programmes are complex and applied to heterogeneous subjects and circumstances; this is often ignored in healthcare research. As such, realist evaluation is useful in terms of understanding why a programme produces different outcomes when implemented in different settings. It describes what mechanisms (how people interpret and act upon ideas and opportunities presented by the programme) cause which outcome (intended or unintended consequences) and in which context (social and cultural conditions external to the programmes) ([Pawson and Tilley, 1997](#)). At its core, realist evaluation focuses on developing explanations of the consequences of social actions that contribute to a better understanding of why, where, and for whom programmes work or fail to work. Realist evaluations ask not 'what works?' or 'does this programme work?' but ask instead, 'what works for whom in what circumstances and in what respects, and how?' ([Pawson and Tilley, 2005](#)). Therefore, I have placed emphasis on:

- a) identifying the **mechanisms** that produce observable programme effects
- b) testing these **mechanisms** and the other **contexts** that may have impacts on the effects observed

**Figure 2.1 The realist evaluation cycle (adapted from (Pawson and Tilley, 1997))**



Realist evaluation stresses four key linked concepts, as illustrated by Figure 2.1 for explaining and understanding programmes: mechanism, context, outcome patterns and context-mechanism-outcome configurations (CMOCs).

According to Funnell and Rodgers (2011) there is merit in using multiple methods in realist evaluation. Realist synthesis uses purposeful sampling of diverse kinds of evidence available, and is inherently a process of building, testing and refining programme theory. In realist evaluation, programmes are theories, they are embedded, active, and are parts of open systems. A key requirement of realist evaluation is thus to take note of the different layers of social reality that make up and surround programmes. The triggers of change in most programmes are in the reasoning and resources of those effected by the programme. Programmes cannot be fully isolated or kept constant, as they are permeable and dynamic. Successful programmes can even change the conditions that made them work in the first place.

## 2.1 Mechanisms

The definition of 'mechanism' is essential in realist evaluation, as it affects the depth of analysis. Most authors use the definitions provided by Pawson and Tilley (1997). Mechanisms describe what it is about programmes that bring about any effects, and are often hidden. Many programmes have multiple components, but the term 'mechanism' is not used to distinguish these components. Programmes also often involve long sequences of steps before the outcome, which are also not considered mechanisms. A mechanism is also not the programme, the name of the programme or the presence or absence of the programme. A mechanism refers to the ways in which any one of the components brings about change. Mechanisms are usually embedded in structures (durable configuration of elements or series of steps). A mechanism describes what goes on within the programme to change behaviour. The definitions presented by Pommier et al. (2010), Rycroft-Malone et al. (2010), Ogrinc and Batalden (2009) and Clark et al. (2005) are aligned with Pawson and Tilley's view. Greenhalgh et al. (2009) define a mechanism as 'stakeholders' ideas about how change will be achieved and the mechanisms they described included descriptions of the actual programme. I did not refer to this definition in the thesis, as it conflated activities and modes of implementation with Pawson and Tilley's definition of mechanisms. In the view of Marchal et al. (2012), realist evaluation looks for mechanisms at individual, group, organisational and societal levels, which I took into consideration.

## 2.2 Outcomes (change)

Outcomes comprise the intended and unintended consequences of programmes, resulting from the activation of different mechanisms in different contexts. The nature and source of these differences is a key focus in realist evaluation. Sometimes it is difficult to be sure outcomes are caused by the programme, as they are valued and perceived differently by different parties; thus these perceptions should be investigated ([Øvretveit, 2002a](#)). Due to variations in context and mechanisms, any programme is likely to have mixed outcomes.

For the remainder of this thesis, I will refer to outcomes as ‘changes’. A change is a difference in a phenomenon of interest between two time points, often referred to as the before and after difference. I have selected the term ‘change’ to replace ‘outcome’, as it is commonly used to refer to clinical outcomes for patients. Using change rather than outcome invites a clarification of ‘change to what or whom?’ Change does not assume one thing caused the difference, whereas outcome assumes a difference is resultant of an action. Change emphasises the difference between two times, and leaves open the likely further changes. Change occurs when programmes, combined with the right contexts, release generative mechanisms. The context-mechanism-outcome configurations (CMOCs) is used as an analytical tool to analyse the data and unearth the mechanism ([Marchal et al., 2012](#)). Realists do not conceive that programmes work, instead, the actions of stakeholders make them work, and the causal potential of a programme provides reasons and resources to enable programme participants to change.

For the purposes of the study, I used a simple definition of change, provided by Dawson ([2003](#)): ‘New ways of organising and working’. A fundamental aspect of change is that it is a process, not an event. Various synonyms for change appear in the management literature, such as reform, renewal and improvement. Change can be regarded as a fact, but it is not the same as improvement, which is based on judgement ([Owen, 2005a](#)). Lewin ([1951](#)) suggested change is a three-stage process: unfreezing the old behaviour (or situation), moving to a

new level of behaviour, and refreezing the behaviour at the new level. Thus change entails moving from one equilibrium to another ([French and Bell, 1998](#)).

As advocated by Dawson ([2003](#)), change as a process takes time and is shaped by context. There is a subtle interplay between the content (the what) of change, the context (the where) of change, and the process (the how of change). The experience of change will vary over time and across individual groups. There can never be a single authentic story of change, there will always be multiple narratives and competing histories of change. In contrast to the dominant approach, which emphasises the importance of quantitative analyses ([Ledford et al., 1990](#)), the process approach is concerned with the collection of qualitative data which facilitates a more detailed understanding of the complex and dynamic process of change.



### 2.3 Context

Realist evaluation uses contextual thinking to address the issues of for whom and in what circumstances a programme will work. Defining 'context' and separating mechanism from context remains a difficult issue in the literature ([Marchal et al., 2012](#)). The effects of programmes vary across settings, indicating that conditions besides the programmes themselves influence outcomes ([Øvretveit, 2011a](#)), ([Kaplan et al., 2010](#)). This way of viewing context acknowledges that contexts and programmes influence each other ([Hovlid and Bukve, 2014](#)).

Kaplan et al. ([2010](#)) defined context as all non-programme factors (not part of a quality improvement programme itself) ([Øvretveit, 2011b](#)). This was in agreement with the view of Marchal et al. ([2012](#)) that it is better to consider contexts an external to the programme, which are present even if the programme does not lead to an outcome, but may have an influence on the outcome. Context can also be defined as the relevant circumstances that might affect the success or otherwise of improvement efforts ([Marshall and Bamber, 2011](#)). Realist evaluators focus on the contexts that may limit or enhance a programme's effectiveness, moderate the effects that have been observed with other groups, or cause differences in the populations served ([Henry, 2005](#)). Contexts of the providers participating in quality improvement and the organisations where quality improvement takes place must be considered when studying quality improvement ([Chao, 2007](#)).

Bate et al. ([2008](#)) examined case studies of healthcare organisations in the USA and Europe that had earned reputations for sustained achievement of quality improvement and concluded that all the successful organisations shared a talent for adapting solutions to their own organisational context.

Pettigrew ([1987](#)) distinguishes between the inner (immediate, intra-organisational, micro) context which includes organisational and divisional cultures, group norms, leadership, local champions and outer (social, political,

macro) contexts. Pettigrew's extensively used framework focuses on three key dimensions of strategic change. The first one refers to the content of the chosen strategy (the what of change), the second one is the process and management of change (the how) and finally, the context in which the strategy unfolds (the why).

Later, Pettigrew ([1992](#)) developed a model for the management of strategic change, which centres on receptive and non-receptive contexts for change. Receptive contexts are defined where context and management action 'seem to be favourable, associated with forward movement'. Non-receptive contexts are where a combination of conditions effectively creates blockages or resistances to change. Bate et al. ([2008](#)) argue that creating a 'receptive context' for change can encompass features such as strategic vision, good managerial relations, visionary staff in pivotal positions, a climate conducive to experimentation and risk taking, effective data capture systems, associated with the capacity to embrace new ideas and implement innovations. Another fundamental role for creating a receptive context for change and learning is to provide funding and resources required to support quality improvement programmes and implement service improvements.

Taking the starting point as Pettigrew et al.'s ([1992](#)) notion of receptive and non-receptive contexts for change which, it is argued now needs to be combined with more contemporary psychological perspectives such as Weiner's ([2009](#)) notion of readiness for change, it is argued that more attention must be paid to the multiple levels of outer and inner context and how these combine to impact upon the success and sustainability of quality improvement efforts.

Bate, Mendel and Robert's framework Organizing for Quality ([2008](#)) moves beyond listing successful contexts to identifying how these contexts unfold to bring about successful quality improvement efforts. This framework is one of the first that applies organisational theories to disentangle the how of improving quality works ([Robert et al., 2008](#)), ([Robert et al., 2011](#)). It highlights common organisational domains important to address in considering whether a

group has the readiness to and is ready to engage with a quality improvement programme.

### **2.3.1 Readiness to change**

Readiness to change, similar to Lewin's (1951) concept of unfreezing, is reflected in organisational members' beliefs, attitudes, and intentions regarding the extent to which changes are needed and the organisation's capacity to successfully make those changes (Armenakis and Bedian, 1999), (French et al., 2004). It is a critical precursor to an organisation's ability to successfully implement change (Weiner, 2009), (Holt et al., 2008), and a lack of organisational readiness for change is an important context for understanding why implementation efforts fail (Holt et al., 2008). Weiner (2009) describes organisational readiness as referring to organisational members' change commitment and change efficacy to implement organisational change with 'readiness' bringing a state of being both psychologically and behavioural prepared to take action (i.e. willing and able).

Ming-Chu and Meng-Hsiu (2014) suggest that companies should consider how to enhance employees' organisational readiness for change (Ferlie et al., 2005). They argue that if managers incorporate the concept of organisational change into employees' daily tasks, employees can translate this into personal values. Weber and Weber (2001) also argued that a higher degree of support in the work environment influences greater readiness for change.

### **2.3.2 Resistance to change**

Employees may feel supportive of, resistant to, or ambivalent towards organisational change (Deetz, 2008), (Oreg and Sverdlik, 2011), (Wanberg and Banas, 2000). Resistance to change is recognised as a critical context that influences the success or failure of any organisational change effort (Piderit, 2000). It can be defined as an employee's behavioural response to managerial control and the treatment that they receive during the change process (Coetsee and Stanz, 2007). Although the literature tends to conceptualise resistance as a

negative response to change, it can actually be helpful, particularly if employees can shed light on concerns regarding intended change ([Coetsee and Stanz, 2007](#)) ([Brian et al., 2015](#)). Thus managers must first confront and reduce employees' resistance to change, because this disposition is an obstacle to successful organisational change ([Furst and Cable, 2008](#)). It is important to identify organisational barriers to change to understand exactly what hinders the successful implementation of programmes and attempts to change practice, so that issues can be addressed proactively as well as reactively ([Brian et al., 2015](#)).

According to Grol ([1994](#)) it is also crucial to allow sufficient time for change within external peer review programmes. Difficulties can be compounded when new initiatives are not given a diagnosis phase or enough time to 'bed in' ([Dixon-Woods et al., 2012](#)), ([Halladay and Bero, 2000](#)). Time should be set aside for identifying the barriers to change, which may be related to the care providers and their characteristics but more often to the setting in which they work ([Berwick, 1989](#)), ([Grol, 1992](#)); for developing a plan for improvement with specific programmes; for managing the change process well; and for evaluating the results.

The scale of resources required to support improvements is often underestimated ([The Health Foundation, 2011](#)), ([Burgoyne et al., 2009](#)), and without adequate financial support, infrastructure, managerial skills and dedicated time, efforts to improve quality can quickly run into difficulties ([Needham et al., 2009](#)). Engaging staff and overcoming a perceived lack of ownership are among the biggest challenges in improvement efforts ([Scott et al., 2003](#)). Boundaries between professional, disciplinary and managerial groups present important obstacles to change, and consensus within one profession is not always shared by others ([Ling et al., 2010](#)). Middle managers and frontline staff can be especially difficult to engage in improvement, as they already face numerous, complex, competing clinical and organisational demands, often with inadequate staffing, limited resources and equipment shortages ([The Health Foundation, 2011](#)).

## 2.4 Context-mechanism-outcome configurations

Realist evaluation is characterised by theory testing and refinement. Knowledge development is used to construct and test theories about mechanisms as they occur in contexts ([Mark et al., 2000](#)). Context-mechanism-outcome configurations (CMOCs) comprise models indicating how programmes activate mechanisms, among whom, and in what conditions to bring about alterations in behavioural or event or state regularities. These propositions bring together mechanism variation and relevant context variation to predict and explain outcome pattern variation. Realist evaluation thus develops and tests CMOCs empirically.

A range of methods is used to present CMOCs. Some authors use narratives to describe the individual elements and interactions of the CMOCs ([Byng et al., 2008](#)), ([Clark et al., 2005](#)), ([Marchal et al., 2010](#)). Others use tables ([Greenhalgh et al., 2009](#)), ([Tolson et al., 2005](#)). A logic model might represent the programme in terms of inputs (resources, time, effort, preparation etc.), activities (meeting standards), outputs (QI), and outcomes (change). However, it does not explain what it is about the programme that will produce or lead to the outcomes of interest, and is not clear what it is about the various components that make the programme work. Logic models often imply that the programme by itself is sufficient to produce these impacts, whereas realist evaluation assumes the programme cannot just produce the intended impacts without any assistance from other programmes or favourable contexts.

Ogrinc and Batalden ([2009](#)) developed a realist hypothesis grid that contains elements of plausible mechanisms, context and outcomes and can be used to generate potential CMOCs during a realist evaluation. Similarly, Byng et al. ([2008](#)) used predictor-outcome matrices that incorporated CMOCs and clarified the analytical induction process. Realist matrices are a complementary approach that I used to focus on one or more causal mechanism in a programme theory, and explore what it is about the programme that makes this causal mechanism work.

### **2.4.1 Programme theory**

Realist evaluation seeks to build initial programme theories, test and refine them ([Mark et al., 2000](#)) ([Pawson and Tilley, 2005](#)). The initial sets of programme theories are propositions which span context, mechanism and outcome and drive the remaining aspects of the realist evaluation ([Pawson and Tilley, 1997](#)). Programme theory is derived from two components, a Middle-Range Theory (MRT) and a theory of action. An MRT encompasses central drivers from which change comes about for individuals or groups. A theory of action could derive from a formal evidence-based theory or unstated tacit understanding about how things work, and the theory of action explains how programmes are constructed to activate these changes ([Funnell and Rodgers, 2011](#)). Pawson and Tilley use the definition of MRT provided by Merton ([1968](#)):

‘theories that lie between the minor but unnecessary working hypotheses ... and the all-inclusive systematic efforts to develop a unified theory that will explain all the observed uniformities of social behaviour, social organization and social change’.

The MRT can be formulated based on existing theory, past experience and previous evaluations or research studies. The result is discussed with the stakeholders and finally results in a testable MRT. The account of how mechanisms that explain how a programme leads to particular outcomes is formulated as an MRT. While no exact distinctions are provided by the literature, in the understanding provided by ([Marchal et al., 2012](#)), the MRT in realist evaluation is situated at a more abstract level than what is called the ‘operational’ programme theory in theory-driven evaluation, or the logic model in ‘theory of change’ (TOC).

### **2.4.2 Black box**

Pawson and Tilley ([1994](#)) claim quality improvement programmes are seen by some experimentalists as black boxes. To understand why programmes work (assuming they do work), one needs to know which social and behavioural mechanisms are active and in which contexts. Opening up the black box is helpful to understand why programmes have (no) effects or unintended effects.

Marchal et al. ([2012](#)) argue that realist evaluation provides a sound framework to examine how context and mechanisms influence the outcomes of a programme, and open the black box ([Blaise and Kegels, 2004](#)) ([Clark et al., 2005](#)) ([Greenhalgh et al., 2009](#)) ([Ogrinc and Batalden, 2009](#)) ([Tolson et al., 2005](#)). It rapidly takes us beyond the ‘one problem, one cure’ view of social programmes ([Pawson and Sridharan, 2009](#)).

A possible limitation of the realist evaluation for studying quality improvement changes are that the concepts of context, mechanism and outcome are not well defined and only illustrated in a few studies. It is also unclear exactly how ‘mechanism’ is elucidated: it does not just describe the programme components or implementing actions, but how this higher-level conceptualisation is created is unclear – of how the actions work ‘generative mechanism’, which is different from their interaction with context.

As qualitative process descriptions are lacking, quality improvement programmes are often described as ‘black boxes’ ([Schouten et al., 2008](#)), ([Wilson et al., 2003](#)). Knowing what actually occurs in setting up and carrying out a quality improvement programme would seem crucial for interpreting the effectiveness results ([Schouten et al., 2008](#)), ([Wilson et al., 2003](#)), ([Øvretveit and Gustafson, 2002](#)), ([Broer et al., 2010](#)). There is increasing recognition of the value of process evaluations using qualitative methods alongside trials of complex programmes, which address a broader set of questions and can enhance the scientific and policy value. It has therefore been argued that randomised controlled trials alone are limited in serving the purposes of health services research ([Ward et al., 2003](#)), and realist evaluation is critical of this.

To conclude, I have used realist evaluation in this thesis to develop, test and refine CMOCs aimed at strengthening the operation of existing external peer review programmes, using the definitions outlined in this chapter.

## Chapter 3 Systematic Literature Review

I conducted a systematic literature review of external peer review programmes across inpatient and community healthcare settings to help me formulate the research questions necessary for my thesis, design the studies that would be undertaken and to be able to place the results of my research in the context of what is already known.

### 3.1 Aims

My overarching aim was to synthesise existing knowledge about how change is brought about in external peer review programmes. I wanted to identify which programme and non-programme (contextual) factors are suggested by previous literature, theory or evidence to be critical for the success of the programme ([Øvretveit, 2004](#)). I have not focused on the impact of external peer review (unless it was considered to be a facilitator or barrier of success), as this has been extensively covered in previous literature reviews ([Kilsdonk et al., 2015](#)). The specific objectives of the systematic literature review were to:

- 1) Summarise what is already known about how change occurs in external peer review to generate improvements in service quality to identify any gaps in the literature
- 2) Investigate which programme factors are important for generating improvements in service quality in external peer review
- 3) Identify the non-programme (contextual) factors which facilitate or impede their success



## 3.2 Design

There are two types of literature review: the narrative review and the systematic review. Narrative reviews take a conceptual consideration of literature with less methodological rigidity assigned to search methods. Often, they are not replicable and they do not provide a transparent presentation of numbers and patterns of findings that can be offered by systematic reviews. Systematic reviews follow a strict protocol of clear aims, explicit search strategy, detailed inclusion and exclusion criteria, standardised study quality assessment and a systematic synthesis of studies ([Higgins and Green, 2011](#), [Clarke, 2011](#), [Chalmers and Altman, 1995](#)). I chose to undertake a systematic review as part of the realist evaluation, as I wanted to generate a comprehensive and authoritative account of factors that influence external peer review programmes ([van der Knaap et al., 2008](#)).

Historically, systematic reviews have focused on reviewing the results of intervention studies such as randomised controlled trials, however researchers have recently affirmed the place of qualitative literature in systematic reviews ([Dixon-Woods and Fitzpatrick, 2001](#)) and established that healthcare research particularly benefits from this structure ([Delamothe and Smith, 1996](#)). I decided that this rigorous approach to understanding what influences the success or failure of external peer review programmes would be of benefit, providing contrast to the present anecdotal nature of much of the literature.

### **3.3 Search Strategy**

#### **3.3.1 Setting**

I only included studies based in hospitals, primary healthcare organisations, community-based healthcare organisations and other healthcare providers. I excluded programmes that featured public health or health promotion interventions aimed at those not in contact with healthcare services. I also excluded studies in educational, laboratory, pathology, social care or non-healthcare settings.

#### **3.3.2 Types of interventions**

I only included literature focused on external peer review programmes in which peer reviewers were external to the host service or unit under review ([Burnett et al., 2007](#)). I excluded studies if services did not use the feedback from external peer reviews to improve service quality. Feedback could include any recommendations; delivered in a written, electronic or verbal format. I only considered studies where external peer review programmes targeted quality improvements which directly impacted service users. I excluded external peer review programmes conducted by statutory bodies, or for the purpose of audits, regulation, self-review or continued professional development (CPD). I excluded studies assessing service quality as whole systems or policies at a regional or national level.

#### **3.3.3 Relevant outcome measures**

I considered any measure of service quality, such as the following (but not limited to):

1. Measures of service user outcome (e.g. condition-specific measures of outcome related to service users' health)
2. Measures of healthcare organisational change (e.g. organisational performance)
3. Measures of perceived changes by stakeholders involved in the process
4. Unanticipated or adverse consequences

### 3.3.4 Types of studies

I included all types of qualitative and quantitative research, including:

1. Descriptive studies such as case reports, case studies and accounts
2. Observational studies such as cross-sectional, cohort and case-control studies
3. Intervention designs such as RCT, quasi-experimental and adaptive clinical trials
4. Qualitative studies

After preliminary searches, I identified a few opinion pieces in which experts described their experiences of external peer review programmes. These papers provided an important source of information about the context of programmes so I also included them.

### 3.3.5 Inclusion & exclusion

I subjected literature identified by the search strategy to a process of study selection. I created specific criteria (displayed with definitions and rationales in Table 3.2) from background reading and by identifying five key reference papers (information displayed in Table 3.1).

**Table 3.1 Five key reference papers**

Reference	Title
<a href="#">(Braithwaite et al., 2010)</a>	Health service accreditation as a predictor of organisational performance; a blinded, random, stratified study
<a href="#">(Piper et al., 2006)</a>	Experience of six years of a regional peer review scheme in rheumatology
<a href="#">(Roberts et al., 2010)</a>	A randomised trial of peer review: the UK National Chronic Obstructive Pulmonary Disease Resources and Outcomes Project
<a href="#">(Salmon et al., 2003)</a>	The Impact of Accreditation on the Quality of Hospital Care: KwaZulu-Natal Province, Republic of South Africa
<a href="#">(Worrall, 2011)</a>	The service context for clinical guidelines: supporting guideline implementation by assuring and improving the quality of service in which clinicians work

**Table 3.2 Inclusion criteria, with definitions and rationales**

Criterion	Definition	Rationale
Studies focus on healthcare	Social care, education or non-healthcare settings excluded	Healthcare was target for this study
Programme setting: inpatient and community medical services	Whole systems, statutory bodies, policies, programmes assessing service quality at a regional or national level excluded Non-service user clinical environments such as laboratories excluded Public health or health promotion quality assessments excluded	Study aimed to identify contexts that influenced quality of care delivered to service users
Programme: external peer review	Explicit mention of external peer review (peer review networks or accreditation schemes), involving benchmarking / checking standards, site visits and feedback from peer reviewers	Needed to provide feedback about service used to improve service quality
External assessment by peer reviewers	Audits, ISO, EFQM, regulation mechanisms and other systems that did not include people excluded Programmes that did not involve professionals working in healthcare (service users and other stakeholders may be involved in these teams) excluded Audit, self-review, chart-reviews and internal peer reviews excluded	Peer reviews had to be carried out by people (peers) An assessment of service quality not competency
Written in English	Studies written in other languages excluded	Needed to be understood by author
Humans	Studies which did not involve humans as the outcome (i.e. mechanised or regulatory processes) excluded	Studies had to improve service quality for humans

### 3.3.6 Identifying and modifying search terms for identification of studies

I followed guidance from Tanon et al. ([2010](#)) who outlined the most efficient search strategies for systematic reviews in patient-safety (closely related to quality improvement) when developing the search terms. I tried to balance sensitivity, specificity and precision to capture as many relevant studies as possible on the topic area. I used a multi-method search strategy, similar to other relevant reviews ([Greenfield and Braithwaite, 2008](#)) ([Alkhenizan and Shaw, 2012](#)) ([Shouten et al., 2008](#)) ([Grimshaw et al., 2003](#)).

I began by identifying facets of the research question as: 'external feedback', 'healthcare' and 'quality improvement'. I selected OVID as the initial search engine to pilot search terms as OVID featured a helpful tool to 'explode' terms and link them to Medical Subject Headings (MeSH). I tested multiple iterations and combinations of all search terms to achieve the best level of specificity and sensitivity. For example, the terms 'health' and 'care' separately identified a substantial number of general articles. Therefore, I selected 'healthcare' as one term. I piloted broad search terms in OVID and produced over 6148 references; the majority were not relevant to the search criteria. I then expanded keywords into a list of search terms, initially brainstorming for synonyms with my supervisor and using keywords from relevant articles. I mapped search terms to MeSH wherever possible and 'exploded' them to ensure all subheadings were included. This greatly increased the number of papers identified by the search. All databases required slightly different MeSH terms; I therefore used variations of the original search strategy. This search strategy generated over 1062 references, few of which addressed the aims of my systematic review. I therefore narrowed the results of the search by selecting narrower search terms within the three facets from the initial search: 'Peer Review' or 'professional peer review', 'accreditation' or 'collaborative', plus 'quality assurance, health care/mt, og, st [Methods, Organisation & Administration, Standards]'. I identified 542 of the 889 records by this strategy in OVID. I truncated and combined search terms using Boolean operators, and I limited the search to article titles.

Table 3.3 displays a summary of the final search strategy, which was run on 26<sup>th</sup> September 2015, producing 889 references. After removing duplicates, I limited results to papers that studied humans and were written in the English language. I searched the databases without time limits, and included studies regardless of publication status. This returned 542 references from OVID, which I subjected to inclusion and exclusion criteria.

**Table 3.3 Final search strategy used in OVID for the review**

#	Search term	Number of articles
1	"Peer Review" / or professional peer review.mp.	6148
2	Accreditation/	11169
3	collaborative.mp.	29401
4	1 or 2 or 3	46493
5	Quality Assurance, Health Care/mt, og, st [Methods, Organization & Administration, Standards]	15206
6	4 and 5	889
7	Limit 5 to English language	782
8	Limit 6 to humans (not valid in Psych Info so records retained)	675
9	Remove duplicates	542
10	Sensitivity check with Worall, Braithwaite, Roberts, Salmon, Piper	PASSED

At each stage, I crosschecked the search lists produced with the five key reference papers as a method of checking sensitivity to ensure my search terms were adequate. At all stages, this sensitivity check was passed.

### **3.3.7 Electronic database search**

Table 3.4 outlines the various electronic databases that I used. I began with the OVID databases, then proceeded to EBSCO databases and Web of Knowledge. Finally, I ended the search with the Cochrane Databases and electronic theses. I last update the search on the 26<sup>th</sup> September 2015.

**Table 3.4 Electronic databases**

Database	Years Searched	Date last searched
MEDLINE (Ovid)	1950 to September 2015	26 <sup>th</sup> September 2015
EMBASE (Ovid)	1980 to September 2015	26 <sup>th</sup> September 2015
PsycINFO (Ovid)	1806 to September 2015	26 <sup>th</sup> September 2015
CINAHL, Business Source Complete (EBSCO)	1980 to September 2015	26 <sup>th</sup> September 2015
Science Citation Index, Social Science Citation Index, ISI Conference Proceedings (Web of Knowledge)	1970 to September 2015	26 <sup>th</sup> September 2015
Cochrane Central Register of Controlled Trials (CENTRAL) 2015, Issue 8, part of <i>The Cochrane Library</i> . <a href="http://www.cochranelibrary.com">www.cochranelibrary.com</a> , including the Cochrane Effective Practice and Organisation of Care (EPOC) Group Specialised Register		26 <sup>th</sup> September 2015
Cochrane Database of Systematic Reviews (CDSR) Cochrane Library 2015 Issue 8		September 2015
Database of Abstracts of Reviews of Effectiveness (DARE) Cochrane Library 2015, Issue 4		March 2015
Electronic Theses Online (EThOS) <a href="http://www.ethos.ac.uk">www.ethos.ac.uk</a>		26 <sup>th</sup> September 2015)

### 3.3.8 Identification of additional materials

As shown in Figure 3.1, I used a second search strategy by contacting 29 sponsoring bodies in the UK ([Academy of Medical Royal Colleges, 2015](#)), authors and key experts mentioned by relevant papers, regarding any published

or unpublished work. I mainly searched websites, and where possible, engaged in discussion with key authors. While only four documents were retrieved from this search, it enabled me to identify which bodies in the UK were undertaking research in this area.

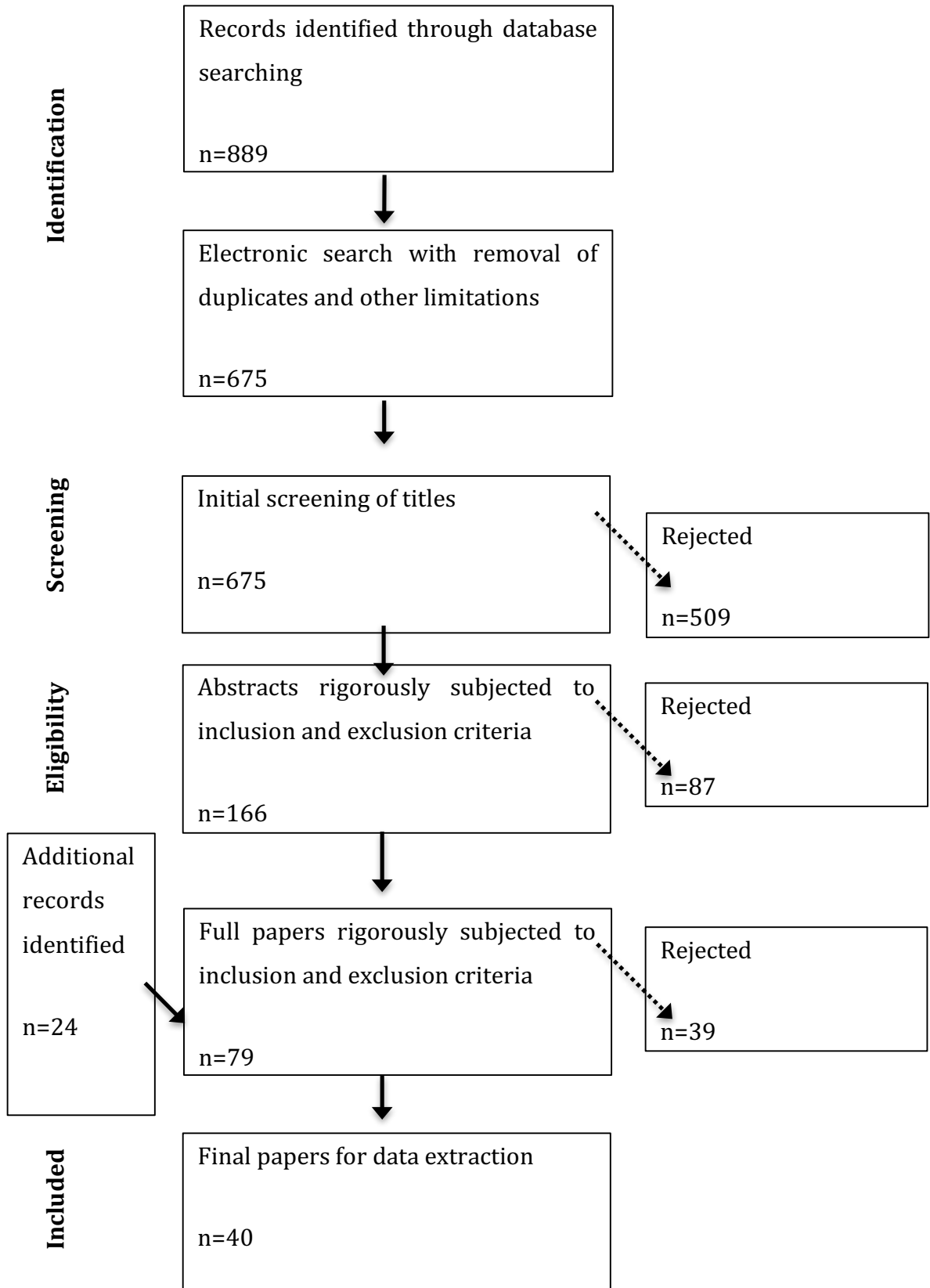
I used a 'snowballing' technique as a third strategy ([Greenhalgh and Peacock, 2005](#)) whereby I hand-searched reference lists of all included studies and key materials (discussion papers, articles and reports). I followed up these references using internet search engines to locate documents which met my inclusion and exclusion criteria. I identified a further twenty documents, including opinion pieces, reports and peer-reviewed articles.

### **3.3.9 Abstract selection**

Figure 3.1 presents the numbers of articles included and excluded at each stage of the review process ([Liberati et al., 2009](#)). I screened titles of 889 references from the electronic search, and discarded those that clearly did not meet the inclusion criteria. Of the remaining 166 references, I studied these abstracts in detail. I used a 'criterion-led method', whereby I ordered the selection criteria into a hierarchy of importance, and applied these one at a time to the final abstracts. Next, I examined abstracts were using three possible scores, awarding scores of either 0 or 1 depending on fulfilment of each criterion. At each stage, papers I scored 0 were rejected, and the reasons for each score were noted. Following this process, 79 abstracts remained for further review. 24 further abstracts were included from the other two search strategies and I reviewed a total of 103 papers for further assessment. The remaining 40 studies, which met the inclusion criteria, are reported in detail in Table 3.5 reporting included studies.



Figure 3.1 Preferred reported items for systematic reviews (PRISMA) flow chart



### **3.4 Data extraction, quality assessment and data synthesis**

I extracted the data from included studies, collecting information on authors, study aims, the type of external peer review programme, area of healthcare, coordinating body, study methods, sample size and respondent groups.

#### **3.4.1 Quality assessment**

Quality assessment ensures that a balanced view of the studies' results is presented in the context of their methodological strengths or weaknesses. Each study was assessed using a quality appraisal tool developed by Kmet et al. (2004) which comprised of two checklists to assess qualitative and quantitative research. I chose this tool as it had been validated, allowed for analysis of both qualitative and quantitative methods, and was considered to be generalisable to study designs. Using respective checklists, I calculated total scores for each paper. All articles were scored on 24 questions with a number between 0–2 (0 = Not reported, or reported but not fulfilled at all; 1 = could not tell if reported and fulfilled to partial satisfaction; 2 = reported and fulfilled to full satisfaction) with an additional not applicable (N/A) option for quantitative studies. Qualitative articles were subjected to ten questions, and quantitative articles were subjected to 14 questions. Mixed methods studies were subjected to all 24 questions. The total scores are expressed as percentages in Table 3.5.

#### **3.4.2 Data extraction and management**

I extracted data using EndNote X7, Microsoft Excel and MAXQDA 11. I used EndNote X7 and MAXQDA 11 to manage data and annotate the individual papers. I then used an Excel spreadsheet to extract details of study methods, sample populations and results. There is currently a lack of established methods for synthesis of qualitative research (Britten et al., 2002). Relevant synthesis strategies include narrative description, cross-case thematic analysis (Miles et al., 2013), meta-ethnography and meta-synthesis (Dixon-Woods et al., 2004). I performed a narrative synthesis of the results with consideration of the risk of bias and the quality of the studies. I was unable to undertake meta-analysis due to marked heterogeneity in design and outcomes of the studies.

In this synthesis, I aimed to maintain the original meanings, interpretations and raw data offered by the articles, also endorsed by Walshe and Downe ([2004](#)). I used descriptive groupings of the data to present patterns of recurring topics across the dataset, similar to thematic narrative synthesis. A difficulty in this approach is bias against information that does not arise more than once ([Estabrooks et al., 1994](#)), although I mitigated against this by reporting salient outlier results.

### **3.5 Overview of Results**

I have presented the key study findings from 40 included papers in Table 3.5, grouped according to the type of external peer review programme, by author, publication year and country. Firstly, I have presented literature on accreditation schemes, and then secondly on peer review networks, and finally a paper that presents both types.

Co-ordinating bodies included the King's Fund Centre, the College Centre for Quality Improvement (CCQI), the Australian Council on Healthcare Standards (ACHS), the Joint Commission on Accreditation of Healthcare Organisations, Zambia Health Accreditation Council, North Carolina Local Health Department Accreditation, American College of Obstetricians and Gynecologists, Trent Small Hospital Accreditation Scheme, Irish Health Services Accreditation Board, National Agency for Healthcare Accreditation and Evaluation, Canadian Council on Health Services Accreditation, Council for Health Service Accreditation of Southern Africa, Joint Advisory Group on GI Endoscopy, National Cancer Action Team, Association of Paediatric Anaesthetists of Great Britain and Ireland, American College of Physician Executives, Comprehensive Cancer Care North Netherlands, the Royal College of Physicians, West Midlands Rheumatology Service and Training Committee, UK National COPD Resources and Outcomes Project (NCROP) and some papers included a mixture of bodies.

**Table 3.5 Included literature characteristics**

First author; Year [Country]	Study Aims	Co-ordinating Body	Area	Study Methods	Sample Size (No. of Organisations)	Respondent Groups	Quality Assessment Score for Qualitative Studies (%)	Quality Assessment Score for Quantitative Studies (%)
<b>Accreditation</b>								
<a href="#">(Balkizas, 1995)</a> [UK]	To describe accreditation during formation of a nursing development unit (NDU)	King's Fund Centre	Community Learning Disability	Case study	(1 NDU)	N/A	40	-
<a href="#">(Baskind et al., 2010)</a> [UK]	To explore effects of accreditation on standards and how staff achieved change	CCQI	Mental health	Qualitative (interviews)	n=8	Local project lead from wards that had failed accreditation	85	-
<a href="#">(Braithwaite et al., 2010)</a> [Australia]	To determine association between accreditation and clinical and organisational performance	ACHS	Health Services	Quantitative (surveys)	(19 accreditation agencies)	N/A	-	75
<a href="#">(Braithwaite et al., 2012)</a> [Mixed]	To describe how organisational attributes of accreditation differ between LMICs and HICs	Mixed	Mixed	Quantitative (survey)	(44 accreditation agencies)	National and international healthcare accreditation providers	-	86.36
<a href="#">(Brasure et al., 2000)</a> [USA]	To explore why rural hospitals are less likely to be accredited than urban hospitals	JCAHO	Rural Hospitals	Mixed methods (routine data, survey, telephone follow-up)	n=24 (913 non-accredited rural hospitals)	Non-accredited hospital administrators	80	-
<a href="#">(Bukonda et al., 2002)</a> [Zambia]	To describe the development of the Zambia Hospital Accreditation Programme from	Zambia Health Accreditation	Hospitals	Mixed methods (document review, interviews, focus	N/A	Stakeholders, consultants, key implementers	60	39.29

	1997 – 2000	on Council		groups, hospital visits, discussion)				
<a href="#">(Davis et al., 2011)</a> [USA]	To examine the extent to which accredited local health departments (LHDs) conducted accreditation activities	North Carolina LHD Accreditation	LHDs	Mixed methods (evaluation report data, online surveys)	(48 accredited North Carolina LHDs)	Local health directors	80	72.73
<a href="#">(Devers et al., 2004)</a> [USA]	To describe programmes' progress towards implementation	JCAHO	Hospitals	Qualitative (interviews)	n=87 interviews with leaders, n=226 interviews other	Leaders of large hospitalsl employers, brokers, health plans, medical groups	65	-
<a href="#">(Doyle and Grampp, 2014)</a> [Ireland]	To explore hospital accreditation as a quality tool	IHSAB	Hospital	Mixed methods (interviews, focus groups, surveys, walks of the facilities)	n=73 interviews, n=7 focus groups, n=72 surveys, n=2 walks (3 accredited sites)	Team members and team leaders of four core teams, all care teams in each hospital	90	68.18
<a href="#">(Duckett, 1983)</a> [Australia]	To analyse the impact of accreditation on Australian hospitals	ACHS	Hospitals	Qualitative (interviews)	(23 hospitals)	Senior staff	75	-
<a href="#">(Gibbs and Cheetham, 1988)</a> [USA]	To describe the Voluntary Review of Quality Care (VRQC)	ACOG	Obstetrics & Gynecology	Qualitative (questionnaires)	(14 participating hospitals, 100 hospitals that expressed preliminary interest but did not	Members of the review team, hospital representative	35	-

					participate)			
( <a href="#">Greenfield et al., 2010</a> ) [Australia]	To explore the experience of accreditation participants	ACHS	Teaching hospital	Qualitative (interviews)	n=30 interviews (1 large public teaching hospital)	Staff who had been formally interviewed (board representative, senior executive, manager, frontline staff)	80	-
( <a href="#">Hinchcliff et al., 2013</a> ) [Australia]	To examine stakeholders' perspectives about influences of accreditation participation	ACHS	Mixed	Qualitative (focus groups, interviews)	n=258	Various	100	-
( <a href="#">Hurst, 1997</a> ) [UK]	To discuss characteristics of accreditation, especially implications of voluntary status focusing on TSHAS	TSHAS	Hospitals	Qualitative (literature review, interviews, questionnaires)	n=20 interviews with informants, n=38 questionnaires from staff participating in TSHAS	Key informants (TSHAS peer reviewers and managers) and staff participating in TSHAS	65	-
( <a href="#">Jaafaripooyan et al., 2011</a> ) [Various]	To identify performance measures to facilitate evaluation of accreditation	Mixed	Mixed	Qualitative (interviews and open-ended questionnaires)	n=25 experts, n=120 health professionals	Experts in healthcare accreditation-associated institutions, hospitals' professionals in a developing country (i.e. Iran)	100	-
( <a href="#">Kern, 2002</a> ) [USA]	To examine differences between surveyed accredited and non-accredited facilities	JCAHO	Long-term care	Quantitative (survey)	(13654 long-term care facilities)	N/A	-	30.77
( <a href="#">Lemieux-Charles et al., 2000</a> )	To complement CCHSA's AIM (Achieving Improved Measurement) project to integrate performance	CCHSA	Acute care healthcare organization	Mixed methods (questionnaires and telephone)	(319 teams)	Teams preparing for accreditation or who had been accredited in the last	60	64.28

[Canada]	indicators into accreditation		s	interviews)		two years		
<a href="#">(O'Connor et al., 2007)</a> [Ireland]	To evaluate role and participation of lay peer reviewers	IHSAB	Mixed	Mixed methods (focus groups, interviews, questionnaires)	n=4 lay peer reviewers, n=16 peer reviewers, n=56 hospital staff	Lay peer reviewers, peer reviewers, applicant organisations, hospital staff	75	54.55
<a href="#">(Paccioni et al., 2007)</a> [Canada]	To describe and understand the effects of accreditation on organisational control	Mixed	Primary care centre	Mixed methods (group interview, interviews, questionnaires, documents, non-participant observation)	n=14 (A), n=21 (B), n=328 (questionnaires), n=60 (observations)	Qualitative (administrators directly involved in accreditation), quantitative (all professionals and employees not directly involved)	90	81.82
<a href="#">(Pomey et al., 2004)</a> [France]	To examine the dynamics of change that operated following preparations for accreditation	ANAES	Hospital	Mixed methods (interviews, questionnaires, documents and observation of means)	n=67 interviews, n=1755 questionnaires (1 UHC)	Various	90	77.27
<a href="#">(Pomey et al., 2010)</a> [Canada]	To evaluate how accreditation helps to introduce organisational changes	CCHSA	Healthcare organisations	Mixed methods (document review, interviews and focus groups)	(5 healthcare organisations)	Managers (interviews), staff involved with accreditation (focus groups)	85	72.73
<a href="#">(Salmon et al., 2003)</a> [Republic of South Africa]	To assess effects of accreditation on hospitals' processes and outcomes in a developing country	COHSASA	Public hospitals	RCT	(20 hospitals)	Public hospitals, interviews with nurses	-	91.67
<a href="#">(Stebbing,</a>	To describe the processes of the development and quality	JAG	GI	Descriptive	N/A	N/A	40	-



<a href="#">2011</a> [UK]	assurance of endoscopy units		endoscopy	account				
<a href="#">(Touati and Pomey, 2009)</a> [France, Canada]	To identify differences between French and Canadian accreditation experiences	Mixed	Mixed	Qualitative meta-analysis	(2 countries)	N/A	60	-
<a href="#">(Valori et al., 2013)</a> [UK]	To describe service accreditation to propose that it should be professionally led	Mixed	Mixed	Descriptive account	N/A	N/A	30	-
<b>Peer review network</b>								
<a href="#">(Bray, 2013)</a> [UK]	To summarise stroke service's peer review network and the Anglia heart and stroke network	RCP	Stroke Services	Descriptive account	n=1 (18 acute trusts)	East of England acute trusts	10	-
<a href="#">(Burnett et al., 2007)</a> [UK]	To assess value and effectiveness of a peer review network	National Cancer Action Team	Cancer services	Mixed methods (document review, workshop, interviews, structured questionnaires)	n=289	Patients, carers and staff involved in Peer Review, staff who have acted as reviewers, various stakeholders	80	68.18
<a href="#">(Butterfield et al., 2012)</a> [UK]	To describe external pathway peer review for improving quality of NHS care	Mixed	Mixed	Qualitative (survey)	n=40	Health sector leaders	75	-
<a href="#">(Crean et al., 2003)</a> [UK]	To assess the process of interdepartmental peer review network of paediatric anaesthesia departments	APA	Anaesthesia	Descriptive account	(6 departments)	Paediatric Anaesthetic Departments	50	-

<a href="#">(Edwards, 2011)</a> [USA]	To determine the association between peer review network programme factors associated with higher subjective quality impact, and better objective performance	ACPE	Acute care hospitals	Quantitative (survey, objective data)	n=296 (1017 non-federal acute care CMS hospitals)	Staff self-identified as holding leadership roles	-	90.91
<a href="#">(Kilsdonk et al., 2014)</a> [Netherlands]	To examine impact of participation and implementation of recommendations on treatment patterns and survival	Comprehensive Cancer Care North Netherlands	Cancer	Quantitative (hospital data)	n=45705 (30 hospitals)	Patients diagnosed with primary invasive epithelial colorectal cancer (Netherlands Cancer Registry)	-	86.36
<a href="#">(Page and Harrison, 1995)</a> [UK]	To outline the planning, implementation, key results, benefits and difficulties	RCP	Respiratory medicine	Qualitative (questionnaires)	(20 units)	Staff from respiratory units	30	-
<a href="#">(Piper et al., 2006)</a> [UK]	To report experience of a rheumatology peer review network	W. Midlands Rheumatology Service and Training Committee	Rheumatology	Qualitative (survey)	n=59 (12 rheumatology units)	Health professionals	50	-
<a href="#">(Rivas et al., 2008)</a> ( <a href="#">Rivas et al., 2010</a> ) ( <a href="#">Rivas et al., 2012</a> ) [UK]	To explore participants' experiences of NCROP (National COPD Resources and Outcomes Project) and change management	NCROP	Respiratory medicine	Qualitative (change diaries, semi-structured interviews)	n=43	Hospital respiratory consultants, nurses, general managers	90	-
<a href="#">(Roberts et al., 2010)</a> ( <a href="#">Roberts et al., 2012</a> )	To evaluate if peer review network of respiratory units improves COPD services	NCROP	Respiratory medicine	RCT	(100 hospital units (54 units intervention,	N/A	-	83.33

[UK]					46 units control))			
					(82 hospital units)			
<a href="#">(Suñol et al., 2009)</a> [EU]	To explore the association between quality improvement implementation in hospitals and their success in meeting quality requirements	Mixed	Hospitals	Mixed methods (questionnaire, on-site audit)	n=89 (389 hospitals)	Acute care hospitals	100	91.67
<b>Mixed approaches</b>								
<a href="#">(Worrall, 2011)</a> [UK]	To describe the context within accreditation and quality networks	CCQI	Mental Health Services	Opinion piece	N/A	N/A	30	-

### 3.5.1 Research settings

Some of the studies included in this review compared more than one country (n=2), Pan-European (n=1), and the remaining originated from nine countries. Almost half (n=17) were conducted in the UK, with some originating from the rest of Europe (Ireland n=2, Pan-European n=1, Netherlands n=1, and France n=1). The USA generated the next most frequent number of papers (n=7). The remainder of the research originated from Australia (n=4), Canada (n=2), the Republic of South Africa (n=1) and one qualitative meta-analysis compared Canada and France (n=1). Types of external peer review under study were accreditation (n=23, 58%), peer review network (n=16, 40%) and those that compared both (n=1, 2.5%).

The settings were mostly hospitals (n=13), or a comparison of multiple specialties (n=8). A few were set in healthcare organisations (n=2), a long-term care organisation (n=1), a primary healthcare organisation (n=1), and a local health department (n=1). Some articles were set in specific settings or contexts: cancer services (n=1), mental health services (n=2), three separate studies in respiratory medicine (n=6), gastrointestinal endoscopy units (n=1), obstetrics and gynaecology hospital departments (n=1), stroke services (n=1), and a paediatric anaesthesia department (n=1). One study related to a community nursing development unit catering to service users with learning disabilities.

### 3.5.2 Research designs

Experimental designs capable of producing evidence of causality were employed in only two separate studies (n=3), one of which included an additional qualitative component ([Rivas et al., 2008](#)) ([Rivas et al., 2010](#)) ([Rivas et al., 2012](#)) ([Roberts et al., 2010](#)) ([Roberts et al., 2012](#)). This included 3 papers from 2 experimental studies (randomised-controlled trials), and 31 papers from observational studies, this included ten mixed methods papers, fifteen qualitative papers, five quantitative papers and one case study. In addition, there were five descriptive papers and one opinion piece.

The earliest research paper was published in 1983, and 68% of the studies had been published in the last ten years, since 2005.

### **3.5.3 Study quality**

The most common study biases were those incurred from cross-sectional designs and descriptive accounts.

Table 3.5 demonstrates a wide variation in study quality, ranging from 20% to 100% scores of the quality assessment tool.

### 3.5.4 Overview of themes

In Table 3.6 I illustrate the main themes generated from the literature review synthesis. This is separated into four categories: the process of change, programme factors, outer non-programme (contextual) factors and inner non-programme (contextual) factors.

**Table 3.6 Themes from systematic literature review**

<b>Process of change</b>	<b>Programme factors</b>	<b>Outer contextual factors</b>	<b>Inner contextual factors</b>	<b>Readiness for change</b>
Ability	Programme structure	Healthcare system	Organisational culture	Multidisciplinary involvement
Key persons	Self-review	Coordinating body	Non-financial resources	Management support
Types of changes	Standards	Economy	Local context	Teamwork
Length of membership	Visit	Mandatory	Geographic location	Communication
Sustaining change	Peer reviewers		Size	
	Feedback			
	Additional approaches			
	Tailoring			
	Additional support			

### **3.6 Results: Aim 1: Summarise what is already known about how change occurs in external peer review to generate improvements in service quality**

To begin with, I sought to address issues of how, who, what ([Zikmund et al., 2012](#)) brought about changes in external peer review. Then, I presented a synthesis of literature on the length of membership required for changes to occur and how changes can be sustained.

#### **3.6.1 Ability of external peer review to bring about change**

The UK National COPD Resources and Outcomes Project (NCROP) was the largest published RCT on a peer review network in the UK. The project was a tripartite initiative from the Royal College of Physicians (RCP), British Thoracic Society and British Lung Foundation, that aimed to evaluate if participation in a peer review network brought about faster changes in service development than the usual mechanisms that operate within the NHS ([Roberts et al., 2010](#)). Only participants from 10% of intervention units, and 16% of control units reported negative changes as a result of participating in NCROP such as staffing, relationships and configuration; where as 74% of intervention sites, and 30% of control sites reported positive service changes ([Roberts et al., 2010](#)). Findings from NCROP suggested that shared learning, experience and materials was a really important causal mechanism of bringing about change through peer review networks ([Roberts et al., 2010](#)). In a qualitative sub-study of NCROP, just under half the intervention group site participants described at least one change that had occurred as a result of participation, that had not been on their change agenda prior to NCROP ([Rivas et al., 2008](#)), and participations from all but one intervention site described generic changes that resulted from NCROP ([Rivas et al., 2012](#)). Overall, participants tended to underestimate the degree of change in their service, and there was great variation in what was labelled as change.

In another RCT which assessed the effects of accreditation on public hospitals' processes and outcomes in South Africa, significant positive change was observed in 20 of 21 elements with sufficient intervention hospitals to make a statistical test; and no meaningful change occurred in any control hospitals

([Salmon et al., 2003](#)). The findings suggested that the changes were attributed to participation in accreditation. Similarly, a more recent qualitative meta-analysis indicated that those undertaking accreditation in Canada and France felt it presented an opportunity to introduce change ([Touati and Pomey, 2009](#)).

Further literature of peer review networks reported on the ability of peer review networks to bring about change. Findings from a quantitative study by Kilsdonk et al. ([2014](#)), to examine the impact of participation and the extent of implementation of peer review network recommendations, provided some indication that participation increased process-related quality of care in the Netherlands. Based on findings from the mixed methods study by Suñol et al. ([2009](#)) to explore the association between peer review network implementation in EU hospitals and their success in meeting quality requirements, implementation was suggested to promote positive change in organisations. Over 70% of respondents in a mixed methods study assessing the effectiveness and value of peer review networks in UK cancer services, agreed that the approach identified service shortcomings and acts as a catalyst for change ([Burnett et al., 2007](#)). The overwhelming view from this report was that the peer review network had helped bring about real service improvement that would not have happened otherwise, but it has also grown and become unwieldy over the years. Respondents provided accounts of service improvements and changes that staff had been trying to bring about before participation, but had previously received no support for. This view was echoed by the descriptive account of the Royal College of Physicians' (UK) National Stroke Services Peer Review Scheme. An expert reported that peer review networks such as this scheme have the potential to act as a catalyst for changes that proved too difficult before ([Bray, 2013](#)).

However, some participants from NCROP sites reported no changes were attributed to participation as they felt they already had service development well underway before joining ([Roberts et al., 2010](#)). Participants from very few sites reported that they were either already so immersed in a culture of change, or so dispirited by previous change failures that they considered the NCROP to



be 'pointless' (this view was twice as common in the control group), or to only have a small impact on services ([Rivas et al., 2008](#)). Some NCROP participants also felt peer review networks lacked the necessary 'bite' to promote change ([Rivas et al., 2008](#)), due to their inherent lack of power to overcome barriers of change ([Rivas et al., 2010](#)). NCROP highlighted the temporal nature of change. For example, an exchange of ideas brought about through a peer review network may later lead to measurable change, but there is limited ability to qualitatively or quantitatively measure or evaluate these effects ([Rivas et al., 2008](#)). Similar findings were reported by the mixed methods evaluation study on how accreditation introduces organisational change in Canadian services by Pomey et al. ([2010](#)). In a particular case, despite the written report containing recommendations, respondents did not consider accreditation to be a driver of change. They considered it a recurrent introspective exercise that instigated or enhanced other quality improvement measures and identified areas where quality ought to be improved. Similar themes were also seen in additional literature on peer review networks. In a qualitative study reporting the experiences of a peer review network in rheumatology, participants expressed the realisation that change does not automatically follow participation, especially in cases where management feel unable to deliver due to financial pressures ([Piper et al., 2006](#)). This was in agreement with a descriptive account of a peer review network of paediatric anaesthesia departments that suggested although the approach identified areas for improvement, making changes was very difficult ([Crean et al., 2003](#)). Despite good intentions to change practice, other time and resource pressures often meant some recommended changes were 'too difficult to do' ([Crean et al., 2003](#), [Butterfield et al., 2012](#)).

One of the survey respondents in the qualitative study by Butterfield et al. ([2012](#)) took a different view to the majority, arguing that peer review networks should take a much more proactive approach. They argued that to be able to drive change, peer review networks need to 'close the loop' on poor performers, and more is needed than the current feedback that is offered. They argued that accreditation is more likely to result in eradication of poor performers.

### 3.6.2 Key persons in making change

A few consultants in the qualitative sub-study of NCROP considered one of the most important facilitators of change was having enthusiastic people responsible for services in both primary and secondary care, who would be proactive and also integrate services ([Rivas et al., 2008](#)). A similar theme was found in accreditation. When exploring the experience of accreditation participants, Greenfield et al. ([2010](#)) reported that motivated staff tend to think and participate beyond natural silos within healthcare organisations in Australia. But the literature uncovered opposing views about change champions between accreditation schemes and quality improvement programmes. According to Greenfield et al. ([2010](#)) 'change champions' were seen to show engagement and desire towards participating in accreditation, as it offered them learning opportunities and enabled their development. Their experience and sense making was predicted to positively shape the norms and attitudes of their colleagues. Similarly, in a more recent qualitative study to examine stakeholders' perspectives about factors influencing participation of accreditation Hinchcliff et al. ([2013](#)) also suggested that effective implementation at an organisational level may require the support of 'change champions'. However, findings from Rivas et al. ([2010](#)) demonstrated that champions and clinicians with an interest in management did not seem to be associated with more changes in a peer review network. A descriptive account of a peer review network of paediatric anaesthesia units challenged this view by suggesting local hospital leaders could be useful in ensuring feedback was communicated to frontline staff in a suitable and useful manner. The account indicated a champion could help to build a positive reputation among participants, however it was based on expert opinion ([Crean et al., 2003](#)).

NCROP participants also observed that when clinicians, managers or commissioners within a successful collaboration vacated their posts, services sometimes degenerated, change became blocked and changes aimed at improving services were less likely to happen ([Rivas et al., 2010](#)).

A mixed methods study that examined the dynamics of change that operated in accredited hospitals suggested that changes during self-review were made primarily by those lower down or working in less prestigious structures. These same people also expected the most from self-assessment as a potential tool for organisational change ([Pomey et al., 2004](#)).

### **3.6.3 Types of changes**

According to the qualitative sub-study, over half of all intervention sites reported that NCROP 'pushed' forward improvements that were already planned or in progress. The term 'galvanised (into action)' was used by several sites, and metaphors of speed, such as 'accelerated processes', 'sped up changes', 'change that would have been slower' without NCROP were reported ([Rivas et al., 2008](#)). In NCROP, a shift in thinking that led teams to feel comfortable about and value the exchange of ideas between sites was the most commonly cited generic change resulting from participation in the peer review network ([Rivas et al., 2012](#)). Such exchange was said to help sites develop plans for quality improvements by drawing on models that their paired site had used or recommended. Although such exchanges often occurred only as one-off events, during peer review visits, data suggested that the experience encouraged teams to consider this a useful strategy that had the potential to be sustained subsequently.

Although few sites had stayed in touch after the peer review visits to network, mentor each other or continue to exchange ideas; one consultant reported how continued networking led to additional improvements ([Rivas et al., 2012](#)). Additional improvements were also seen to occur in accreditation, based on findings from a mixed methods study by Davis et al. ([2011](#)) which examined the extent to which 48 accredited Local Health Departments (LHDs) in the USA conducted accreditation activities. 67% reported conducting quality improvement activities after achieving accreditation, processes included the Institute for Healthcare Improvement's (IHI) model for improvement, Lean, Six Sigma, and quality improvement tools such as Pareto charts. In a qualitative study on accreditation of inpatient mental health services, reflection on practice

was reported to have provided inspiration for further changes to the organisation and structures on the ward that were not accreditation standards, but ones which the ward itself had recognised as weakness ([Baskind et al., 2010](#)).

Approximately one third of generic changes were not explicitly labelled as change by participants of NCROP ([Rivas et al., 2012](#)). This indicated that the changes are not always recognised by participants of external peer review.

According to Duckett ([1983](#)), the area of physical facilities and safety, because of the relative ease of changing these environmental aspects displayed great change. It was suggested by the authors that this area could have also received attention to provide overt demonstration of the hospital's desire to comply with key accreditation standards.

#### **3.6.4 Length of membership needed for change to take place**

Change diaries from NCROP indicated that developments were at differing stages of completion including: further discussions, submissions of business cases, cases being accepted and change having actually taken place. It is thus possible that the quantitative evaluation at one year was too early to detect difference in completed changes in service provision, where previously the pace of change within the NHS has previously been documented as requiring more than one year to see service developments ([Roberts et al., 2010](#)). Interview participants also acknowledged this, ([Roberts et al., 2012](#)) ([Rivas et al., 2010](#)), thus a re-evaluation was carried out to assess if any changes that had not been demonstrated at 1 year may become apparent over a 3-year period. During the re-survey, some units changed from partially meeting or not meeting an indicator in 2007 to meeting it in full in 2010, while some units changed in the opposite direction ([Roberts et al., 2012](#)). Participants from 100% of intervention and 95% control units reported at least one change for the better since 2007 in further change diary returns in 2010 ([Roberts et al., 2012](#)). However, specifically with regard to the quantitative impact of a peer review

network, there was only slight evidence at 3 years of a significant divergence between the intervention and control group ([Roberts et al., 2012](#)).

Similarly, the authors of the RCT to assess the effects of accreditation in South Africa, Salmon et al. ([2003](#)) felt the trial may also have been too short to capture the outcomes of accreditation. The first quality-indicator survey occurred, on an average, 10 months after the COHSASA baseline survey in intervention hospitals. It is possible that hospitals had already made considerable progress that was not captured because the first round of the survey was too late to be a true baseline. This may explain the lack of effect of accreditation on the selected quality indicators. A Canadian study corroborated this theme and showed that changes within organisations differed according to the number of years of participation in accreditation ([Lemieux-Charles et al., 2000](#)). Changes varied according to whether an organisation was in its first accreditation cycle, had experienced several cycles, or had participated in accreditation for over 10 years.

In a mixed methods study to examine the dynamics of change in accreditation, changes in attitude during accreditation did not come about after just a few months, and a climate of trust had to take root in order to maintain this capacity for reflection ([Pomey et al., 2004](#)).

### **3.6.5 Sustaining changes**

One respondent partaking in a peer review network of cancer services noted that activity peaked for the visit and then trailed off as there was no mechanism for maintaining activity between review cycles ([Burnett et al., 2007](#)). In this sense, peer review networks were considered weak by participants from NCROP as they do not equate to 'must do' targets, thus managers and commissioners do not have an obligation to follow through ([Rivas et al., 2008](#)). Participants reported that the NCROP design did not give anyone power to ensure that changes were sustained, or even carried out at all.

However, the same can also be said for accreditation as there were reported failures of self-review teams to remain systematically functional between accreditation visits ([Touati and Pomey, 2009](#)). As services undertaking accreditation provide limited information on a regular basis, between cycles, sufficient to satisfy the accrediting body; this has led some to argue that accreditation awards should be for shorter periods of time to ensure continual monitoring to maintain compliance with essential standards ([Kern, 2002](#), [Hurst, 1997](#)). To challenge the assumption that changes are not sustained once the visit is complete, Lemieux-Charles et al. ([2000](#)) assessed whether teams preparing for accreditation made more changes than teams that had been accredited in the past two years. The findings did not support this view, and statistically significantly indicated that teams that had been accredited in the past 18 months had made more (2.5 changes on average) compared with prospective teams (1.5 changes).

According to experts and health professionals in a qualitative study, flexibility of an accreditation programme to adapt to the changes in its environment and accommodate the feedback of different stakeholders could ensure its sustainability and relevance. It was suggested that openness to changes could transfer the programmes into a learning organisation which always incorporates feedback in its development process and stays up-to-date ([Jaafaripooyan et al., 2011](#)).

In a survey distinguishing how organisational attributes of accreditation differ between low and middle income countries (LMICs) and higher income countries (HICs). Findings suggested that HIC accreditation programmes are more likely to be linked or associated with government, which could enable them to overcome barriers to sustainability ([Braithwaite et al., 2012](#)).

### **3.7 Results: Aim 2: Investigate which programme factors are important for generating improvements in service quality in external peer review**

For the purpose of this review, I defined programme factors as intrinsic parts (stages) of external peer review programmes.

#### **3.7.1 Programme structure**

The survey by Edwards ([2011](#)) to determine the association between peer review network programme factors associated with higher subjective quality impact with better performance, provided significant evidence that well designed peer review networks can improve quality and patient safety in the USA. This was echoed in accreditation schemes. Respondents from a multiple case study evaluating how accreditation helps to introduce changes in Canada considered that highlighting problem areas helped the institution set priorities and accelerate implementing change because of the pre-determined structure of accreditation, which required participants to answer to the accrediting body regarding matters where change was expected ([Pomey et al., 2010](#)). In a specific case highlighted by the authors, staff created a template to monitor changes that were required and changes that were implemented following the hospital's most recent accreditation feedback report. This exercise enabled them to link accreditation standards to changes actually made.

Action plans are planning documents which list the resources and time required to achieve certain goals. These were considered important in structuring change in NCROP ([Rivas et al., 2008](#)). Findings showed that in 2007, the intervention units were asked to prepare service development plans to achieve change based on peer review visits. In 2010, 80% of the data was received back from these intervention units, and action plan themes predominantly related to changes that were made. Most aims were fully achieved in at least a third of units, and partially achieved in a similar percentage ([Roberts et al., 2012](#)).

### 3.7.2 Self-review

Throughout the extant literature, the importance of self-review was repeatedly highlighted. When analysing the impact of accreditation on Australian hospitals, changes which were due to accreditation were all reported to be initiated prior to the visit by interview participants ([Duckett, 1983](#)). The preparatory period before the visit was said to have seen the most work and change. This was also observed in a longitudinal explanatory single-case study conducted in France to examine the dynamics of change that operated following preparations for accreditation. 69.6% of those surveyed believed that irreversible changes occurred at the level of the hospital in preparation for accreditation ([Pomey et al., 2004](#)). The changes in question were concerned primarily with the introduction of changes in practice (30.4%) and the implementation of new changes in the organisation of care (3%). Only once the self-review was completed, staff at the strategic and operational levels recognised the importance of the changes to be made and the necessity of implementing major changes. Pomey et al. (2004) reported that during self-review, tacit learning took place at the individual and institutional levels. Professionals acquired new models of thought (i.e. new vocabulary, the development of a sharper sensibility for the needs of patients and their families, the discovery of self-review, and an awareness of the interdependence between professionals and departments). According to the authors, self-review was one of the most encouraging moments for implementing change in accreditation, albeit with a variable impact on the different sections of the organisation and the different professionals involved.

This was also echoed in a mixed methods study in Canada, whereby it was clear from respondents that the most important changes implemented during the accreditation cycle had been identified during self-assessment ([Pomey et al., 2010](#)). The study showed that the self-review phase lent itself well to self-reflection and the identification of problem areas. This was the phase that built consensus for the changes that the institution saw as most important and most legitimate. Valori et al. (2013) highlighted the importance of preparation for an



accreditation visit in a descriptive account, and argued that programme should do more to support services during this phase of the process.

The qualitative sub-study of NCROP ([Rivas et al., 2008](#)) found that if participants were aware their services were about to be extensively reviewed, the action of undergoing self-review could lead to change through self-reflection. This was echoed in the descriptive account of hospitals participating in a peer review network of stroke services, where it was suggested that the impending visit stimulated positive changes in pathways and services before the review itself ([Bray, 2013](#)).

It was also noted that differing interpretations of self-review by different teams could act as a possible barrier to achieving success in accreditation, according to a mixed methods study exploring hospital accreditation as a quality tool ([Doyle and Grampp, 2014](#)). Some service providers also highlighted the negative impact that self-review could have on other tasks such as contact time with patients ([Doyle and Grampp, 2014](#)); ([Rivas et al., 2008](#)); ([Davis et al., 2011](#)) ([Touati and Pomey, 2009](#)).

### 3.7.3 Standards

In a qualitative study from the UK, survey respondents outlined that peer review networks have the potential to drive changes resulting in ongoing quality improvement by increasing standards year on year ([Butterfield et al., 2012](#)). According to respondents of a mixed methods study to assess value and effectiveness of peer review networks, it was undoubtedly felt that standards helped to drive changes in UK cancer services to meet the Calman Hine recommendations and the Improving Outcomes Guidance ([Burnett et al., 2007](#)). However, 46% of respondents felt that the standards were not flexible enough to provide fair comparisons, and 57% thought that the standards did not give a true reflection of a team's performance. 67% of respondents believed the number should be reduced. These views were repeated in workshops and interviews with respondents.

The standards developed for accreditation in both Canada and France were reported to have helped health care organisations select domains to improve ([Touati and Pomey, 2009](#)). Focus group and interview participants commonly suggested that implementation was best enabled by standards focused on issues directly relevant to the daily activities of frontline health professionals, in an examination of Australian stakeholder perspectives of factors influencing accreditation participation ([Hinchcliff et al., 2013](#)).

### 3.7.4 Visit

According to a mixed methods Canadian study by Pomey et al. (2010), the accreditation visit phase resulted in relatively few changes, except when accreditors pointed out deviations to regulations or when security was at stake .

NCROP participants generally reported that visits were more successful when focussed on one specific topic (Roberts et al., 2010). However, in interviews with key informants and staff participating in the accreditation of small and community hospitals in the UK, when peer reviewers concentrated on one part of hospital services at the expense of others, it was reported to threaten the validity of accreditation and alienate staff groups who had prepared but were unable to meet peer reviewers (Hurst, 1997). Visits were reported to offer poorly performing teams with the opportunity to see models of good practice and learn from more developed services. Respondents reported that this encouraged (rather than demoralised) teams; and better performing hospitals were in turn encouraged to maintain high quality services (Rivas et al., 2008, Rivas et al., 2010). The opportunity to reflect in light of others' experiences helped to see where improvements were needed (Roberts et al., 2012). Many teams reported that visits were a validating and reassuring experience that showed they were 'not alone' in problems they faced, and also that they were doing well (Rivas et al., 2012).

NCROP participants described the hospital walk through as one of the most enlightening parts of the visit. It was reported to be a learning experience that enabled best practices to be seen and copied; that would not have been picked up on otherwise (Rivas et al., 2008). Most participants reported to have liked to have spent more time on wards and in clinical areas. However, a minority thought although the experience was interesting, it did not necessarily inform the process. According to a descriptive account, including a walk through the integrated care pathway in a quality network for stroke services visit enabled peer reviewers to meet staff who worked outside the stroke unit and were crucial to its function, such as staff in A&E and radiology (Bray, 2013).

According to a descriptive account of accreditation for endoscopy units in the UK, 'meet and greet' and 'presentation' activities at the start of some visits acted as ice breakers, facilitating rapport between the host service and peer reviewers. These activities provided the host team with an opportunity to showcase areas of pride, give an honest account of obvious shortcomings, or even highlight areas where they needed support or advice ([Stebbing, 2011](#)).

Conversations between peers that took place during tea breaks, lunch or visits were reported to enable participants to share good practice ([Roberts et al., 2010](#)) and exchange ideas ([Page and Harrison, 1995](#)). According to a qualitative study of respiratory units, this was usually characterised by sharing 'tricks of the trade' with peers in the same specialty and discussing how services and training could be improved. Participants from NCROP reported staff with similar roles could converse across organisations ([Rivas et al., 2010](#)). Although such exchanges often only occurred during one-off events or visits, change diaries suggested that the experience encouraged teams to consider this a useful strategy outside external peer review, that had the potential to be sustained subsequently. In a case study of accreditation of a nursing development unit in the UK, the sharing of experiences and collaboration over professional issues was seen to influence the development of good practice ([Balkizas, 1995](#)). The relaxed exchange of ideas was considered as equally productive as the peer review visit itself. Surveyed health professionals participating in a peer review network of rheumatology services, found networking was conducted in a positive learning environment where strengths could be transferred between units; they also found it very helpful in promoting multidisciplinary team working ([Piper et al., 2006](#)).

According to a descriptive account, service user interviews during peer review network visits provided revealing insights into the cohesion of stroke teams ([Bray, 2013](#)). However, service user representation was criticised by one respondent participating in a peer review network of cancer services. The respondent suggested when single hand-picked service users are chosen for

interview during a visit, they cannot accurately represent service user experience ([Burnett et al., 2007](#)).

### **3.7.5 Peer reviewers**

Interviewees from a qualitative study on peer review networks by Butterfield et al. ([2012](#)) highlighted that peer reviewers should have the authority to be able to challenge poor performance and not be willing to collude on poor practice. One interviewee argued that where services are not providing adequate care, reviewers need to hold 'bruising' reviews and prove resilient to ensure that changes are made.

During a descriptive account of endoscopy accreditation in the UK, preparing and monitoring a skilled team of peer reviewers was viewed as pivotal to achieving consistency and credibility of accreditation ([Stebbing, 2011](#)). The account highlighted the risk of reputational damage that could be caused by poor or inconsistent assessors. Greenfield et al. ([2010](#)) highlighted the importance of perceived skill and credibility of peer reviewers in Australia, which was thought by interviewees to be carefully scrutinised by the host service undergoing accreditation. Interviewees also reported when staff were not given recognition by peer reviewers, it acted as a barrier to services remaining engaged in accreditation.

There was a majority agreement that service user representation in NCROP had been tokenistic and problematic ([Rivas et al., 2008](#)). Service users demonstrated loyalties to their own service by over-commenting on their own service instead of the host service, or were seen to use the visit as an opportunity to find out more about what was available elsewhere. In a study evaluating the role and participation of lay peer reviewers in Ireland, stakeholders considered the inclusion of service user peer reviewers to be beneficial in terms of lending credibility and balance to the accreditation process and as a catalyst for driving changes and improvement ([O'Connor et al., 2007](#)). However, in a blinded, random, stratified study looking at the association between accreditation and organisational performance in Australia, consumer involvement was not

significantly associated with any organisational characteristic ([Braithwaite et al., 2010](#)).

During the qualitative study of a rheumatology peer review network in the UK, although the health professionals that were interviewed reported being comfortable to be reviewed by those they knew well, this was seen to affect objectivity ([Piper et al., 2006](#)). This was echoed in the descriptive account of a peer review network of paediatric anaesthesia departments where it was noted that near-neighbouring peer reviewers could be problematic, objectivity could be lost, and according to one particular view, 'inconsistent' with competition in the NHS ([Crean et al., 2003](#)).

### **3.7.6 Feedback**

Findings from NCROP suggested that feedback through peer review networks was an important mechanism through using external validation as a negotiating tool ([Roberts et al., 2010](#)). According to staff from respiratory units in a qualitative study of a peer review network, reports in many instances were used to support claims for additional resources, facilities or staff ([Page and Harrison, 1995](#)). Similarly, reports as part of the accreditation of inpatient mental health services' were said to provide a platform to negotiate more successfully with senior managers to gain further resources for wards ([Baskind et al., 2010](#)). The view of one Director of Nursing in a qualitative study analysing the impact of accreditation on Australian hospitals by Duckett ([1983](#)) argued that accreditation gave good support for what staff had been asking for, and nursing staff seized this as an opportunity for change.

Changes that were a direct consequence of peer reviewer recommendations was another key theme from NCROP ([Roberts et al., 2010](#)). Written feedback was considered important in framing improvement plans according to an expert opinion of the peer review network of stroke services in the UK ([Bray, 2013](#)). According to the qualitative meta-analysis by Touati and Pomey ([2009](#)), in both the French and Canadian accreditation systems, the principal means of accomplishing changes have been through recommendations. After receiving

the written report for accreditation, Canadian organisations were reported to have responded to report recommendations to achieve accredited status ([Pomey et al., 2010](#)). However, interviewees describing the impact of accreditation on Australian hospitals expressed that the report itself was only described as the main instrument of change in hospitals that had either failed to be accredited or where medical matter had been the focus of the feedback ([Duckett, 1983](#)).

According to the qualitative study reporting the experience of a peer review network of rheumatology services, there was no clear consensus on whether the report was viewed seriously by UK Trusts ([Piper et al., 2006](#)). Opinions were also divided about allied health professionals not having access to the report and not being involved in implementing pertinent recommendations.

### **3.7.7 Additional quality improvement activities**

Additional quality improvement activities were offered by organisations who coordinated external peer review. According to the expert opinion provided by Worrall ([2011](#)), additional activities such as conferences provided an important space for staff whose specialist services were often isolated, to come together.

During the national hospital accreditation programme in Zambia ([Bukonda et al., 2002](#)), educational surveys were the strongest feature of the accreditation programme. These were essentially visits with no accreditation decision made at the end (thus comparable to peer review network visits), to familiarise the hospital with accreditation standards and enable staff to appreciate how their hospital functions. They provided an opportunity to receive constructive suggestions on how to achieve accreditation, not just a report card of inadequate performance.

### **3.7.8 Tailoring**

Differing views were expressed about whether external peer review programmes should offer consistency across the country, or whether peer reviewers should be flexible in their interpretation of the evidence provided

([Burnett et al., 2007](#)) ([Crean et al., 2003](#)). According to the mixed methods study on a peer review network of cancer services, it was suggested with some standards it is possible to be precise, with others less so. Participants felt if peer review networks are used to provide a national picture for monitoring, then consistency is vital. If however, the purpose is quality improvement, then flexibility is acceptable ([Burnett et al., 2007](#)). Another argument was that those being peer reviewed should determine the focus of the review, so they can achieve maximum benefits. When appropriate and feasible, adjustments in the Voluntary Review of Quality Care (VRQC) programme were made on the basis of comments made by hospital representatives and reviewers, this led to success in the peer review network ([Gibbs and Cheetham, 1988](#)). With regards to accreditation, some experts felt high performing services that delivered consistently excellent outcomes should earn autonomy from cycles ([Valori et al., 2013](#)). Some participants from hospitals undergoing accreditation in Zambia indicated the need for tailored technical assistance for implementing changes ([Bukonda et al., 2002](#)).

### **3.7.9 Additional support**

During the qualitative sub-study of NCROP, it was revealed that staff often took on extra duties or worked longer hours for no additional pay. More than half of control and intervention sites interviewed reported using 'goodwill' as a key mechanism to improve quality. Some sites were reliant on this: 'If I were to withdraw goodwill, then there's no services at all.' Others reported that it should only be used temporarily, but longer term use could lead to problems ([Rivas et al., 2008](#)).



### **3.8 Results: Aim 3: Identify the non-programme (contextual) factors which facilitate or impede their success**

There are a complex set of outer and inner contexts at play in external peer review, and it is difficult to single out impact ([Kilsdonk et al., 2014](#)). Overall it has been found that inner and outer contexts (non-programme factors) can facilitate or impede the success of external peer review.

#### **3.8.1 Outer contexts**

##### **3.8.1.1 Healthcare System**

Findings from the NCROP re-survey suggested that the healthcare system facilitated changes through modifications in commissioning, the QIPP agenda (Quality Innovation, Productivity and Performance) as well as specific national guidelines requirements and recommendations ([Roberts et al., 2012](#)). Service re-design was seen by participants to act as a barrier that precluded effective service delivery. NCROP was undertaken during a time of great reform in NHS, which may have influenced study results as structural changes altered the culture of teams ([Rivas et al., 2008](#)). Findings from NCROP also suggested that peer review networks were associated with more changes in hospital based services than community based services ([Roberts et al., 2010](#)).

In a quantitative study to determine the association between accreditation and self-reported clinical performance and ratings of organisational performance, there was no significant difference between accreditation ratings between public and private sector organisations ([Braithwaite et al., 2010](#)). This finding was unique, and not present elsewhere in the extant literature. A separate quantitative study describing the differences in organisational attributes of accreditation between LMICs and HICs observed that a change in government or a lack of legal standing could act as a barrier to successful or sustainable accreditation ([Braithwaite et al., 2012](#))

Participants during the qualitative sub-study of NCROP reported how particular clinical conditions could 'lose out' if not considered higher priority by the government ([Rivas et al., 2010](#)). In the UK, a lack of acute sector medical focus

was also seen as a barrier to achieving success through peer review networks in cancer services ([Burnett et al., 2007](#)).

In the USA, the Medicare social programme was seen as a major, but indirect, facilitator of achieving success through accreditation. This is because hospitals must be accredited by the Joint Commission or undergo regulatory review by the Centre for Medicare & Medicaid Services (CMS) to be able to participate in Medicare. As Medicare accounts for 40% of hospitals' revenues, it provides a strong incentive to become accredited. However, the legal culture of blaming individuals rather than systems ([Devers et al., 2004](#)) and serious disagreements between the department and external entities such as the health department was seen to hinder success of accreditation ([Gibbs and Cheetham, 1988](#)).

#### 3.8.1.1.1 Other quality improvement programmes

Some healthcare stakeholders in an Australian qualitative study examining perspectives of factors that influence participation of accreditation argued that when accreditation is aligned with other regulatory programmes this can facilitate success ([Hinchcliff et al., 2013](#)). However, where multiple programmes aimed at quality improvement are operating in the same system but are not well coordinated, this may lead frontline staff to disengage, according to participants who have reported their previous experiences ([Crean et al., 2003](#)) ([Doyle and Grampp, 2014](#)) ([Bukonda et al., 2002](#)).

During the qualitative sub-study of NCROP, some clinicians considered audit and pilots to be a 'necessary evil', whilst others showed a cultural shift to evidence-based change that aligned them with managers. Audits were described as 'very powerful levers for funding' and even considered to be critical tools for change within the peer review network. They were extensively used by teams that were more proactive in achieving change or had learned how to work in harmony rather than tension with primary care trusts. Successful teams explicitly stated that an audit approach set them apart from less successful respiratory departments ([Rivas et al., 2010](#)). Successful change was also associated with the appropriate use of audit data within organisations enabling

robust business cases for change to be presented by clinicians to their acute trust managers that recognised the power of cost-effectiveness and cost-benefit arguments. However, the same benefits were not seen across organisations ([Rivas et al., 2010](#)). According to Worrall ([2011](#)), standards used in College Centre for Quality Improvement (CCQI) external peer review programmes specify that clinical audit should be conducted and support should be offered.

Worrall ([2011](#)) also provided the expert opinion that services may have greater influence within their organisations and with commissioners if they are accredited or are participating in a peer review network .

### ***3.8.1.2 Coordinating body***

In the UK, in a descriptive account assessing the process of a peer review network, service providers felt external peer review services should not be run by a regulatory body such as the Care Quality Commission (CQC) ([Crean et al., 2003](#)), as it could lessen the opportunity for exchange of ideas and discussion, which was seen as important to participants ([Page and Harrison, 1995](#)). For similar reasons, there was unhappiness reported about devolving management of Trent Small Hospital Accreditation Scheme (TSHAS) to NHS Trusts, in fear of programmes dissipating, losing robustness and credibility. Key informants and staff participating in TSHAS felt that it may similarly alter the focus from developmental to inspectorial ([Hurst, 1997](#)).

The descriptive account by Crean et al. ([2003](#)) provided arguments both for and against Royal Colleges in the UK coordinating external peer review. The account detailed cases where external peer reviews conducted by Royal Colleges have uncovered evidence of poor practice but no action was taken to inform regulators. This could suggest Royal Colleges lack sufficient powers; however, it was also argued that with appropriate contracts on terms of reference, transparency and dealing with poor performance, Royal Colleges remain suitable coordinating bodies.

### **3.8.1.3 Economy**

Findings from NCROP indicated that funding changes acted as a barrier that precluded effective service delivery ([Roberts et al., 2012](#)). This was echoed by a survey comparing the organisational attributes of accreditation between LMICs and HICs, where it was suggested that a premature end to core funding by international donors could act as a barrier to successful and sustainable accreditation for LMICs in particular ([Braithwaite et al., 2012](#)).

According to stakeholder interviews involved with Joint Commission accreditation in the USA, an absence of strong market incentives to improve, and pressures caused by a difficult economic environment could act as barriers to achieving success through external peer review ([Devers et al., 2004](#)). This could cause healthcare organisations to cut back or stop participating in external peer review programmes, even when they have been members for some time, according to a mixed method evaluation in Canada ([Pomey et al., 2010](#)). Limited financial resources were also reported by participants as the biggest reason for not participating in accreditation ([Brasure et al., 2000](#), [Devers et al., 2004](#), [Doyle and Grampp, 2014](#), [Gibbs and Cheetham, 1988](#)).

### **3.8.1.4 Mandatory**

In a longitudinal explanatory single-case study of accreditation in France, despite being mandatory at the level of the organisation, the hospital's obligation to embark on accreditation was not accompanied by an awareness of how much accreditation could eventually serve as an agent of change ([Pomey et al., 2004](#)). In this study, pressure from mandatory accreditation was reported as having increased supervision and administrative checks in some instances, resulting in a potential source of conflict, causing professionals not to buy into the programme ([Pomey et al., 2004](#)).

According to a qualitative study describing peer review networks ([Butterfield et al., 2012](#)), some respondents argued that providers are more likely to respond if they request peer review networks themselves. Despite the general consensus that participation should be voluntary, some respondents felt there should be more tools available to encourage participation. This was felt to be important, as some described that often poorer performing services are also less likely to volunteer to participate in external peer review. They also felt highly performing providers may be more incentivised as participation can provide a mechanism to demonstrate the quality of their service ([Butterfield et al., 2012](#)) ([Crean et al., 2003](#)).

### **3.8.2 Inner contexts**

Inner contexts may facilitate or impede external peer review programmes.

#### ***3.8.2.1 Organisational culture***

According to stakeholders of an accreditation programme in Australia, implementation may be more effective when programme aims, requirements and benefits are conceptually unified and articulated using language and formats that appeal to the cultures and normative practices of different professional groups ([Hinchcliff et al., 2013](#)). In a quantitative Australian study by Braithwaite et al. ([2010](#)), a positive correlation was found between organisational culture and accreditation performance ( $r=0.618$ ,  $p=0.05$ ). In another Australian study, participants described that a promoted shared desire for improvement and a quality and safety culture can enhance accreditation, as participation contributes to development of a collaborative organisational culture ([Greenfield et al., 2010](#)). In Canada, where many mergers have occurred, accreditation also has facilitated the creation of a new organisational culture by stimulating staff from different sites to collaborate in the programme ([Touati and Pomey, 2009](#)). According to findings from another mixed methods study in Canada ([Pomey et al., 2010](#)), an organisational culture that was considered to be open to change was cited as an important condition for the implementation of change in one particular case. Where the hospital encouraged a high degree of autonomy, this was seen to facilitate the implementation of change during accreditation. A quantitative study examining the differences between Joint Commission accredited and non-accredited facilities in the USA observed the process of preparing for accreditation can create a quality-oriented culture ([Kern, 2002](#)).

Similar themes were found in literature of peer review networks. Participants described a number of successful respiratory teams participating in NCROP were enthusiastic and motivated with strong cultures of commitment to service improvement ([Rivas et al., 2008](#)). In a quantitative study of quality improvement programmes in the USA, adverse organisational culture was associated with lower quality ([Edwards, 2011](#)).

### **3.8.2.2 Non-financial resources**

Non-financial resources such as building and site characteristics ([Doyle and Grampp, 2014](#)), staff ([Davis et al., 2011](#)) ([Roberts et al., 2012](#)), facilities and equipment ([Doyle and Grampp, 2014](#)) and good IT services ([Rivas et al., 2008](#), [Devers et al., 2004](#)) were reported to influence the ability of external peer review programmes to bring about change.

### **3.8.2.3 Local Context**

Worrall ([2011](#)) provided an expert opinion that services that are undergoing major structural change will find it hard to adopt new practice, and subsequent change following participation to external peer review. In addition, services that are struggling, e.g. due to staffing problems or unfeasibly high caseloads, are unlikely to prioritise implementation of clinical guidelines, as they are often focused on survival mode.

Differences in context can mean even in standardised programmes, effects are likely to vary. In the RCT to assess the effects of accreditation in South Africa, it was suggested that local context could impede the role of accreditation in improving service quality ([Salmon et al., 2003](#)). On occasion, in NCROP, barriers were reported around difficulty in communicating with service users, possibly because of ethnicity and language barriers ([Rivas et al., 2008](#)).

### **3.8.2.4 Geographic Location**

Previous research of Joint Commission accreditation in the USA has highlighted differences in levels of participation in rural and urban areas ([Kern, 2002](#)) ([Brasure et al., 2000](#)). Reasons for these differences were unclear, but it is possible that logistical factors, such as travel requirements make it more difficult to deliver external peer review programmes to more remote services. Patterns varied with rurality of area, geographic location and time in a mixed methods study by Brasure et al. ([2000](#)). Furthermore, it was observed that in rural settings, there are not sufficient numbers of physicians in the same specialty to review their peers. However, in a contrasting quantitative study conducted in Australia, there was no significant difference found between

accreditation ratings of organisations in different locations; namely metropolitan, regional and rural locations ([Braithwaite et al., 2010](#)).

#### **3.8.2.5 Size**

In a qualitative study to describe VRQC accreditation in the USA, the size of both the hospital and the obstetrics and gynaecology department played a role in requests for external review ([Gibbs and Cheetham, 1988](#)). Members of the review team and hospital representatives felt that hospitals with small to moderate-size teams frequently had difficulty developing performance standards and enforcing compliance. Relative to hospitals, primary care organisations are smaller, but still required two or three days of peer reviewers' time.

Although interviewees from some large hospitals in the USA believed they had advantages ([Devers et al., 2004](#)); the quantitative study by Braithwaite et al. ([2010](#)) in Australia showed no significant difference between accreditation ratings of small, medium and large organisations. This was also seen in a descriptive account of a peer review network of respiratory units in the UK ([Page and Harrison, 1995](#)). Apart from recommendations for more staff, single units received no more recommendations than larger units.

#### **3.8.3 Readiness**

Kilsdonk and colleagues ([2014](#)) concluded from their quantitative study on hospital data in the Netherlands, that quality focussed hospitals participating in a peer review network are more likely to work on continuous quality improvement and behave as early adopters.

In a case highlighted by Pomey et al. ([2010](#)), to evaluate how accreditation in Canada helps to introduce organisational changes, managers and professionals that were young and dynamic were cited as an important condition for the implementation of change.



### **3.8.3.1 Multidisciplinary involvement**

Quantitative findings from a mixed methods study in Canada indicated staff not directly involved in accreditation had significantly different perceptions ([Paccioni et al., 2007](#)). The accreditation dynamic was limited to those involved, through teams and various committees formed. Process and final outcomes were reported as not understood or as well absorbed by staff who did not participate.

Weak, unengaged and uninformed medical participation was observed across the extant literature ([Touati and Pomey, 2009](#)) ([Duckett, 1983](#)) ([Burnett et al., 2007](#)). When doctors participated in Canadian accreditation, only a few interested were reported to have taken part ([Pomey et al., 2010](#)). French hospital physicians' input was reported to have diminished in self-review over time without them actually boycotting the programme ([Pomey et al., 2004](#)). Interview accounts of Australian accreditation described that medical staff were usually only involved after a visit when the report contained recommendations which affected them ([Duckett, 1983](#)). Respondents noted that the areas that showed the least change were those associated with medical staff. An RCT assessing the impact of a hospital accreditation programme in South Africa indicated doctors' behaviours were very difficult to change ([Salmon et al., 2003](#)). It was suggested that hospital staff found it easier to implement standards that did not require doctors' involvement, and these were reported to have changed first. However, there was also a significant inverse correlation between nurses' perception of their participation in decision making and final accreditation scores in this study. This suggested when doctors and managers led preparation for accreditation, they experienced better success rates.

In both mixed methods studies ([Pomey et al., 2004](#)) ([Pomey et al., 2010](#)), in instances where clinicians were shown to disengage, participants reported this was due to more of an organisational than professional matter, as clinicians saw accreditation to be in the realm of nurse managers than their own. Nurse managers were mainly more involved, with department physician heads to a lesser extent ([Pomey et al., 2004](#)). Nurse managers were also reported as the

most involved ([Pomey et al., 2010](#)) and motivated to take part on account of their assumed leadership roles ([Pomey et al., 2004](#)).

Differences of opinion between views of administrators and professionals ([Paccioni et al., 2007](#)) led to observed tension and power struggles ([Pomey et al., 2004](#)). Previously stymied teams were the most likely to report external peer review as helpful in reducing the divide between acute trust managers and their clinicians in NCROP ([Rivas et al., 2010](#)). Some participating nurses described systemic or technical changes that their consultant colleagues were not aware of, at the five sites where both were interviewed ([Rivas et al., 2012](#)). Similarly, there was not always consensus between consultants in the same rheumatology unit in another UK peer review network study as to whether recommendations had been acted on ([Piper et al., 2006](#)).

### **3.8.3.2 Management support**

Institutional commitment to quality and patient-safety improvement was reported by interviewees to have an important impact on the success of external peer review ([Devers et al., 2004](#)) ([Pomey et al., 2004](#)). According to the expert opinion provided by Worrall ([2011](#)), clear support from the organisation's board and senior executive were considered to facilitate the success of external peer review. According to stakeholders involved in accreditation in the USA, board management and clinical leadership were especially important, as resources and cooperation were more likely to follow ([Devers et al., 2004](#)). Descriptive accounts of a peer review network of paediatric anaesthesia departments in the UK described how poorest performers in the peer review network also lacked clinical and managerial support ([Crean et al., 2003](#)).

In a descriptive account of UK endoscopy accreditation visits, there is an expectation that senior management are well represented during verbal feedback. If senior management are not present, it indicates poor organisational culture and support for the endoscopy service ([Stebbing, 2011](#)). This was echoed with findings from a mixed methods study of USA accreditation where

lack of recognition from direct managers or supervisors suggests accreditation is an independent activity rather than an integral part of system ([Doyle and Grampp, 2014](#)).

### **3.8.3.3 Teamwork**

Good team working within the respiratory departments was considered to be critical for changes to work, and depended on individuals in teams, according to NCROP participants ([Rivas et al., 2008](#)) ([Rivas et al., 2010](#)). Teams with a positive attitude to problem-solving were reported to achieve more changes. Participants described how successful collaborations were possible across and within organisations.

### **3.8.3.4 Communication**

According to local project leads of wards that had failed accreditation in the UK, poor communication was viewed as detrimental to effective participation ([Baskind et al., 2010](#)).

### 3.9 Discussion

The results of this review provide a wealth of views and opinions from stakeholders who have designed, delivered or participated in external peer review programmes. In contrast to the number of reports describing external peer review programmes, there has been relatively little systematic research that has specifically examined the context in which external peer review programmes are most and least likely to help services improve the quality of the care they provide. Much of the research focussed on the programme factors of external peer review, and more broadly the benefits and outcomes of external peer review, which this review did not focus on. Included papers used a wide range of study methods; most of which were observational (n=31).

Through data extraction; data, methods and descriptions were used to categorise and tabulate information according to the three original aims. The most frequently mentioned programme factors were self-review, the visit, peer reviewers and feedback. Non-programme (contextual) factors which were reported to influence the success or failure of the programmes were separated into two categories: outer and inner contexts, to be able to inform an initial MRT.

Of the 40 papers identified in this systematic review, the majority investigated accreditation (n=25), 14 papers referred to peer review networks and 1 opinion piece compared different approaches.

The most robust research was conducted by Roberts et al. ([2010](#)) ([2012](#)) on NCROP, a 1-year and 3-year national evaluation of a peer review network COPD in the UK. This was the largest ever voluntary peer review network to be evaluated to date in the UK, and provided much of the data for this systematic review. Findings after 3 years indicated a slight association with improved quality of care, services delivery and changes that promote quality improvement. However, length of membership to be able to bring about changes was considered a salient theme across the literature. As a qualitative

sub-study was conducted in parallel to the main trial, researchers were able to triangulate findings from quantitative data from the main trial and an independent analysis of change diaries, and found the same picture emerged from the three different types of analysis ([Rivas et al., 2010](#), [Rivas et al., 2008](#), [Rivas et al., 2012](#)). Another RCT which assessed the effects of accreditation on public hospitals' processes and outcomes in South Africa, also observed significant positive changes but only in intervention hospitals ([Salmon et al., 2003](#)).

The most important programme factors identified in this review were self-review, the visit and feedback. Self-review, or before the visit, is where most of the change was seen to have occurred for accreditation. Some of the literature suggested that change was not achieved during visits or reports unless something not right or accreditation had been failed. Self-review was seen to promote teamwork and promote a culture of change. However, it was felt by participants that these dynamics were limited mostly to those directly involved. Sharing and learning was seen as an important mechanism that bought about change during visits, enabling staff involved to share good practice. Peer reviewers were considered a crucial programme factor, and their perceived skills, credibility and conduct was referred to frequently in the extant literature. Feedback is where most of the change was seen to have occurred for peer review networks, through using external validation as a negotiating tool. Some of the key studies, such as NCROP, reported direct changes attributed to peer reviewer recommendations received through feedback. The literature uncovered opposing views about change champions between accreditation schemes and quality improvement programmes.

The focus of visits, and more widely standards was a key point of debate in the literature. In interviews with key informants and staff participating in the accreditation of small and community hospitals, when peer reviewers concentrated on one part of hospital services at the expense of others, it threatened validity of accreditation and alienated staff groups who had prepared but were unable to meet peer reviewers ([Hurst, 1997](#)). This was

interesting as it contradicted findings from the NCROP in which visits that focussed on one topic were seen to be more successful ([Roberts et al., 2010](#)).

The most common outer context seen to influence external peer review was the healthcare system. The way healthcare systems' structure, prioritise and incentivise, was seen to heavily influence changes attributed to external peer review. There was much debate in the literature regarding the national coordination of external peer review, and the extent to which they feed into regulatory processes.

The most common inner context seen to influence external peer review was organisational culture. Programmes were seen to create and promote an organisational culture oriented to quality improvement, as well as observed changes facilitated by organisational culture.

Readiness for change, especially the sub-themes of multidisciplinary involvement and management support were key. Involvement in external peer review was linked to understanding, and the literature presented a debate between the involvement of different groups of staff such as nurses, doctors, administrators and managers. Although frequently mentioned, there was little research on how to overcome or assess clinical resistance to change or tension between clinical and non-clinical staff. Management support was also an important predictor and facilitator of success. If management were not seen in feedback sessions, or observed to have an active role by peer reviewers, it was an indication that external peer review was not a central part of the organisation. This would then have further implications on staff understanding, engagement, and ability to request resources and implement changes following recommendations.

The main limitation of this review is the quality and study design of the included papers. As most studies were observational, it was difficult to assign rank to the non-programme (contextual) factors in order of importance. Another

considerable limitation was the heterogeneity of use of the terms: 'external peer review', 'accreditation' and 'peer review network'.

As there were a variety of specialties observed (see Table 3.5) and different combinations of multi-professional teams were used, outcome data could not easily be synthesised or generalised. There was also a great deal of variety in relation to the countries and local contexts where the studies were published. Given the difference in healthcare systems, caution was used when drawing comparisons between programmes from different countries. Having taken that into consideration, it could be argued that the inner non-programme contexts which influence the success of external peer review can be generalised across countries, healthcare systems and specialties.

### **3.10 Conclusion**

In summary, I have demonstrated through this review, that while many external peer review programmes have been set up with the aim of improving the quality of healthcare, few studies have examined the process of how they work to bring about change.

Existing studies have predominantly focussed on accreditation and the outcomes of external peer review, with limited empirical research into the process of change which result in their success or failure, or on the impact of non-programme factors that may influence them. The review findings demonstrate that over the past decade, the number of quality improvement programme studied has increased. The review identified clear literature gaps, and a need for a multi-method approach, providing me with a logical platform to embark on further work in this field.

Although some key mechanisms within programme factors have been identified, further research is necessary to support this from the perspective of stakeholders involved in external peer review, especially frontline staff. Therefore, one of the thesis aims will be to evaluate stakeholder views of the causal mechanisms associated with programme factors that are essential for quality improvement to take place through external peer review.

It was challenging to categorise and review the non-programme factors into outer and inner contexts, as this had not been done previously in this field of research. Further research needs to be done in this area to understand how these differing contexts can influence the success of external peer review. Readiness for change constructs are an important part of the change process, and more research is needed this area. Therefore, research undertaken for the purpose of the thesis will investigate which readiness for change constructs are necessary to effect change in external peer review, with a specific focus on readiness for change.



The modest evidence that exists, highlights the current use of external peer review comprises a patchwork of programmes, some of which have been short lived and of uncertain benefit. There is currently no consistent strategy in place for the use of external peer review to improve and assure standards, and subsequently there has been very little evaluation of the non-programme (contextual) factors which lead to the potential success or failure in a system-wide approach to high-quality care.

The mixed method study that I intend to carry out aims to supplement the information provided by the literature review, with a view to providing a much more complete picture with which to inform future external peer review programmes. Pawson ([2006](#)) has argued that a focus solely on outcomes does little to develop a cumulative understanding of complex social programmes. Accounts of how they work, possibly in terms of mechanisms, can reach a better understanding of how theory may be improved ([Robson, 2011](#)).

## **Chapter 4 Research Context**

This chapter introduces the research context of the Royal College of Psychiatrists' College Centre for Quality Improvement (CCQI), which is one of the largest providers of external peer review programmes in the UK. Within the range of programmes offered, I chose to undertake a realist evaluation of peer review networks and accreditation schemes from forensic inpatient and community-based mental health settings. I conclude this chapter by focussing on the Evaluation of Low Secure Units (eLSU) study, which ran parallel to my PhD, where some of the data used in my thesis was collected.

#### 4.1 Royal College of Psychiatrists – CCQI

The Royal College of Psychiatrists is the professional body that represents the interests of psychiatrists in the UK. One of the main aims of the College is to promote the delivery of better services for people with mental illness. The Royal College of Psychiatrists accommodates a Centre for Quality Improvement (CCQI) that coordinates a range of external peer review programmes ([Barnes and Paton, 2012](#)). The work of the CCQI is funded through a mixture of grants and subscriptions from healthcare organisations, and funding from the Department of Health ([Barnes and Paton, 2011](#)). The CCQI aims to improve the quality of care provided for service users with psychiatric disorders through audit-based quality improvement programmes, external peer review programmes (peer review networks and accreditation programmes) and the Prescribing Observatory for Mental Health (POMH-UK). POMH-UK was set up in 2005, and aims to improve the quality of prescribing practice by seeking to promote and support the optimal, safest use of existing medications in psychiatric practice through focussed, audit-based, quality improvement programmes ([Barnes and Paton, 2012](#)).

According to a cross-sectional survey, 12 out of 39 of the eligible external peer review programmes identified in the UK were coordinated by the CCQI, making it one of the largest providers of external peer review programmes ([Shah et al., 2010](#)). Thus, I chose to undertake realist evaluation on the external peer review programmes coordinated by the CCQI. I specifically selected peer review networks such as the Quality Network for Forensic Mental Health Services (QNFMHS) for Low and Medium Secure Services ([NHS Commissioning Board, 2013](#)), Accreditation for Inpatient Mental Health Services (AIMS) specifically for psychiatric intensive care units (PICUs) as these are also locked wards, and the Memory Services National Accreditation Programme (MSNAP) which are for community-based memory services.

In the programmes, clinicians and experts by experience agree standards of quality for their services, carry out visits of other services and assess against

standards using a combination of quantitative and qualitative evidence ([Davies and Killaspy, 2015](#)). In addition to these standards, all healthcare providers within England are subject to registration through the Care Quality Commission (CQC). Any Department of Health secure service standards must therefore be interpreted alongside the CQC *Specialist Mental Health Services: Provider Handbook* ([Care Quality Commission, 2015](#)).

## **4.2 Forensic Inpatient Mental Health**

UK mental health services have developed substantially over the past decade ([Appleton, 2012](#)). Mental health services can be divided between inpatient and community services. An inpatient mental health service is defined as a unit with hospital beds that can provide 24-hour nursing care. These services can provide care for service users detained under the Mental Health Act 1983 (MHA 1983), with a consultant psychiatrist or other professional acting as responsible clinician. Inpatient mental health services are currently provided by NHS or independent sector providers, and may be located in a hospital campus or a community setting.

Forensic mental health is an area of specialisation that involves the assessment and treatment of services users who are both mentally disordered and whose behaviour has led, or could lead, to offending ([Mullen, 2000](#)). They also provide treatment and care to people who present with serious challenging behaviour, who are often temporarily or indefinitely contained to protect the public ([Sugarman and Dickens, 2015](#)). The forensic pathway aims to promote and enable recovery and independence of the service user whilst ensuring protection of the public. The pathway includes the provision of appropriate levels of physical, procedural and relational security within a range of environments including high, medium and low secure inpatient facilities as well as within community settings that serve the public and the criminal justice system ([NHS Commissioning Board, 2013](#)), ([Department of Health, 2007](#)). The core tasks of secure services include assessment, management and treatment of people, whether in the community, in hospitals or in the criminal justice system,

who present a significant risk of harm. The assessed level of risk will inform the level of secure service required ([Department of Health, 2012a](#)).

According to Rutherford and Duggan ([2008](#)), prisoners and offenders charged with offences are most often transferred directly to high and medium levels of secure forensic hospitals. Inpatient services at each level will have a range of ward functions depending on the nature of the service provided; these may include assessment wards, intensive care, high dependency wards, rehabilitation services (including slow-stream provision for patients requiring longer-term care) and general treatment wards ([Department of Health, 2012a](#)).

The lead role for commissioning of clinical health services belongs to various local and regional bodies within the NHS, for social care by local authorities, and for public health by the NHS (and now local authorities) ([Miller and Rees, 2014](#)). Private and voluntary sector providers have important involvement in particular aspects of delivery such as residential, specialist treatment, home support and advocacy services ([Mental Health Strategies, 2012](#)), and successive governments have been keen to enhance this diversity through attempts to introduce competitive procurement, payment systems and individually held budgets. Voluntary sector organisations are important partners for the wider commissioner process, through their community networks and understanding of specialist needs ([Department of Health, 2006](#)), ([Department of Health, 2012b](#)).

#### **4.2.1 High-secure**

There are only three NHS providers of high-secure beds (Ashworth, Broadmoor and Rampton Hospitals). These are designed for service users detained under the MHA 1983 who 'pose a grave and immediate danger to the public' ([Rutherford and Duggan, 2008](#)).

#### **4.2.2 Medium secure**

Medium secure services are specifically designed to meet the needs of adults detained under the MHA 1983 with the potential to abscond, who 'pose a

serious danger to the public' ([Rutherford and Duggan, 2008](#)). They require care and treatment in a secure setting to ensure they are safely managed ([Appleton, 2012](#)), ([Department of Health, 2007](#)). Access to medium secure services typically follows a court appearance, referral for rehabilitation, psychiatric intensive care unit (PICU) or from general mental health services, or transfer from high secure care.

The relatively open culture of medium secure services triggered the development of forensic mental health into a professionally and academically credible specialism, supporting multidisciplinary care. Emerging ideas of closest-to-home care the 'least restrictive alternative' as outlined in the *Reed report* ([Chiswick, 1992](#)), and escalating compulsory admissions both from prisons and the community ([Rutherford and Duggan, 2008](#)) fuelled demand for more local secure provision (low secure) ([Department of Health, 2002](#)).

#### **4.2.3 Low secure**

Low secure services for adults are provided for those who have long-standing and complex mental health problems, and cannot be safely or successfully cared for in acute inpatient mental health wards. These service users who pose a significant danger to themselves or others are usually detained under the MHA 1983, and present a level of risk greater than general mental health services could safely address ([Department of Health, 2002](#)), ([Appleton, 2012](#)), ([Rutherford and Duggan, 2008](#)). Service users will not require the level of physical security provided by medium secure services ([Department of Health, 2012a](#)).

The mix of individuals transferred from medium secure services, individuals diverted from custody and acutely disturbed non-offenders in local mental health services is a clinical challenge. This is now addressed by intensive care units, longer-term low secure units and newer 'enhanced' low secure units for acutely unwell prisoners. This difficulty between general and forensic services, is heightened by an intense pressure to move service users through a stepwise rehabilitation ([Sugarman and Dickens, 2015](#)).

#### **4.2.4 Psychiatric intensive care units (PICUs)**

Those charged with an offence and detained under MHA 1983 may be transferred to other mental health services such as PICUs and community forensic mental health services ([Rutherford and Duggan, 2008](#)). In some cases, service users may also be referred from prisons or rehabilitation wards ([Appleton, 2012](#)). Psychiatric intensive care is for compulsorily detained adults who are in an acutely disturbed phase of a serious mental disorder. A PICU is a secure ward that usually receives service users who cannot be managed on acute inpatient wards due to the level of risk posed to themselves or others.

The service user's length of stay is normally short (ranging from a few days to a few weeks, depending on their needs) and they are usually returned to acute wards as soon as their risk has reduced and more intensive treatment has started ([Department of Health, 2002](#)), ([Appleton, 2012](#)).

### **4.3 Memory Clinics**

A memory clinic or memory service is defined as a community-based multidisciplinary team (either provided for by the NHS or a private organisation) that assesses and diagnoses dementia, and may provide psychosocial interventions for dementia ([The Royal College of Psychiatrists, 2013](#)). This can include Community Mental Health Teams for Older People. There is currently no agreement about what such services should be called.

In 2011, The NHS Information Centre ([2011](#)) estimated 337 memory clinics in the UK. To achieve a more accurate picture of memory clinics, National Audits were conducted by the Royal College of Psychiatrists ([2013](#)), ([2015](#)) which estimated there to be approximately 214 memory clinics in the UK in 2013 and 222 in 2014.



#### 4.4 eLSU study

The Department of Health ([2007](#)) published *Best Practice Guidance: Specification for Adult Medium-Secure Services*, which led to the development of the Quality Network for Forensic Mental Health Services (QNFMHs). The QNFMHs developed standards for medium secure services, which were reviewed and updated in 2014 ([Royal College of Psychiatrists' Centre for Quality Improvement \(CCQI\), 2014](#)), where the groupings used complement the Department of Health ([2007](#)) best practice guidance. In 2012, the Department of Health completed a consultation as part of a review of the existing 2002 national minimum standards for general adult services in psychiatrics intensive care units and low secure environments ([Department of Health, 2002](#)). This review led to the publication of *Low Secure Services* ([Department of Health, 2012a](#)) to support the effective commissioning and delivery of services and the formulation of a QNFMHs for low secure units (LSUs).

To evaluate the impact of early and delayed participation and assess the true effectiveness of participation, the Evaluation of the Low Secure Units (eLSU) Study was an on-going randomised controlled evaluation running parallel to the generation of my thesis. The study was a two-armed, stepped-wedge design, parallel group, researcher-masked, clustered randomised controlled evaluation of early or late participation in a peer review network. Services randomised to late participation in the network had the opportunity to join the network once all follow-up data had been collected after one year. Compliance with a selection of key standards of care delivered by LSUs that participated in the evaluated network (intervention group) were compared against those that did not join in the first year (control group). This was repeated over three cycles (years). The primary outcome, the QELS checklist, was a measure of compliance with key environmental standards ([Aimola et al., 2016](#)).

## **Chapter 5 Aims**

In this thesis, I will examine how external peer review programmes aim to improve the quality of mental health services, and explore programme and non-programme factors that support their effective use. Following a systematic review of the extant literature, and initial discussions with stakeholders, I decided to include an examination of inner contexts relating to the organisation and culture of services prior to their participation in an external peer review programme, in particular their 'readiness for change'.

I will use the knowledge gained through this research to develop, test and refine CMOCs aimed at strengthening the operation of existing external peer review programmes and making suggestions for the development of future programmes.

To achieve this research question, I will address the following objectives:

1. To explore stakeholder views of the causal mechanisms of external peer review programmes
  - a. To use focus groups to examine practitioners' (CCQI staff) views of when, how and what changes take place in a range of external peer review programmes coordinated by CCQI
  - b. To use semi-structured interviews to examine subjects' (staff working in services) views of when, how and what changes take place in two peer review networks (QNFMHS MSU and QNFMHS LSU) and one accreditation scheme (AIMS PICU) coordinated by CCQI
2. To identify contexts which influence the ability of services to make use of their participation in external peer review programmes
  - a. To use focus groups and semi-structured interviews to examine which programme and non-programme factors influence the ability of member services of CCQI accreditation and peer review networks to make best use of participation
3. To assess whether 'readiness for change' influences the ability of services to make effective use of an external peer review programme
  - a. To use qualitative data to explore practitioners' (CCQI staff) and subjects' (staff working in services) views of how 'readiness for change' influences improvements in quality that are attributed to participation in external peer review
  - b. To use quantitative cross-sectional data to examine whether there is a relationship between certain 'readiness for change' constructs and indicators of quality improvement in services participating in accreditation
  - c. To use quantitative longitudinal data to examine whether 'readiness for change' scores can predict if a service is more likely to improve in quality following membership of a peer review network

## **Chapter 6 Research Methodology**

This chapter will provide an overview of the research methodology used in this study, with a specific focus on the critical realism approach I have used. This is followed by the types, analysis, strengths and limitations of qualitative and quantitative methodology. Finally, the approach to mixed methods research is presented.

### **6.1 Introduction**

Methodology can be defined as a 'thinking tool' which influences how a research question is framed ([Giddings and Grant, 2007](#)). I have located this study within the broad fields of social science research, management and clinical medicine. I characterised this through data collection, using a range of methods. During this chapter, I will develop the theoretical and philosophical positions to illustrate the methodological pathway I have taken. I have addressed this in terms of the research methods used, and the rigour associated with data analysis. I have demonstrated my awareness of my role as an external researcher, and the ethical issues that arose over the course of the research.

When considering which research methodology to adopt, three influences became important to me: how worldviews differ given the way we think the world is (ontology), how we gain knowledge of what we know (epistemology), how we think it can be investigated (methodology and research techniques) ([Creswell, 2009](#)) ([Fleetwood, 2005](#)). Although inherently linked, I considered each in turn.

## 6.2 Critical Realism

According to the work of Bhaskar, critical realism first originated in the late 20th century, as a response to positivist direct realism and postmodernist nominalism. It continues to occupy a middle ground between these two positions ([Reed, 2005](#)). Critical realism is an integration of realist ontology (the real world exists independently of our perceptions, theories and constructions) with constructivist epistemology (our understanding of this world is inevitably constructed from our own perspectives and standpoint) ([Saunders et al., 2016](#)).

A specific form of realism, critical realism aims to recognise the reality of the natural order and the events and discourses of the social world. It holds that we will only be able to understand, and thus change the social world if we identify the structures at work that generate those events and discourses.

I believe as researchers, we need to look for the bigger picture, as what we see is only a small part. Bhaskar ([1989](#)) argues that we can only understand what is going on in the social world once we understand the social structures that have given rise to the phenomena we are trying to understand. Critical realist research therefore focuses on explaining observable organisational events by looking for underlying causes and mechanisms through which deep social structures shape everyday organisational life ([Saunders et al., 2016](#)).

### 6.2.1 Ontological position

Ontology embraces the basic nature of social entities and reality; whether these are dependent or independent of individual consciousness ([Bryman, 2012](#)). According to Connelly's ([2007](#)) claim, critical realism and realist evaluation share a realist ontology. Critical realism proposes a 'stratified ontology', and assumes what we can observe is produced by underlying generative forces which may not be immediately observable ([Sayer, 2009](#)), ([Pawson and Tilley, 1997](#)).

Reality is the most important philosophical consideration for critical realists, and a structured and layered ontology are crucial ([Fleetwood, 2005](#)). Critical realists view reality as external and independent, but not directly accessible through our observation and knowledge of it ([Saunders et al., 2016](#)). Critical realists highlight how often our senses deceive us, and claim there are two steps to understanding the world. First, sensations and events are experienced; then there is the mental processing after the experience, when we 'reason backwards' (retroduction) from our experiences to the underlying reality that might have caused them ([Reed, 2005](#)) ([Bhaskar, 1978](#), [Bhaskar et al., 1998](#)), [McEvoy and Richards, 2006](#)).

### **6.2.2 Epistemological position**

Epistemology comprises the study of the scope, nature and utility of knowledge. As this study was located in the real world, which is complex, political and constantly changing; I reflected this epistemological stance in the research intentions and design.

The two main epistemological paradigms are positivist and anti-positivist approaches ([Bryman, 2012](#)). Realism shares two features with positivism: a belief that the natural and social sciences can and should apply the same kinds of approach to the collection of data and explanation, and a commitment to the view that there is an external reality to which scientists direct their attention ([Bryman, 2012](#)). Critical realists also embrace epistemological relativism ([Reed, 2005](#)), which recognises that knowledge is historically situated, and that social facts and social constructions are agreed by people rather than existing independently ([Bhaskar, 1989](#)). This implies that critical realist notions of causality cannot be reduced to statistical correlations and quantitative methods, and a range of methods is acceptable ([Reed, 2005](#)). In taking this position, critical realism accepts a form of epistemological relativism or constructivism.

### **6.2.3 Methodological position**

The stances taken have in turn influenced my methodological position and the overall research strategy I adopted ([Burrell and Morgan, 1979](#), [Morgan and Smircich, 1980](#), [Bryman, 2012](#)).

Robson ([2011](#)) distinguishes between ‘fixed’ and ‘flexible’ designs. The former is theory-driven; research is conducted to test and confirm the researcher’s theory and hypotheses, and mostly conducted under controlled conditions. Flexible designs are mostly exploratory in nature with less control over variables that produce the findings. I used a flexible design in this study and acknowledged the multifaceted nature of the research area in the methodological approach. My use of a flexible design allowed a mixed method approach to data collection and analysis.

### **6.2.4 Research design**

Bryman ([2012](#)) underlines that the function of the research design is to provide a framework for the collection and analysis of data. Inherent within this, is a reflection of the weight and significance given to various dimensions of the research process, including the degree of appreciation and acceptance of social phenomena and their interconnections, and the extent to which causal relationships between variables may be expressed ([Thurston et al., 2008](#)). Research design also encompasses the meaning and understanding of behaviour in social context and the extent generalisations can be made to larger groups. It therefore underpins the philosophical and theoretical basis of critical realism on which I conducted this research.

Prior to adopting any one research design, I considered the nature of the research question and objectives. My research is particularly relevant for addressing issues of who, how, when, where, and what ([Zikmund et al., 2012](#)). Descriptive research was relevant as it focused on exploring testing the MRT in stakeholders whilst scouting external peer review programme reality, through developing and testing CMOCs.

According to Øvretveit and Gustafson (2004), I have based this thesis on the advised steps for studying a quality improvement programme. I began by conceptualising external peer review programmes, and then reviewed previous research about similar programmes to develop programme theories in the form of an initial MRT. I identified research questions which arose from previous research and which could be of interest to stakeholders. I considered whether the programme could be controlled in its implementation, and whether comparisons could be made with similar control sites or sites that had differing levels of implementation. I planned methods to investigate how external peer review programmes were actually carried out in practice, the different programme factors and levels of implementation involved, and I gathered data about the sequence of activities and how external peer review could change over time. I noted any differences between the planned programmes and the programmes in action, and subjects' explanations for this as well as other explanations. I examined the component parts over time, the main outcomes, contexts and causal mechanisms which appeared to be critical in producing the outcomes. I specified the limitations of the study, the degree of certainty about the findings, and the answers to the research questions.



### **6.3 Qualitative Methodology**

Qualitative methods are useful in capturing and interpreting the specific constructions and construals that individuals make about how the world works. Critical realism places and interprets these within both a specific theoretical and a particular real-world context ([Patton, 2015](#)).

#### **6.3.1 Types**

In mental health services research, interviews and focus groups are typically the most common methods of qualitative research. They both involve the clarification of subjective meaning, experience, beliefs, and attitudes. This can be undertaken either through one-on-one interviews or small, facilitator-led, group discussion ([Whitley and Crawford, 2005](#)). Both focus groups and semi-structured interviews were used for the qualitative arm of this research study.

#### **6.3.2 Analysis**

Qualitative data analysis is a process of bringing order, structure, and interpretation to a mass of collected data ([Marshall and Rossman, 2011](#)). Unlike the quantitative research process, qualitative data collection and analysis processed are not completely separate and sequential. Qualitative analyses attempt to preserve the textual form of the data gathered and to generate analytical categories to explain themes. These are collected together, compared, and re-analysed to develop hypotheses or theoretical explanations. Pope and Mays ([2006](#)) mention three broad analytical approaches: thematic analysis, grounded theory and framework approach. They argue that these approaches range from a broadly inductive approach to a more deductive approach. I chose to use a framework approach for analysing data in this study ([Pope and Mays, 2006](#), [Ritchie and Lewis, 2003](#)). I chose this framework approach because it permits the consideration of previous theories and frameworks, which are vital in the analysis of data for a realist evaluation. The framework approach tends to be more explicit and informed by *a priori* reasoning in comparison with thematic analysis and grounded theory ([Lincoln and Guba, 1999](#)).

I used the logical form of creating new ideas, referred to as abduction. I used the data as a starting point and arrived at any new ideas through interpretation and de- and recontextualisation ([Estabrooks et al., 1994](#)). This could be a combination of old and partly familiar ideas, or the discovery of an idea that has never existed. The decisive point is that this idea, which in this form is new, explains or explains better something that was previously unexplained or unclear ([Reichertz, 2014](#)). Through deduction, I used the procedure of subsumption. This proceeds from an already known context of features, that are from a familiar rule and seeks to find this general context in the data to obtain knowledge about the individual case ([Reichertz, 2014](#)).

To make generalisations to develop CMOCs, I took the characteristics of a small selection of elements of a specific group to be representative of the characteristics of that group, and the existence of certain qualitative features in a sample implies the presence of other features. This was referred to as a qualitative induction, which was only a probable form of inference ([Reichertz, 2014](#)). Inductive inferences are tenuous, as they are not truth conveying but only more or less probable.

### **6.3.3 Strengths**

Focus groups and semi-structured interviews are flexible, easy to use and inexpensive. They provide an excellent method of collecting rich, diverse and insightful data. Their broad focus enables sufficient flexibility for new concepts and ideas to emerge ([Britten, 1995](#)). Non-verbal cues can also offer new insights. Responses to open-ended questions reflect an individual's personal reaction to the phenomenon under investigation, rather than one elicited by way of a forced choice between predefined options, and do not force consistency on people's thinking.

Focus groups can be attractive as a research method as they provide access to a wide variety of experiences. They provide a forum to explore different opinions, reflection on common practices and investigate participants' assumptions. Sim ([1998](#)) outlined that focus groups can:

- Provide information on the 'dynamics' of attitudes and opinions in the context of the interaction between participants
- Encourage more spontaneity in the expression of views than alternative methods
- Provide a 'safe' forum for the expression of views
- Participants may feel supported and empowered by a sense of group membership and cohesiveness

Semi-structured interviews have the potential to uncover and generate a detailed understanding of subjects' experiences relating to a particular research issue ([Rubin and Rubin, 2005](#)). According to Rapley ([2004](#)), semi-structured interviews provide the opportunity for social encounters where participants produce retrospective and prospective accounts of their past or future actions, experiences, feelings and thoughts.

#### **6.3.4 Limitations**

When conducting qualitative research, observer bias can occur in the researcher, individual interviews and focus-group facilitation. Observer bias is defined as one-sided viewpoint or specifically grounded standpoint on a phenomenon ([Olsen, 2012](#)). These are intrinsically present in many social situations. To overcome and challenge bias, I tried to develop a balanced, nuanced, carefully tolerant account. Both qualitative data collection and analysis can be prejudiced by the 'Hawthorne effect' ([Holden, 2001](#)), where the presence of the interviewer may influence participants' behaviour or responses. Issues, such as the researcher's role in observing, the adequate performance of recording equipment, the time to locate documents, are concerns that I addressed in this study, to mitigate against this.

## **6.4 Quantitative Methodology**

Quantitative data are collected on close-ended questions based on predetermined response scales, or categories. A quantitative questionnaire, for example, can ask respondents to rate their answers to questions to categories or on a scale using instruments that measure individual attitudes, and the scores are recorded in a close-ended fashion.

### **6.4.1 Types**

When testing the effectiveness of an intervention in clinical medicine the randomised controlled trial (RCT) is considered to be the 'gold standard' approach. Due to the dynamic nature of external peer review programmes, an RCT can only identify statistical associations between variables but cannot alone establish causality ([Bowling, 2014](#)). Three types of observational study are commonly used in healthcare research: cohort, case control and cross-sectional studies.

#### ***6.4.1.1 Cross-sectional data***

In cross-sectional designs, data are collected at a single point in time. This is often employed in conjunction with a survey method of data collection.

#### ***6.4.1.2 Longitudinal data***

A longitudinal study refers to an investigation where respondent outcomes and possibly treatments or exposures are collected at multiple follow-up times ([van Belle et al., 2004](#)). They are 'pre-post' studies in which a single baseline measurement is obtained, an intervention is administered, and a single follow-up measurement is collected. The change in outcome measurement can be associated with the change in the exposure condition. There are a variety of different models designed specifically for longitudinal data. These models fall into the following categories: random-effects, fixed-effects, dynamic and marginal models.

### **6.4.1.3 Clustered data**

Clustered data arise when data from the whole study can be classified into several different groups, referred to as clusters. Each cluster contains multiple observations, giving the data a 'nested' or 'hierarchical' structure, with individual observations nested within the cluster. The key feature of clustered data is that observations within a cluster are 'more alike' than observations from different clusters ([Galbraith et al., 2010](#)).

A cluster randomisation trial is one in which social units, or clusters of individuals, rather than individuals themselves, are randomised to different intervention groups. Trials randomising clusters, have become particularly widespread in the evaluation of non-therapeutic interventions in healthcare. The units of randomisation are diverse, including hospital wards and medical practices ([Donner and Klar, 2000](#)). Stepped wedge randomisation was used in the eLSU study ([Hemming et al., 2015](#)). This is when two groups have been randomised to receive the intervention at different time points and in sequence; groups that had not yet joined the peer review network remained as controls until one year had passed, this time point was predefined by the randomisation design as it was considered unethical to delay their participation any longer.

### **6.4.2 Likert scale and Likert-type items**

The original Likert scale used a series of questions with five response alternatives: strongly approve (1), approve (2), undecided (3), disapprove (4), and strongly disapprove (5 scales) ([Likert, 1932](#)). Likert combined the responses from the series of questions to create an attitudinal measurement scale. His data analysis was based on the composite score from the series of questions that represented the attitudinal scale; he did not analyse individual questions. While Likert used a five-point scale, other variations of his response alternatives are deemed to be appropriate ([Boone and Boone, 2012](#)). The Likert scale is based upon the assumption that each item on the scale has equal attitudinal value, importance or weight in terms of reflecting an attitude ([Kumar, 2014](#)). This is also the main limitation of the scale, as statements seldom have equal attitudinal value. The response categories in Likert-type

items and Likert scales have a rank order, but the intervals between values cannot be presumed to be equal ([Jamieson, 2004](#)).

Clason and Dormody ([1994](#)) described the difference between Likert-type items and Likert scales. They identified Likert-type items as single questions that use some aspect of the original Likert response alternatives. While multiple questions may be used in a research instrument, there is no attempt by the researcher to combine the responses from the items into a composite scale. Numbers assigned to Likert-type items express a 'greater than' relationship; however, how much greater is not implied ([Boone and Boone, 2012](#)). Due to these conditions, Likert-type items fall into the ordinal measurement scale ([Pett, 1997](#), [Blaikie, 2003](#)).

A Likert scale, is composed of a series of four or more Likert-type items that are combined into a single composite score during the data analysis process. Combined, the items are used to provide a quantitative measure of a character. Typically, researchers are only interested in the composite score that represents the character or personality trait. Likert scale items are created by calculating a composite score (sum or mean) from four or more Likert-type items. Therefore, the composite score for Likert scales should be analysed at the interval measurement scale.

Likert-type and Likert scale items can be considered categorical data. Answers are easy to score, responding to 'closed' questions that facilitate a quick response and make data easy to analyse. They do, however, restrict the respondent in their reply.

### **6.4.3 Analysis**

Quantitative research methods use statistical analysis to interpret and present data to inform the association between variables. Statistical approaches can be described as descriptive or inferential.

Descriptive statistics refers to the information on the spread of study data, such as the mean, median, standard deviation and inter-quartile range. I used the median or mode as the 'measure of central tendency' for ordinal data, as the arithmetical manipulations required to calculate the mean (and standard deviation) are inappropriate for ordinal data ([Altman, 1991](#)) ([Blaikie, 2003](#)) where the numbers generally represent verbal statements. Ordinal data may also be described using frequencies or percentages of response in each category ([Blaikie, 2003](#), [Boone and Boone, 2012](#)).

In general, parametric tests are preferable to non-parametric tests because a larger variety of tests are available and, if the sample size is not very small, they provide approximately 5% more power than non-parametric rank tests to show a statistically significant difference between groups. For non-parametric tests or distribution free tests no assumptions are made about the distribution of data.

Inferential statistics allow 'inferences' or deductions to be made from data. This is usually to test hypotheses or relate finding to the population beyond those who formed the study samples. This analysis seeks to establish whether you have got what you expected to find.

An interaction is when the effect of one variable on a second one depends on the value of a third variable. When such interactions exist, and they are extremely common in social research, they mean that any generalisations we seek to make about causal processes are limited. It is consistent with the critical realist view that mechanisms do not operate universally and that the research task is to specify the contexts in which they work.

#### **6.4.4 Strengths**

Quantitative research is appropriate in situations where there is pre-existing knowledge, which will permit the use of standardised data collection methods, and in which it is aimed to test hypotheses ([Bowling, 2014](#)).



Surveys can be carried out at one point in time (cross-sectional) or at more than one point in time (longitudinal). A major advantage of surveys is that they are carried out in natural settings, which allows statistical inferences to be made in relation to the broader population of interest and thus allows generalisations to be made ([Bowling, 2014](#)). This will increase the validity of the study. Longitudinal surveys are of value for studying the effects of new programmes, as greater precision can be obtained when measuring change than with a cross-sectional survey.

#### **6.4.5 Limitations**

Cross-sectional studies can only point to statistical associations between variables; they cannot alone establish causality ([Bowling, 2014](#)).

Respondents from longitudinal samples can also become conditioned to the study, and even learn the responses that are expected of them (as they become familiar with the questionnaire); they may remember, and repeat, their previous responses; they can become sensitised to the research topic and hence biased in some way; there can be a reactive effect of the research arrangements. This is known as the 'Hawthorne' effect as people change in some way simply as a result of being studied ([Roethlisberger and Dickson, 1939](#)) ([Bowling, 2014](#)).

A difficulty in the analysis of longitudinal data is known as 'response shift'. This refers to the scale of values which people use to make judgements, and the way in which it changes as changes in the variable of interest occur (beta changes). Occasionally, an individual's entire conceptualisation of the target variable might change (gamma change) ([Sprangers and Schwartz, 1999](#)). Response shift can be a problem if changes that are detected from self-reporting are not 'real changes' (alpha changes), but reflect beta or gamma changes ([Bowling, 2014](#)).

## 6.5 Mixed Methods Research

True mixed methods research involves a genuine ‘integration of the data at one or more stages in the process of research’ ([Creswell and Plano Clark, 2011](#)). It involves the planned mixing of qualitative and quantitative methods at a predetermined stage of the research process. Although these two approaches have previously been construed as incompatible, there is increasing acknowledgement that they can be complementary in healthcare. The core assumption is that the combination of approaches provides a more complete understanding of a research problem than either approach alone ([Creswell, 2014](#)). A balance of the flexibility of qualitative exploration with the fixed characteristics of theoretical grounding and hypothesis-testing inherent to many quantitative approaches is required in a mixed methods design ([Creswell and Plano Clark, 2011](#)).

Tashakkori and Teddlie ([2010](#)) suggested that most authors embrace pragmatism as the worldview for mixed methods research. Pragmatism draws on ideas such as ‘what works’, valuing both objective and subjective knowledge. However, the critical realist perspective has also been discussed ([Maxwell and Mittapalli, 2010](#)) as it validates and supports key aspects of both quantitative and qualitative approaches, and can facilitate communication and cooperation between them ([Greene, 2007](#), [Mark et al., 2000](#)) ([Bhaskar, 1978](#), [Bhaskar et al., 1998](#)).

### 6.5.1 Convergent parallel study design

Qualitative and quantitative data can be collected either sequentially or concurrently. In sequential studies one data collection method follows the other, whereas, in concurrent studies, data are collected at the same time. The four basic mixed methods designs are the convergent parallel design, the explanatory sequential design, the exploratory sequential design and the embedded design ([Creswell and Plano Clark, 2011](#)). The most well-known approach to mixing methods is the convergent parallel design; its purpose is ‘to

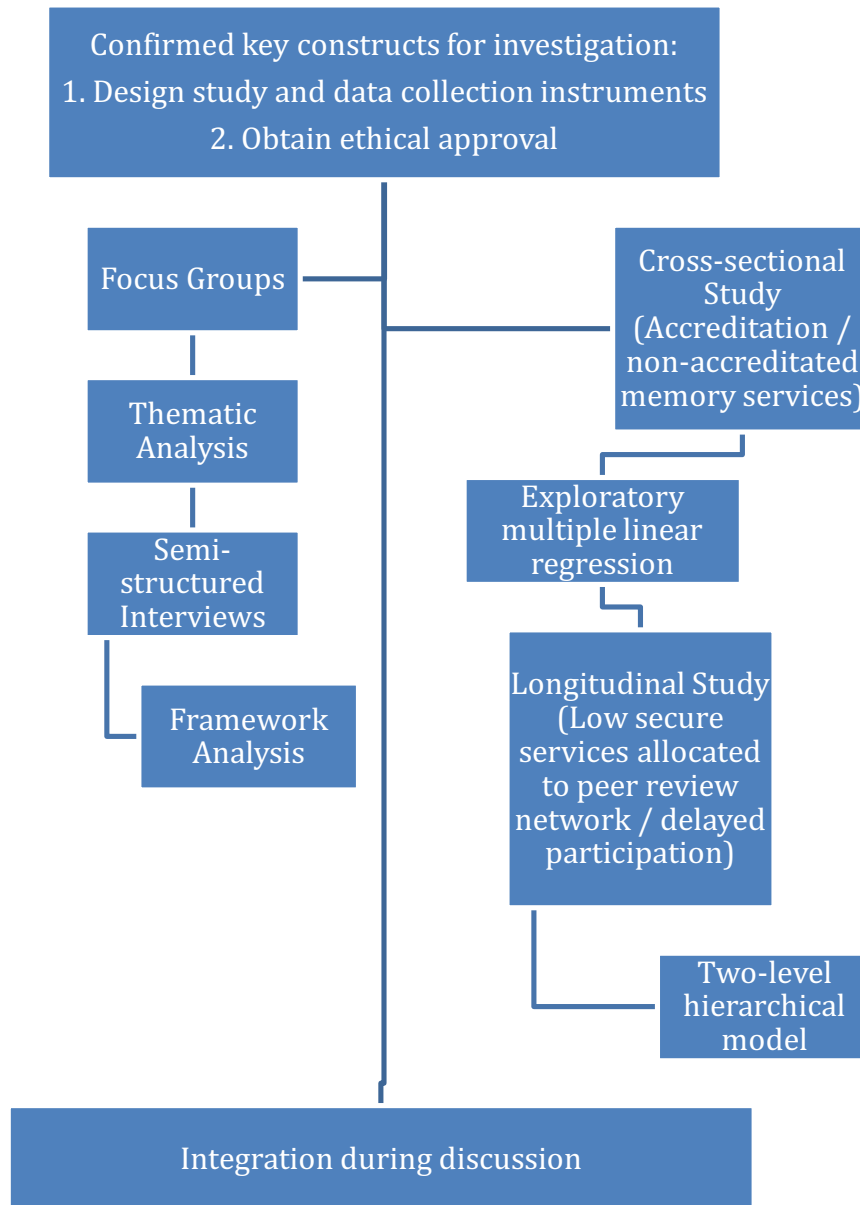
obtain different but complementary data on the same topic' ([Morse, 1991](#)) ([Creswell and Plano Clark, 2011](#)).

A strand encompasses the basic process of conducting research: posing a question, collecting data, analysing data, and interpreting results based on that data ([Teddlie and Tashakkori, 2009](#)). Mixed methods studies include at least one quantitative strand and one qualitative strand. The level of interaction is how much the two strands are kept independent. I collected two independent data sets to avoid bias in one form of data collection that follows another ([Creswell and Plano Clark, 2011](#)). I used an independent level of interaction as quantitative and qualitative strands were concurrent but kept separate from each other; one strand did not depend on the results of the other ([Creswell and Plano Clark, 2011](#)). The two parallel strands were conducted and analysed separately and only brought together during interpretation, this was the only point in the research process where I mixed the data. I utilised a qualitative priority and placed greater emphasis on the qualitative methods; quantitative methods were used in a secondary role ([Creswell and Plano Clark, 2011](#)).

In addition to matching the design to the purpose of the research, I made the following considerations when deciding to use the convergent design ([Creswell and Plano Clark, 2011](#)). Although I had a limited timeframe, I felt there was value in collecting and analysing both quantitative and qualitative data to understand the research problem. I possessed skills in both quantitative and qualitative methods of research and thus felt I could manage extensive data collection and analysis activities.

I based the validity of using the convergent parallel approach on establishing both quantitative validity (e.g. construct) and qualitative validity (e.g. triangulation) for each strand ([Creswell, 2014](#)) (see Figure 6.1).

**Figure 6.1 Convergent parallel study design (adapted from (Creswell, 2014))**



### 6.5.2 Integration

Triangulation can provide contrasting data sources to bolster confidence that the data are 'on the right lines' ([Denscombe, 2007](#)). I triangulated the findings from the four different studies by directly comparing and contrasting quantitative statistical results with qualitative findings for corroboration and validation purposes to develop CMOCs ([Creswell and Plano Clark, 2011](#)). The data were collected from different groups to increase the confidence in the data. I used a side-by-side comparison approach, which is presented in the discussion section of this thesis ([Creswell, 2014](#)). This merging step involved directly comparing the separate results to interpret to what extent and in what ways the two sets of results converged, diverged from each other, related to each other, and combined to create a better understanding in response to the study's overall purpose of the research ([Creswell and Plano Clark, 2011](#)). I also used mixing to illustrate the quantitative results with qualitative findings, synthesising complementary quantitative and qualitative results to develop a more complete understanding, and comparing multiple contexts within external peer review ([Creswell and Plano Clark, 2011](#)). Through triangulation, and the development of CMOCs, the plausible patterns that explained how the external peer review programme led to the observed results were confirmed. In a final step, these were translated into the more abstract level of the CMOCs, which was modified as necessary ([Marchal et al., 2012](#)). Typically, the comparison does not yield a clean convergent or divergent situation, and differences exist on a few concepts and themes ([Creswell, 2014](#)). Equifinality was a useful concept that I used, given that its typical causal model implies a few causal paths to an outcome ([Goertz and Mahoney, 2012](#)).

### 6.5.3 Ethics

Walliman (2005) describes two different aspects to consider when dealing with ethical issues. The first concerns the researcher's values of honesty, frankness and personal integrity or ethical behaviour. The second concerns the responsibility to the participants of research, including privacy, confidentiality and courtesy. I adhered to compliance with legal and professional requirements such as Ethics Committee Approval, Data Protection and professional codes of practice. The CCQI Research Ethics Committee approved CCQI research ethics for the staff interview aspect of the qualitative strand in June 2013 (Ref: 2013-1). I considered any potential harmful, adverse or risk-taking practices identified by staff during interviews, and arrangements were in place for identifying this and reporting if necessary.

As the nature of this data collection required time to be spent at sites not typically visited by the public (forensic mental health services), a gatekeeper was necessary. This was the role of my supervisor, who acted as an individual inside the organisation (CCQI) and was supportive of the proposed research and could essentially 'open up' the organisation. Practitioners and subjects were made aware of the nature of this relationship.

It was implied that completing the questions in the Memory Services Audit 2013 suggested consent had been given. Signed informed consent was gained from all other research participants before the respective questionnaire, focus group or interview took place (see Appendix F). Within the participant information sheet (see Appendix E), all relevant issues were addressed such as privacy, confidentiality, courtesy, feedback mechanisms and how the data protection requirements were met. Subjects were given ample time to read through the participant information sheet and then each subject was asked to sign the accompanying consent form. In addition, every interviewee was given a copy to retain.

All potential subjects for the qualitative study were written to by e-mail, explaining the aims, rationale and objectives of the study seeking to gain their

written consent. The procedures for how I proposed to collect the data and protect the gathered information were explained in detail in the participant information sheet (see Appendix E), as the research involved collecting data in places where individuals work ([Olsen, 2012](#)).

I entered sites in a respectful manner and tried not to disrupt the flow of activities. I tried to remain sensitive to exploring subjects' personal assessments and views of the phenomena within the study, which without the appropriate levels of both confidentiality and anonymity, could present the risk of embarrassment, conflict and loss of standing ([Bryman, 2012](#)). Participants were also informed that they could opt out of the study at any time should they wish to. I offered a reflective conversation to any participants who chose to opt out. Practitioners and subjects were clearly advised that transcribed data from the focus groups and semi-structured interviews could be reviewed. All research participants were clearly notified that the final audience for the thesis would be open access, thus identifiable interview material would be anonymised.

Downie and Calman ([1998](#)) point out, it is not acceptable to collect as much information as possible 'for the sake of it' but, on the contrary, researchers should decide which information is strictly necessary for analysis at a later date. I remained mindful that I would be party to sensitive information, particularly by being present on the sites of low secure, medium secure and PICUs. I kept all data collected during the study in a secure location, and took full responsibility to ensure complete privacy and confidentiality of all participants and of meeting the 'Data Protection Act' ([Department of Constitutional Affairs, 1998](#)).

## 6.6 Reflexivity: role as an external researcher

The key to validity in qualitative inquiry is being reflexive ([Finlay, 2002](#)). As a researcher, I have strived to remain aware of the ways in which my socio-cultural background and experiences may have influenced this research, and have sought to minimise such biases and remain as objective as possible.

Prior to joining the Centre for Mental Health and gaining an honorary research contract at the Royal College of Psychiatrists, my experience of the UK adult mental health sector was very limited. As a full-time PhD student, I held the position of an external researcher to the organisations I was collecting data from ([Saunders et al., 2016](#)). My lack of status in relation to the organisations (the CCQI and mental health services) meant I needed to remain sensitive to the issue of goodwill for practitioners and subjects giving up their time; thus, I tried to foster and recognise this at each stage in the study. My gatekeeper (supervisor) played an important role by creating awareness of my research in the organisation, adding credibility, and introducing my research to the relevant people ([Saunders et al., 2016](#)).

My ability to demonstrate research competence and integrity, and in particular, my ability to explain my research project clearly and concisely, was also critical at each level of access. To strengthen this, I undertook training in both facilitating focus groups, and qualitative data collection and analysis prior to collecting study data. This proved to be beneficial, as participants were willing to accept me as being objective, without a covert or organisation focussed agenda ([Saunders et al., 2016](#)).



## 6.7 Initial Middle Range Theory (MRT)

I formed a simple model, presented in Table 6.2, of the component parts of the external peer review programmes under study. These were the activities or stage (programme factors) that were integral to the programme. Without an understanding of a programme's underlying theory, success is left largely to chance, and measurement of the resultant outcomes is suspect at best ([Arnold, 2015](#)). The initial MRT was built up from programme documents using pre-existing publicly available descriptions, and findings from the systematic literature review. This enabled me to conceptualise an identifiable set of activities that were expected to lead to some identifiable outcomes ([Patton, 2012](#)), ([Øvretveit and Gustafson, 2004](#)).

This enabled me to develop the realist evaluation and generate some CMOCs to test in subsequent studies which might underlie the role of external peer review programmes in improving the quality of mental health services.

**Table 6.1 External peer review initial programme theory**

<b>Programme Factors</b>	<b>Non-programme (contextual) factors</b>
Joining	Outer: Healthcare system, coordinating body, economy, mandatory  Inner: Organisational culture, non-financial resources, local context, geographic location
Self-review	
Visit	
Verbal feedback	
Written feedback	
Peer reviewing other services	
Additional activities	

## Chapter 7 Qualitative Methods

I used qualitative research to generate 'theories in action'. I explored practitioners' (CCQI staff) views of what happens in practice through portraying their experiences of what happens in the field through facilitating focus groups. I undertook interviews to test the theories that had been developed with practitioners and further understand why subjects (staff working in services that are members of CCQI programmes) implement programmes in a way; and the strategies that they both believe are effective.

Formulation of programme theory from stakeholders' mental models involves drawing out the concepts of how they understood or anticipate the programme to work. Different groups or disciplines might operate from different mental models. I used qualitative methods to articulate mental models of how stakeholders understood external peer review programmes to work, how they would like to see them work, what their perceptions would look like if they were successful, as this was important for realist evaluation ([Funnell and Rodgers, 2011](#)).

## **7.1 Focus Groups**

I used focus groups to elicit practitioners' 'folk theories' on why and for whom the programmes work ([Pawson and Tilley, 1997](#)). I used focus groups over other qualitative approaches to observe the interactions between practitioners as people who work in the same organisation regularly share a common history and vocabulary. During communication, I wanted to elicit the special terms they used to describe their work, or phrases that were specific to the group, as practitioners may well have adapted the programme to try to get the best out of the services, and so will have specific ideas on what it is within the programme that works (mechanisms). They are also likely to have witnessed successes and failures (outcomes), and thus have some awareness of the people and places (contexts) for whom and in which the programme works. I used concept clarification to explore the meaning of these special, shared terms ([Rubin and Rubin, 2005](#)).

### **7.1.1 Planning**

I developed a topic guide from the key themes I had identified in the systematic literature review (see Appendix C). I contacted all the relevant staff who coordinated and worked on external peer review programmes via e-mail to start allocating them per their availability. Practitioners were asked to think about four pre-focus group questions prior to attending their respective session, to prime their thinking towards the topics of discussion (see Appendix H).

### **7.1.2 Data collection**

I ensured that the focus groups were undertaken in a venue that allowed adequate data recording. There were no feasible locations available, so I chose a room with the aim to maximise participation ([Barbour and Kitzinger, 1999](#)). I conducted the focus groups inside the CCQI building during the lunch hour to facilitate ease of participation, in hope of causing least disturbance to daily work activities. I also provided lunch refreshments as an additional incentive.

I conducted one pilot interview with a practitioner who was unable to attend any focus groups, and four subsequent focus groups. I aimed for enough heterogeneity within each focus group to stimulate discussion, but sufficient homogeneity to facilitate comparison between focus groups ([Barbour, 2001](#)). There is no defined number of participants necessary for a focus group to take place ([Sim, 1998](#)). I had three to five participants in each focus group. I chose staff who coordinated different external peer review programmes, with both short-term and long-term tenure, to enable a variety of thoughts and opinions.

I adopted an iterative approach, following up on themes that had been identified in one focus group, to test out in subsequent focus groups if other practitioners agreed, disagreed, or ranked the importance. Group processes and dynamics between participants provided some check as to what was acceptable and realistic, and extreme views could be challenged by the group. As discussed by Morgan ([1997](#)), questions and debates interested the participants and empowered them to contribute fully, thus stimulating areas of discussion, creating insights and revealing hidden meanings which would not have emerged in an individual exercise ([Barbour and Kitzinger, 1999](#)). Open debate and the presence of several perspectives in one room, ensured consistent and agreed views. Raising taboo and difficult areas was encouraged by the perceived safety of the group. The homogeneity of the group allowed capitalisation on the participants' collective shared experiences and revealed relevant cultural values or group norms.

I asked practitioners to search their memory for cases, illustrations and commonalities in respect of 'where change happened in the programmes' and 'what mechanisms and contexts had the most influence on programme success'. These inquiries released a flood of anecdotes, and the examples were remarkable not only for their insight but in terms of the explanatory form which was employed. These 'folk theories' were 'realist' theories and began to identify contexts and mechanisms which were conducive to the outcome of change in external peer review.

Calder ([1977](#)) proposed that when a facilitator reaches the point where they are able to anticipate fairly accurately what the next group is going to say, then there are probably enough groups, similar to theoretical saturation ([Bryman, 2012](#)). After the fourth focus group, I felt I had reached this point.

### **7.1.3 Data analysis**

I transcribed all the focus group data using Microsoft Word. I used unique identification (ID) codes during transcription for each participant. Firstly, I deductively coded the pilot interview using the thesis aims and themes I had generated from the systematic literature review. Then I proceeded with inductive coding ([Olsen, 2012](#)). I used thematic analysis to analyse and synthesise the data; using the software packages MAXQDA 11 and Microsoft Excel. I then used the findings to construct a topic guide to guide semi-structured interviews.

## 7.2 Semi-structured Interviews

I decided to use the data derived from practitioners (CCQI staff) to test out CMOCs in subjects (staff who work in services). Subjects are 'mechanism experts'; the research question will be tested through them, and will be about 'mechanism salience' ([Pawson and Tilley, 1997](#)). However, their fixed positions within a programme will mean that their sensitivity to the influence of context will be greatly limited since the circumstances in which they encounter the programme will be, for them, entirely routine. Subjects normally just experience one journey through a programme and, therefore, may also have little understanding of different outcomes ([Pawson and Tilley, 1997](#)).

I chose to use one-to-one semi-structured interview format, face-to-face with subjects, based on the format's ability to reconcile both structure and flexibility ([Bryman, 2012](#), [Zikmund et al., 2012](#)). This approach allowed me to explore issues arising during the course of the interview, while still using the interview guide to both structure and drive the dialogue ([Silverman, 2011](#)).

I chose this approach over other methods such as structured or telephone interviews as they offered a way of framing a clear area for discussion, but still had flexibility through their use of open questions to obtain a diversity and richness of data ([Oppenheim, 1992](#)). They also allowed direct interaction with the participant and thus an opportunity to consider non-verbal cues ([Britten, 1995](#)).

### 7.2.1 Topic guide development

I developed the content of the topic guide based on themes from the systematic literature reviews, focus group findings, first-hand experience of attending a combined low secure and medium secure service QNFMHS visit, and discussions with CCQI staff ([Rapley, 2004](#)) ([Britten, 1995](#)).

To increase the face validity of the topic guide; the material was reviewed by two other senior practitioners with expertise in the field of qualitative research

in mental health services. Saunders et al. ([2016](#)) suggest that validity for the semi-structured interview has the potential to be high where the interview itself is carefully conducted, which was my aim at all times.

I used a loose structure within a defined scope, and predetermined, open-ended questions were employed. According to advice from Kumar ([2014](#)) the questions were clear, open-ended, non-leading, sensitive and avoided unnecessary jargon. In terms of sequence, the topic guide commenced with background, general and easy to answer questions, before moving on to more sensitive and challenging issues.

### **7.2.2 Purposive sampling strategy**

I applied a purposive sampling technique with the aim of explicitly selecting interviewees who might generate appropriate data ([Mack et al., 2005](#)). I selected a purposive sample of interview participants to obtain a range of individuals with different professional backgrounds and levels of experience. Previous literature showed that senior, ward managers and frontline staff were important subjects which is why I included these specific groups in the purposive sampling strategy. Senior management were at the top of the organisational hierarchy and held organisation-wide responsibilities and practices ([Merali, 2005](#)). Middle managers were in the middle of the organisational hierarchy and had one or more managers reporting to them. For the purpose of this study, I grouped middle managers with senior management, as senior managers were not always available to participate ([Mintzberg, 1973](#)). Ward managers were defined as managers at the first level of the hierarchy who had frontline staff reporting to them. Stickley ([2006](#)) suggested a critical realist framework offered an argument for an alternative to accepted methods of service user involvement. He argued that because mental health frontline staff are often the workers who have the most contact with service users, their philosophy and approaches can represent people who use mental health services in some instances, which is why I included frontline staff in the purposive sampling strategy.



The realist position that meaningful categories or underlying structures exist, also has an implication for evaluation. It justifies the search for programme subtypes and other classifications such as different groups, different external peer review programmes, different geographical locations, and different healthcare structures ([Mark et al., 2000](#)); which is the reason behind including these contexts (based on the systematic literature review) in the purposive sampling strategy. The categories represented a particular salient attribute of the sample, but they only served as an indicator of different contexts which influenced external peer review ([Mason, 2002](#)).

Glaser and Strauss ([1970](#)) wrote that when field researchers studied more than one setting at a time, they framed the issue well. They argued that using multiple comparison groups helped to find out ‘under what sets of structural conditions hypotheses are minimised and maximised’. This helped me to calculate where changes were most likely to occur. Using multiple settings also helped me to find negative cases to strengthen the CMOCs, which I built through examination of similarities and differences across cases. It helped to not only elucidate the specific contexts under which an outcome would occur, but also helped to form the more general categories of how contexts could be related ([Miles and Huberman, 1994](#)).

I used a multistage strategy to select interview participants. This involved choosing settings and groups representing a sample in two stages. The first stage was based on availability; and the second stage was more purposeful to achieve a good distribution from public and private services, and small and large mental health services. Previous studies found it difficult to demonstrate change in the time period they were assessing ([Roberts et al., 2012](#), [Salmon et al., 2003](#)). For this reason, I also incorporated both new and long-standing members of external peer review programmes in the purposive sampling strategy. Two studies from my literature review provided evidence of disparities in achieving accreditation between rural and urban hospitals ([Brasure et al., 2000](#)). For this reason, I attempted to include an even

distribution of services located in both rural and geographical parts of the UK as well as both the North and South of the country.

Previous literature has focused on the benefits and pitfalls of making external peer review a compulsory process. As the low secure peer review network was not compulsory, but the medium secure was a mandatory programme, this provided me with an opportunity to sample from both types of programme. Overall, I sampled member services belonging to three different programmes as illustrated in Table 7.1.

**Table 7.1 Different external peer review programmes**

<b>CCQI Programme</b>	<b>New / long-standing</b>	<b>Type of external peer review programme</b>	<b>Participation</b>
Low secure QNFMHS	new-standing	peer review network	Voluntary
Medium secure QNFMHS	new and long-standing	peer review network	Compulsory
AIMS PICU	from new and long-standing	accreditation programme	Voluntary

Mason ([2002](#)) points out that the focus of sampling is towards achieving a depth of investigation rather than a breadth of coverage. Decisions about who to sample were ongoing throughout the research, facilitating a change of direction if necessary, leaving open the possibility of doing multiple interviews with some subjects if necessary. I undertook this in the case of one accreditation service for which data I collected data both before and after their accreditation visit. I also used extreme case sampling of individuals who I thought could shed light on the phenomenon being studied. For example, I chose to sample a service on the basis that it had failed accreditation the first time and was about to undergo accreditation again.

### **7.2.3 Recruitment**

Announcements were made by programme managers at both low secure and medium secure QNFMHS Annual Forums, followed by an e-mail to the nominated contact persons informing member services that they had been automatically opted in to be contacted by myself to participate. A similar e-mail was also sent to nominated contact persons for PICUs who were members of AIMS PICU.

After a few weeks, I was provided with databases containing contact details for services who did not wish to opt-out at this stage. I contacted the nominated contact persons using the purposive sampling strategy by e-mail between February 2014 and April 2014 to reconfirm their willingness to provide three interview participants per site (six for combined low secure and medium secure sites), and to arrange a time and date to set up the semi-structured interviews. The e-mail detailed the study and included a participant information sheet (see Appendices D and E). I followed up on correspondence with a phone call in several instances. These were, in turn, reiterated in written confirmation to the nominated contact persons, which enabled me to reassure them of the confidential nature and maximum expected length of interviews. I conducted interviews at different points in services' membership cycles.

### **7.2.4 Pilot study**

I decided to conduct three pilot interviews for each type of external peer review programme. This provided an opportunity to test out the topic guide which would be used for the remainder of interviews.

### **7.2.5 Conducting interviews**

I always arrived at least fifteen minutes before the first interview session to allow time for signing in, to show necessary identification, go through security procedures, be shown to a room to set up, ensure the necessary documents were laid out on a table prior to the subjects' arrival, and to check the dictaphone was working.

I welcomed the subject on arrival, and thanked them for their time which served to establish an initial rapport. I presented each subject with a participation information sheet, a research study sheet, a consent form, and a short demographic checklist (see Appendices E, F and G). I requested that they sign a consent form and complete a demographic checklist (collecting information about their job title, which external peer review programme their service was a member of and their tenure). Before recording began, I recapped purpose of the interview and then proceeded to outline the purpose of the study by briefly summarising the contents of the participant information sheet.

In realist interviews, the exchange of ideas is driven by the researcher's theory ([Pawson and Tilley, 1997](#)). I used a teacher-learner function by playing a more active role in 'teaching' the overall conceptual structure of the investigation to the subject. During interviews, I taught subjects the initial MRT ([Pawson and Tilley, 1997](#)). Subjects, having learned the theory being tested, could teach me about programme factors, and related mechanisms and contexts in a particularly informed way.

I sought to actively listen and pay attention for the duration of the interview which, in turn, facilitated the identification of opportunities for further probing of issues where appropriate, as offered by the semi-structured interview format. Moreover, I aimed to build an empathetic, non-judgemental interview atmosphere, which would allow the interviewee's feelings to emerge. I tried to shape an atmosphere that encouraged exploration of feelings. I did this by responding in ways to encourage exploration of feelings, involving reflecting not just on the content but also the feelings that emerged ([Wilkinson et al., 2004](#)). I made every effort not to ask leading or suggestive questions, asking open questions only. I started by asking 'do you remember how you first became aware of this peer review network or accreditation programme?' If necessary, I prompted ideas based on responses and themes from previous interviews. I encouraged subjects to explore the issues they mentioned, which meant in some cases I could obtain a great deal of information about each area I had intended to explore. I tried to respond in a manner that picked up the subject's own

wording, seeking to establish a naturalistic pathway through their worldview ([Wilkinson et al., 2004](#)). Reflecting, mirroring, paraphrasing and summarising techniques were also used to maintain the flow of conversation without imposing my own understanding.

I asked questions to try to evoke vivid descriptions. Vivid anecdotes or examples allowed me to picture the examples, and to respond not just intellectually but also emotionally. To obtain vivid reports I asked for narratives or requested step-by-step descriptions of what had happened ([Rubin and Rubin, 2005](#)). I evoked richness by encouraging interviewees to elaborate. When a subject gave detailed answers, I encouraged them through continuation probes. I suggested that I wanted to hear more details to elicit more elaborated and refined material. Even when a narrative seemed complete, I tried to ask for more examples, because each example would be a little different, adding themes or concepts or putting a new slant on them.

Preferred theories are how subjects would like the programme to operate or think it should operate, how they would like to benefit or see others benefit from the programme, how they would like the programme to operate, and some of the barriers to effective operation. I asked subjects to name which features of the programme they had found most helpful, or at the end of the session were asked for any recommendations or suggestions for improvements that might benefit them ([McEvoy and Richards, 2006](#)).

Interviews lasted from fifteen minutes to one hour and twenty minutes. At the end of each interview, I thanked subjects for their participation and time. Very few subjects became shy or hesitant at the thought of being recorded; most appreciated being recorded as it ensured their messages would be transcribed accurately. I recorded the interviews on a dictaphone and then mp3 audio files were stored on a secure password-protected USB drive. Recordings were replayed and analysed throughout the data collection period to refine CMOCs and determine what questions to ask subsequent subjects ([Rubin and Rubin, 2005](#)).

### **7.2.6 CMOC testing**

I continually revised CMOCs throughout, rather than conducting a discrete phase of data analysis after all data had been collected ([Fowler, 2013](#)). I undertook CMOC refinement by learning subjects' theories, formalising them, teaching them back to new subjects, who were then in a key position to comment upon, clarify and further refine the key ideas. This process was repeated throughout all the interviews ([Pawson and Tilley, 1997](#)).

I asked subjects how they implement external peer review programmes and why they undertook some activities that appeared to be at variance with the original programme structure or omitted parts of the programme structure. I recognised that programmes were implemented differently, in different sites and by different organisations. Therefore, there was a possibility of different programmes all operating under the same external peer review programme banner. Many of these differences reflected appropriate adaptations to local contexts which is often required by programmes with complex aspects operating in complex situations. I tried to understand the theories that were operating in these different contexts to enrich CMOCs by identifying what works for whom, under what circumstances, and why various adaptations were useful.

### **7.2.7 Theoretical saturation**

To determine the sample size for a purposive sample, the 'gold standard' ([Guest et al., 2006](#)) is saturation. To achieve saturation, a researcher collects and analyses cases to the point that sampling additional cases does not provide any new information (informational redundancy) that can be incorporated into the thematic categories (theoretical saturation) ([Lincoln and Guba, 1994](#), [Sandelowski, 1995](#)). However, determining an appropriate sample size to achieve the goal of saturation can be complex and is often mediated by data quality in terms of amount and degree of complexity, sample heterogeneity and resources ([Guest et al., 2006](#)). Recruitment continued until I had reached theoretical saturation ([Pope and Mays, 2006](#), [Ezzy, 2002](#)). This means I carried on sampling theoretically until a category had been saturated with data, until no

new or relevant data seemed to be emerging regarding a category, categories were well developed and the relationships among categories were well established and validated ([Bryman, 2012](#)).

### **7.2.8 Data analysis**

I undertook half of transcription myself, where possible, immediately after interviews had been completed. In instances of ambiguity on recordings, I could remember and fill in the missing material (indicating on the transcript it was from memory). Careful and regular review of the transcripts provided an opportunity to follow up on the material generated from interviews ([Rubin and Rubin, 2005](#)). I produced verbatim transcripts of audio files to incorporate any 'ums' and 'ers' from responses to preserve the full content of interviews ([Sandelowski, 1994](#)), ([Pope and Mays, 2006](#), [Silverman, 2011](#)). Being mindful of the sensitivity and confidentiality of the content of interviews, the remaining half of transcription was undertaken by a private authorised professional under my guidance. I undertook subsequent proofing of these transcripts against original recordings, and repeated this again during the familiarisation of transcripts ([Sandelowski, 1994](#)) ([Bryman, 2012](#)). It was important that I became familiar with the recordings and transcripts, as this was a vital stage of interpretation ([Gale et al., 2013](#)).

MAXQDA 11 is a computer-assisted qualitative data analysis software (CAQDAS) package, which helped me to manage the large amounts of data ([Pope and Mays, 2006](#)). I used it to manage and code transcripts, utilised the data search functions, and record reflective notes in the form of memos. After familiarisation, I read each transcript line by line, and applied descriptive codes developed from the research question using a deductive method ([Crabtree and Miller, 1999](#)). This was then followed with analytic codes in an inductive manner (open coding) ([Boyatzis, 1998](#)) to ensure important aspects of the data were not missed ([Gale et al., 2013](#)). The readiness for change items were inductively coded, the two categories, barriers to change and reported indicators were deductively coded as the patterns emerged from analysis of the transcripts. I based coding in terms of programme factors, underlying

mechanisms, contexts and reported changes ([Marchal et al., 2012](#)) ([Doi et al., 2015](#)). I tried to elicit the meanings perceived by the subjects to gain an understanding of their ways of sense making, thus providing transferable learning. It was important to refer back to the whole transcript and notes made during familiarisation to maintain the fit of the data to the themes ([Furber, 2010](#)) and to ensure that the context of the data was not lost.

I added new codes when new ideas emerged from the data until I produced a final code manual. I grouped similar codes together to form overarching themes, and grouped codes together into clearly defined categories to form a working analytical framework ([Gale et al., 2013](#)). I then applied the working analytical framework by indexing subsequent transcripts using the existing categories and codes.

I used a Microsoft Excel spreadsheet ([Swallow et al., 2002](#)) to generate the matrix. I then summarised the data by category from each transcript and charted it into a framework matrix to identify patterns of meaning, similarities and differences, especially between the subgroups ([Gale et al., 2013](#)). I included as much data as was necessary in charting rather than discarding it if it did not appear to meet the objectives of the study ([Furber, 2010](#)).

I had noted early interpretations and continued to record impressions using analytic memos. Gradually this helped to identify characteristics of and differences between the data, interrogating CMOCs and I mapped connections between categories to explore relationships and causality ([Gale et al., 2013](#)). Comparing and contrasting data was vital, and using the framework approach enabled me to compare data across cases with ease as well as within individual cases ([Gale et al., 2013](#)). Making comparisons across different cases helped me to readily identify exceptions and suggest reasons for these.

I then advanced to the interpretative phase, where the findings of the systematic reviews and focus groups complemented the generation of CMOCs about what mechanisms are operating, in what context, and to produce what



outcomes in external peer review. Framework analysis helped me to remain immersed in the data, as analysis took place at the data-level. This helped to identify connections and patterns, to make systematic comparisons, and develop interpretations.

I used cross-case comparison by inspecting each case in a setting (context) to see if it fell into clusters or groups that shared certain patterns or configurations ([Miles and Huberman, 1994](#)).

### 7.3 Reliability, Validity, Rigour and Authenticity

Reliability and validity are key considerations for any research project. Reliability is concerned with the replicability of research findings using similar methods ([Ritchie and Lewis, 2003](#)). However, in qualitative research, because of the different arguments on reality and effects of context those concepts have greater resonance with confirmability ([Ritchie and Lewis, 2003](#)), consistency ([Denzin and Lincoln, 2008](#)) or dependability ([Marshall and Rossman, 2011](#)). Ritchie and Lewis ([2003](#)) explain that, to ensure these qualities exist, one should conduct internal checks on the quality of data and their interpretation, and provide information about the whole research process. The process of analysis and coding can also be described clearly and applied methodically and systematically ([Green and Thorogood, 2009](#)). Validity is traditionally understood to refer to the 'correctness' or 'precision' of research ([Ritchie and Lewis, 2003](#)). There have been attempts in the qualitative literature to move away from validity and use other terms which are more appropriately related to the correctness of qualitative evidence. It is suggested that 'credibility' and 'transferability' translate more appropriately for qualitative enquiry ([Lincoln and Guba, 1999](#)). Walliman ([2005](#)) offers alternatives, and I ensured these were considered. Data were collected and analysed in a transparent manner. To avoid memory issues, I recorded all qualitative data, used transcripts, and made notes during and after each interview. To ensure the accuracy of recorded data, I compared all transcribed qualitative data against original recordings. I considered my mind-set and preconceptions always.

To increase rigour and credibility of my findings, a senior researcher checked over one of my transcripts using the coding frame. In health science research, this method is used when inquirers want an external check on the highly interpretative coding process ([Creswell, 2013](#)). There was synchrony in the codes used and how things were coded, and he agreed on the appropriateness of the coding frame I had developed. Silverman ([2005](#)) also supports inter-coder agreement.

I checked for misinformation that could have stemmed from distortions introduced by the myself or informants by clarifying points with programme managers at the CCQI ([\(Ely et al., 1991\)](#) [\(Erlandson et al., 1993\)](#) [\(Gelsne and Peshkin, 1992\)](#) [\(Lincoln and Guba, 1985\)](#)).

Reflexive triangulation includes the audience's reactions to triangulation: my reflexive perspective, the perspective of those studied, and the perspectives of those who received the findings ([Patton, 2015](#)). I presented draft findings to multiple audiences, including practitioners at the CCQI and stakeholders at the QNFMHS low secure annual forum to learn how they reacted, what they focused on, what was clear and unclear and what questions were inadequately answered. This form of audience undertaken supplemented data triangulation by increasing confidence in the validity of the findings ([Silverman, 2011](#)). I considered which aspects of data were the most critical and triangulated them first ([Sake, 2010](#)). I solicited practitioners' views of the credibility of the findings and interpretations by presenting integrated findings, and then using this as a way of 'open forum' audience review. This helped to strengthen CMOCs and verify the accuracy of the account, and reduced bias and reactivity (my presence as researcher interfering in the setting or influencing the behaviour of the research participants). This validation method can include a range of techniques in which the investigator's account is compared with the accounts of those who have been investigated to establish the level of correspondence between the two sets ([Pope and Mays, 2006](#)). This allows a check on factual accuracy and allows the researcher's understandings to be confirmed (or amended) by those whose opinions, views or experiences are being studied. I was seeking accuracy, my possible insensitivity, and new meanings during this form of validation, and did not want to present findings back to subjects since Pope and Mays ([2006](#)) have maintained that the account produced by the researcher is designed for a wider audience and might be different from the account of an individual informant because of their different and limited role in the research process. As such, the analysis of the data might take the explanation beyond something that would be immediately recognisable to the subject ([Denscombe, 2007](#)). Programme evaluation constitutes a challenge in

establishing credibility, as the ultimate test of credibility is the response of primary intended users and readers of the evaluation. Their reactions often revolve around face validity ([Patton, 2015](#)). In seriously soliciting intended users' reactions, I ensured my perspective was joined to the perspective of the people who will use these findings.

I used negative case analysis to refine CMOCs as the inquiry advanced ([Ely et al., 1991](#), [Lincoln and Guba, 1985](#), [Miles and Huberman, 1994](#), [Patton, 1980](#), [Patton, 2015](#)), in light of negative or disconfirming evidence. Not all evidence fitted neatly into themes, but it was necessary to report these themes to provide a realistic assessment as where patterns and trends had been identified. My understanding was also increased by considering instances and cases that did not fit within CMOCs. These included exceptions that illuminated the boundaries of CMOCs, and in some instances broadened my understanding, changed my conceptualisation or cast doubt on the CMOC altogether ([Patton, 2015](#)). I reported the bias for conclusions reached about the salience of the negative or deviant cases.

## Chapter 8 Quantitative Methods

I sought to collect quantitative data from two studies, as they provided me with the opportunity to sample services from both peer review networks and accreditation programmes. At first, I implemented five constructs of readiness for change in a cross-sectional study. But due to the inherent limitations of cross-sectional data, I implemented the full tool (all 25 constructs) in a longitudinal study when the opportunity arose as it is a more robust quantitative method.

### 8.1 Cross-sectional Memory Clinics Audit

Between July and September 2013, the CCQI Memory Services National Accreditation Programme (MSNAP) conducted the National Audit of Memory Clinics in the UK. This was designed as a cross-sectional study as it attempted to collect data from managers of UK memory clinics within a short time frame. This highlights the advantages of this type of study, as a large volume of data can be collected in a short time scale in comparison to cohort studies.

The questionnaires included the same domains that were in the 2011 NHS Information Centre audit ([The Royal College of Psychiatrists, 2013](#)). In addition, I decided to embed five separate Likert-type items of readiness for change from the 25-question tool ([Bobiak et al., 2009](#)) in the audit (see Appendix I). Only five items were selected as this was the number deemed to be feasible to complete by managers of memory clinics. The tool was amended for the target population of memory clinics, and I had sought the appropriate permission from authors. These questions were selected as they fitted with key themes that I had identified in preliminary literature as important in influencing readiness for change within external peer review, as I had limited first-hand practical experience of accreditation at the time.

I selected the construct 'things have been changing so fast in this service that it is hard to keep up with what is going on' to ascertain how multiple, fast-paced changes can influence quality improvement. Literature suggested that change is

often dropped into an organization without modifying surrounding organisational systems that influence success ([Gustafson et al., 2003](#)). Various studies have also identified organisational stability as a key feature of organisational readiness ([Burnett et al., 2010](#)).

I chose the construct 'this service tends to be very flexible in dealing with change', as various studies had identified the culture and environment for improvement work as a key feature of organisational readiness ([Burnett et al., 2010](#)).

I selected the construct 'leadership in this service articulates a vision for the service' as Gustafson et al. ([2003](#)) summarised that change is more successful when defined at a high level within the organisation. Performance improvements were reported to be greater in services with leaders who effectively demonstrate support for the new strategy than in services with leaders who do not ([Caldwell et al., 2008](#)).

I chose the construct 'this service is generally willing to try new things to improve patient care' as previous studies had identified past history of successful change as a key feature of organisational readiness ([Burnett et al., 2010](#)). Groups that shared a norm of valuing innovation were more likely to change than those who did not ([Caldwell et al., 2008](#)).

I selected the construct 'this service is typically able to adapt new standards or procedures, even those forced upon us', as supporters and opinion leaders were found to be more likely to be innovative and support change if the organisation's norm is to adopt changes ([Rogers, 1995](#)).

### **8.1.1 Data collection**

The identified contacts from memory clinics were asked to complete the questionnaires via a link to a webpage that had been e-mailed to them. This contained the necessary information and access to the online questionnaire. Direct automatic entry was used as the data were generated from online

surveys and was later made available to me in Microsoft Excel. Before I received the dataset, extreme outliers and null responses were identified and the responder was contacted to ask for clarification, if needed in certain cases. In instances where no answer was received, the data were removed from calculations.

### **8.1.2 Data analysis**

Prior to data analysis I selected a primary and secondary outcome from the standards that were assessed in the audit. I aimed to identify key measures of service quality that were recognised as important by clinicians and were reliably assessed in the audit. I discussed which items I should use with staff working on the audit and selected provision of recommended services and treatments (i.e. home-based assessments, access to specialist post-diagnostic counselling, indication of anti-dementia medication, review of anti-dementia medication, access to cognitive stimulation therapy, access to education and support for carers and access to Life Story work) as the primary outcome. Improving access to recommended treatments and services has been a major focus for the MSNAP programme ([Mainz, 2003](#)).

Speed of access to services, measured through waiting times, are a commonly used measurement in assessing the quality of healthcare ([NHS England, 2015](#)). Waiting times were assessed in the audit and had also been the focus of work of the accreditation service but feedback from clinicians and staff working in MSNAP indicated this outcome was less under the control of Memory Clinics. I therefore selected waiting times, as measured by the average number of weeks' wait between receipt of referral and the person starting their assessment, as the secondary outcome for my analysis.

Data analysis was carried out using version 22 of SPSS Statistics (Statistical Packages for the Social Sciences), Microsoft Excel and SAS 9.4. Alpha was set to 0.05, as this is the most common alpha used and well-established by statisticians ([Fisher and Bennett, 1990](#)).

Transformation of a variable may allow parametric statistics to be used if the transformed variable follows a normal distribution. A logarithmic transformation of the secondary variable of waiting times (measured in days) was necessary to proceed with further statistical analyses. When a distribution has a marked tail to the right-hand side, a logarithmic transformation of the scores is often effective ([Chinn, 1991](#)). This is advantageous as they provide interpretative results after being transformed into original units. I transformed the items that had reverse scoring to assess their internal consistency. I produced descriptive statistics for the whole sample and by grouping variables.

I conducted an exploratory multiple linear regression (ordinary least squares estimation) to investigate potential predictors of the outcome variables, waiting times, and maximum new service users. Whilst a multiple regression does not require the data to be normally distributed, the residuals of the dependent variable should be approximately normally distributed ([Slinker and Glantz, 2008](#)), which was checked with a histogram plot. The Stepwise method was chosen due to the lack of previous relevant research and theory on this topic which made a hierarchical method inappropriate. The Stepwise method is also based on a statistical calculation which is advantageous over the forced entry method.

To look for a correlation between five Likert-type items against primary (provision of services and treatments) and secondary outcomes (waiting times) it was necessary for me to compute correlation coefficients. A correlation coefficient describes how closely two variables are related; the amount of variability in one measurement that is explained by another measurement.

I obtained descriptive and summary statistics first to provide a good working knowledge of the data before embarking on multivariate modelling, and so I could build the model up in a systematic way. Once all variables in the model had been included, I used a logical process to decide which variables to remove, through using a backwards elimination process.



The model was computed to aid comparison between accredited and non-accredited memory clinics, to observe if the accreditation programme has had any effect on both the primary and secondary outcomes and Likert-type items. Thus, I built interaction effects of being or not being a member of an accreditation programme (MSNAP) into the model.

To build a bigger picture, I considered what other variables could be associated with high quality memory clinics. I then used the number of people were seen by services in the last year as a proxy outcome for size, and I included this in the model as a possible confounding factor. I also included participation in research as an additional binary outcome, after consulting with the MSNAP programme manager for consensus.

Following logarithmic transformations, I tested the secondary variable and the proxy indicator for size, the number of new service users that were seen in the last year for normality.

## **8.2 Readiness for Change Longitudinal Study**

As aforementioned, the eLSU study was a randomised controlled evaluation which parallel to my PhD. I collected longitudinal data for this study, from a readiness for change tool which I chose to embed into the randomised controlled trial. In addition, I also used the Quality of Environment of Low Secure Services (QELS) checklist scores which were collected as the primary outcome measure of quality improvement for the purpose of the eLSU study ([Aimola et al., 2016](#)). Both these measures were collected at baseline and follow-up (after 12 months) as part of the eLSU study. The checklists were completed by myself or another researcher collecting data for the wider eLSU study, using data provided by the ward or service LSU manager on site. I used QELS checklist scores as the primary outcome of interest for the longitudinal study, as they were the primary outcome of quality improvement for the eLSU study.

### **8.2.1 Self-administered questionnaires**

I chose to select self-administered questionnaires as they have the least researcher, and are less susceptible to information bias and interviewer effects. Questionnaires were administered with as little variation as possible so that undue bias was not introduced into the process. Standardised procedures were used between the three researchers involved, including myself. As more than one researcher was involved, training was provided so that the questionnaire was administered in a standard way each time.

### **8.2.2 Tool selection**

By using an instrument to quantify services' readiness for change, I wanted to select an appropriate measure to assess which services are most likely to succeed in change efforts towards improved quality, and those in which capacity building may be initially required. When selecting which tool to use, I considered two tools developed by ([Holt et al., 2007](#)) and ([Bobiak et al., 2009](#)).

I selected the 25-question capacity for change self-administered assessment tool ([Bobiak et al., 2009](#)) over another, as preliminary literature review findings indicated that the tool was testing for domains that were considered important (see Appendix I). The tool featured key domains such as: features of the organisational structure (e.g. practice resources), climate (motivation and perceived options for change) and culture (e.g. interpersonal and inter-organisational relationships). The tool was developed from an initial pool of 117 constructs. Following pairwise correlations and Rasch modelling to test validity of the constructs, the remaining 25 constructs were those with a conceptual relationship to readiness for change, greatest clarity in wording, the best fit within the Rasch model and the highest discriminative ability between practices to be sampled. In addition, convergent validity of the tool was assessed against independent assessment.

When selecting the tool, I tried to aim for maximum validity and reliability to achieve a degree of scientific rigour. Validity is the extent to which the questionnaire actually measures what it is supposed to measure ([Carmines and Zeller, 1979](#)). The most common are face validity, content validity, criterion validity and construct validity. Face validity refers to the belief or perception by individuals using or being assessed by an instrument that 'on the face of it' the instrument measures what it is purporting to measure. Content validity is the determination of the content and the representativeness of the items 'contained' in an instrument. It is considered to be a more rigorous test than face validity. Construct validity traditionally refers to the extent to which the measurements obtained by an instrument correlate with expectation (often unobservable qualities). Reliability has been defined as the extent to which the same measure gives the same results in repeated occasions ([Carmines and Zeller, 1979](#)).

When considering the appropriateness of the tool, I had to consider who the respondent would be. As part of the eLSU study, managers of clinical services or wards would be expected to complete the tools, as literature indicated staff in management roles were equipped with the necessary oversight to make judgements relating to readiness for change ([Lehman et al., 2002](#)). I embedded

this tool as part of the manager's checklist process as I wanted it to be completed by each Low Secure ward manager participating in the eLSU study. The tool was easy to implement, and could be undertaken in a quick and timely manner. There was an emphasis on finding short, readable questions that were clear and unambiguous. I felt this was especially important as completion is likely to decline if the tool has too many items which are not considered relevant to the respondent.

In healthcare, there is no 'gold standard' for an acceptable response rate. McColl and Thomas ([2000](#)) suggest that a response rate equal or greater than 85% is excellent, 70% to 84% can be considered very good, with a rate of 60% to 69% being acceptable. A barely acceptable response rate would be below 60%, with a rate that falls below 50% being considered unacceptable.

### **8.2.3 Amendments to the tool**

The use of 'language' particularly 'jargon' was considered in relation to the target group. As the tool was originally developed in the United States for use in primary care facilities, I sought permission from the authors to alter some of the spelling, terms and language to make it suitable for use in the study (see Appendix A). I optimised the design and wording of questions to administer in this study setting. For example: as the tool, would be tested in LSUs in a forensic inpatient mental health environment, I replaced the word 'practice' with 'unit'.

### **8.2.4 Data collection methods**

Myself and another researcher undertook a pilot visit, aimed at testing the acceptability and to observe if there were any difficulties in understanding the Likert scale questions. During this visit, it became clear that it was easier for the tool to be completed in the presence of the researcher in case respondents had any questions or needed clarify a question. This was done on most occasions throughout the eLSU study. On seldom occasions where this was not possible, then it was made very clear from the onset that the researcher could be contacted by via phone or e-mail at any point, if clarification was required.

The tool used in both baseline and follow-up measures were identical. Composite scores for each service were compared between two time intervals: at baseline pre-intervention, and after 12 months for both active and control arms of the randomised controlled evaluation.

### **8.2.5 Data entry**

I used the software packages version 22 of SPSS Statistics (Statistical Packages for the Social Sciences), Microsoft Excel and SAS 9.4. I manually entered and managed data were entered in SPSS version 22, as it was originally collected in the form of paper questionnaires. Single-transfer coding (where the response is already in the form which should be entered in to the computer) was used to enter individual the Likert scale items and then I calculated a composite score for each questionnaire. I carefully compared data entries directly against the original questionnaires and generated 26 variables for each case.

#### **8.2.5.1 Missing data**

I used the signal code of 999 for missing data. SAS 9.4 showed the value specified as missing data and dealt with it intelligently; by computing averages based only on the data present.

### **8.2.6 Data analysis**

I used SAS 9.4 for data analysis, as it had enhanced features for fitting likelihood-based models. Two of the features included mixed linear regression models and mixed nonlinear regression models, which was the reason I chose to use this software package. Unlike the other general purpose packages, SAS 9.4 also allowed for fitting very complex multilevel models (MLMs) and calculated corresponding statistics. MLMs can be conceptualised as regression models occurring at different levels. There are many types, which differ in terms of the number of levels e.g. wards nested within hospitals or clinical services, type of design (e.g. longitudinal with repeated measures), scale of the outcome variable (e.g. continuous or categorical), and number of outcomes (e.g. univariate or multivariate).

Modelling techniques to incorporate ward level data enable inherent correlation within clusters to be modelled explicitly. I used these methods to incorporate the hierarchical nature of the data into the analysis ([Grol et al., 2004](#)). I could have undertaken ward level analyses using adjustments to simple statistical tests to account for the clustering effect. However, this approach would not have allowed adjustment for practice characteristics. I chose multilevel analysis, as it was considered by some as a better method for the analysis of data with complex patterns of variability, with a focus on nested sources of variability (i.e. models that have been fit using the same data and where one model is a subset of the other). In this study, managers who completed the readiness for change tool and QELS checklist belonged to wards (level 1) which were nested within services (level 2), and may have had covariates at the practice level (such as practice size) ([Bell et al., 2013](#)). In the analysis of such data, it is usually considered illuminating to take account of the variability associated with each level of nesting, as one may draw the wrong conclusions if variability is ignored. As such, multilevel models have been developed to properly account for the hierarchical (correlated) nesting of data ([Heck and Thomas, 2000](#), [Hox, 2010](#), [Klein and Kozlowski, 2000](#), [Raudenbush and Byrk, 1992](#), [Snijders and Bosker, 1999](#)).

I first used frequency counts as a simple means of exploring the data and observing central measures of tendency. I then derived logit scores using Microsoft Excel for the readiness for change composite scores (according to the original paper ([Bobiak et al., 2009](#))) to carry out parametric regression analysis.

The outcome variable I used in this study, was a calculation of the change from baseline in QELS checklist score ([Aimola et al., 2016](#)). I calculated this using the difference between log transformed baseline and follow-up QELS checklist scores. The explanatory variable was a composite transformed baseline readiness for change score.

There appeared to be a small clustering effect when looking at colour coded wards (separated by treatment arms) in exploratory scatter plots; whereby

some of the wards of similar services were grouped closely together adding strength to the argument to undertake hierarchical modelling.

I specifically used the PROC MIXED procedure in SAS 9.4 oriented toward general mixed linear models, as it could analyse practically all hierarchical linear model examples for continuous outcome variables. The generalised mixed model orientation was also advantageous as crossed random coefficients could be easily included. I first had to adhere to the formatting requirements of the data files which would be used to conduct multilevel analyses using PROC MIXED. The data file needed to be wide whereby one row of data for each observation with each variable in the data file was contained in separate columns.

Firstly, I used Pearson correlation to obtain a scatter plot of baseline readiness for change scores against baseline QELS checklist scores. This was to observe the spread of data, any initial differences between control and intervention wards at baseline.

To compare the difference in the mean value of three or more independent groups simultaneously, an analysis of variance (ANOVA) which is a parametric test, can be used. In this study, ANOVA was suitable as the outcome variable was a continuous normally distributed variable, and the explanatory variable was categorical with three or more groups. In building statistical models, it is assumed that some of the coefficients are fixed and others are random. Random effects, which are estimated as variance components, are model parameters that are estimated to vary between higher level units, whereas fixed effects are estimates that are modelled to not vary between higher level units.

To formally test the correlation coefficients derived from Pearson correlation, I began to build the linear regression model, random effects ANOVA. At the minimum, a 2-level linear model estimated in PROC MIXED will usually include one random effect. This is considered the key difference between 2-level linear models and single-level ordinary least squares (OLS) models. If the assumption

of random selection has not been met, the correlation coefficient does not describe the true association between two variables that would be found in the general population. Only the coefficients from random population samples have an unbiased value and can be compared with one another.

Although different sources provide different guidelines on the model building process when estimating multilevel models, they all have the same goal, for researchers to estimate the most parsimonious models that best fit their data. I began with an unconditional model, which is a model that has no predictors, and used it to calculate the intraclass correlation coefficient (ICC) which estimated the relationship between baseline readiness for change and change from baseline in QELS checklist score, accounting for differences in control and intervention groups.

I gradually estimated a more complex model while checking for improvement in model fit, after each model was estimated. A major difference between model building approaches related to the fit statistic that was used to assess that the model fit, which depended on model estimation. I considered two main options for estimating parameters in this model: maximum likelihood (ML) and restricted maximum likelihood (REML) as well as whether the models were nested or not. I estimated the model using REML, whereby smaller values represented better fitting models. In estimating the variance components, REML is considered to be less biased than ML, which is one of the reasons why I selected it. In ML, the regression coefficients and the variance components are included in the likelihood function so the fit of entire model is described. In REML, variance components are calculated after the fixed effects have been removed from the model and describe the fit of the random effects ([Hox, 2010](#)). In most cases, the difference between the estimates produced by the two methods is minimal ([Hox, 2010](#)). However, for small samples (in this case) or when the number of fixed effects is large, REML is preferred ([Hayes, 2006](#)).



To observe if the same respondent completing the QELS checklist and readiness for change tool at baseline or follow-up influenced the difference, I continued the model building process by including this as a random effect. I additionally went on to investigate if services in the intervention group of the eLSU study demonstrated higher readiness for change at baseline than those in the control group, as this was indicated in preliminary qualitative findings.

## **Chapter 9 Results of Focus Groups**

In this chapter, I present the results from thematic analysis of the data from focus groups I conducted with practitioners (staff who coordinate external peer review programmes at the CCQI). The findings are largely structured around themes that were generated from the systematic literature review, as they arise within research aims. Furthermore, to preserve the anonymity of the individual respondents, names have been removed, and replaced with identifying codes.

### **9.1 Practitioner Demographic Characteristics**

Table 9.1 presents a summary of the demographic data collected from practitioners at the start of focus groups (see Appendix B). This includes their gender, tenure and which CCQI programmes they had worked on previously.

**Table 9.1 Practitioner characteristics**

Session	Gender	CCQI tenure (years)	Programmes worked on since start
1	F	1 - 5	Accreditation for inpatient mental health services, learning disability (AIMS LD)
2	M	5 - 10	Quality network for inpatient child and adolescent mental health services (QNIC), quality network for community child and adolescent mental health services (QNCC), quality network for perinatal mental health services (Perinatal)
	F	5 - 10	Unit admin, Unit office manager, closing the gap through clinical communitites (CTG), QNFMHS, eLSU study
	F	1 - 5	Memory services national accreditation programme (MSNAP), home treatment accreditation scheme (HTAS), electro-convulsive accreditation service (ECTAS)
	F	1 - 5	Accreditation for inpatient mental health services (AIMS)
	F	<1	QNFMHS
3	F	1 - 5	Electro-convulsive accreditation service (ECTAS), quality networks for eating disorders (QED), Accreditation for inpatient mental health services (AIMS)
	F	1 - 5	Community of Communities (C of C), Quality mark for Elder Friendly Wards (QM)
	F	>1	eLSU study
4	F	1 - 5	Accreditation for inpatient mental health services (AIMS), national audit of schizophrenia (NAS), Prescribing Observatory for Mental Health (POMH-UK)
	F	1 - 5	QNFMHS
	F	>1	QNFMHS
5	F	<1	Community of communities (C of C)
	F	5 - 10	Electro-convulsive accreditation service (ECTAS), Accreditation for inpatient mental health services (AIMS), home treatment accreditation scheme (HTAS), Memory services national accreditation programme (MSNAP)
	F	1 - 5	National audit of schizophrenia (NAS), Prescribing Observatory for Mental Health (POMH-UK), national audit of psychological therapies (NAPT)
	F	5 - 10	Community of communities (C of C), enabling environments (EE)

## 9.2 Overview of Themes

Table 9.2 presents an overview of the themes categorised by the thesis research aims that were generated from focus group data.

**Table 9.2 Focus group themes overview**

<b>Aim 1: Programme Factors</b>	<b>Aim 1: Causal Mechanisms within Programme Factors</b>	<b>Aim 2: Outer Contexts</b>	<b>Aim 2: Inner Contexts</b>	<b>Aim 3: Readiness for Change</b>
Before joining	Networking	Types of programme	Membership length	Indicators
Joining	Training	Healthcare System	Non-financial resources	Leadership
Self-review	Standards	Economy	Local context	Communication
Visit	Peer-aspect	Mandatory	Joining influences	Understanding
Verbal feedback	CCQI organisation			Engagement
Written feedback	Additional staff efforts			Connections to others

### 9.3 When Change Happens

Practitioners were asked to identify when they thought change most likely occurred during the external peer review programmes. They discussed that often realisation from services that they were not meeting certain standards occurred before or during the self-review stage. This was discussed to make it difficult for the CCQI to measure change. Practitioners agreed that services often make changes, those seeking accreditation before self-review. Most practitioners often felt that a lot of services seeking accreditation did not want to sign up to external peer review until they were sure of being able to achieve accreditation. Some practitioners provided accounts of services that had gone through the standards (which are freely available online) prior to joining, and had made changes to prepare.

“So, in some respect we never really get an accurate baseline because people start changing the minute that they see the standards, I think.” [FG 2b]

In contrast, some practitioners felt that services participating in peer review networks often left the self-review stage to the last minute. This highlighted that although change could have also occurred during self-review, it was in a different manner to those seeking accreditation.

Some practitioners often felt with accreditation programmes such as AIMS or ECTAS, most of the changes occurred after self-review but before the visit.

“...they'll submit all their self-review data and while they're doing that they'll obviously obtain a copy of it and see that they're not meeting this standard and not meeting that standard and then before the peer review visit takes place before people come and visit they think 'oh my God, yeah we need to get this sorted, this person needs to go on training, we need to get this done in our documentation, we need to get this signed off'. So, that all happens, and once the peer review visit comes in then the peer reviewers can then look and say 'OK you're not meeting this standard', and they'll say 'oh well actually since the self-review we've done this and we've now changed it'. And that's why a lot of standards for AIMS and ECTAS it

goes from not met to met during that time, and that's what gains them the accreditation." [FG 3a]

This was also perceived in instances where both services belonging to peer review network and accreditation programmes received compiled self-review data that was entered from different members of staff and departments. They were reported to make changes between their self-review and their peer review visit, once they had looked at the overall picture.

Some practitioners felt that most of the changes occurred during or after the feedback stage, as sometimes the momentum for change was stronger.

"...but during the peer review they think 'oh my God', you know there's something we have to do about this. So during the feedback kind of felt this insight into their own problems, so I think that's even stronger than the self-review stage." [FG 3c]

One practitioner felt that once services had received their feedback, they began to digest the experience and the ideas generated from the day. After a few weeks have passed, that is when they begin to start thinking about change.

It was broadly agreed that change could be much more of an iterative ongoing process in peer review networks than accreditation programmes.

The preparation stage was thought to be essential by three focus groups. However, practitioners did pick up on subtle differences between how services belonging to peer review networks and accreditation programmes underwent preparation. There was a sense that much of the preparation for change in accreditation programmes were undertaken at the time of joining or even sometimes prior to joining, and this was often before self-review. Where as in contrast services belonging to peer review networks mostly prepared for changes after receiving feedback. Preparation for self-review was also found to be an informal indicator to some practitioners of how engaged a ward or service was in the programme, and several practitioners stated that services could do better or find more benefits from programmes if they were better prepared.

Feedback was generally mentioned by two focus groups, with a focus on how written reports were used to gain support from senior management in services. However, one practitioner highlighted occasions where there had been differences between verbal feedback delivered on the day and the following written report. In some instances, feedback was felt to be a more substantial change agent than the self-review stage as it was being delivered by a peer. However, other practitioners in another focus group suggested the opposite, that lead reviewers could often be too heavy handed with their feedback.

## **9.4 Mechanisms within Programme Factors**

Practitioners identified possible programme factors which could be linked with causal mechanisms that would be later tested through CMOCs in semi-structured interviews with subjects working in member services.

### **9.4.1 Networking**

Networking was one of the most essential causal mechanisms, mentioned by all practitioners. Sharing resources and learning from one another was deemed to be the most essential aspect of the programme by practitioners, as they felt this was how those participating in peer review gained the most value to instigate change. Practitioners who had attended peer review visits gave accounts of instances where they had witnessed idea exchanges.

“...it’s the cross-pollination and dissemination of ideas throughout the network which I think is really valuable.” [FG 3b]

Services were ‘all in the same boat’ and ‘all struggling’ were common phrases used. However, one practitioner said that if all the services were struggling with the same problem, in some cases the group mentality was to accept the struggle instead of trying to improve:

“But I suppose the main issues are colluding, and if they’re all struggling with something then they all just seem to write it off, ‘well we all struggle with that’, instead of stopping and thinking about it.” [4b]

### **9.4.2 Training**

Training differed between CCQI projects, with some programmes only providing lead reviewer training and others providing training for all reviewers. This theme was mentioned by all practitioners. Some projects regularly updated their training whereas others did not. There was a balance appreciated by practitioners between including feedback and updating training sessions, whilst ensuring there was continuity and a level of standardisation between cycles. The effectiveness of the training was discussed with some practitioners stating



they were unsure how much impact the training had. The time available to deliver training was limited, and one practitioner especially felt a reviewer training week would be especially helpful. However, many practitioners agreed as training was only delivered at the CCQI's London office, this was a geographical and financial barrier (context) that had to be considered.

### **9.4.3 Standards**

Standards were a frequently mentioned key mechanism by all practitioners. However, there were differing opinions with regards to CCQI standards. Comparisons were made to guidelines from other bodies such as National Institute for Health and Care Excellence (NICE), and a few practitioners discussed how services deemed some CCQI standards to be inessential or that different services had standards which were particularly salient to them. On this topic, there was much discussion around the ambiguity of standards, unnecessary or unachievable standards, and a need to develop specific standards for more specialised services, which could help this issue.

In cases of standard ambiguity, most practitioners reported this feedback to their team meetings, but also perceived the nature of some of the standards was subjective.

“Yeah. I mean we have written them so they're not like that. But even so, it's something you might not think is ambiguous and then people go 'actually, that could mean this'. I mean there are, so when we find one of those, we just make a note to revise the wording at the next standards revision. We're having a big standards revision this month, for all our branches and projects.” [Interview 1a]

Many practitioners agreed that CCQI standards began as basic and became harder once services were compliant. The three 'types' of standards (type 1, type 2 and type 3) were frequently mentioned. Focus groups highlighted differences between CCQI projects regarding if standards were regularly reviewed or remained fairly static.

#### **9.4.4 Peer-aspect**

Peer reviewer qualities was the most frequently mentioned programme factor across three of the focus groups. The differing personalities of peer reviewers was often considered to be a weakness of external peer review programmes.

Focus groups uncovered that different CCQI projects had different meanings of 'peer', some felt it was appropriate to include practitioners as peers on the review team, as they were stakeholders of the programmes; whereas others felt it should only be subjects working in services and there was debate about the use of service users and carers from practitioners from different projects. Practitioners who had attended peer review visits exchanged their different experiences. Some projects required practitioners to take a much more active role in the visit than others who were restricted to activities such as timekeeping and ensuring the purpose and credibility of the programmes were upheld. Logistics of this were also discussed across focus groups, such as the practitioners having to mitigate against different agendas of peer reviewers.

#### **9.4.5 CCQI organisation**

There was a sense across three focus groups that CCQI collaboration between projects was a key mechanism. Some practitioners expressed a concern that despite regular quarterly meetings there was not enough communication or knowledge and resource sharing taking place at the CCQI, and if this was improved it could facilitate the CCQI's ability to assist improve quality for their member services. Examples were provided of development of new standards despite similar sets used by other projects. As some wards participated in more than one CCQI quality improvement programme it was felt that knowledge sharing between CCQI teams, could help make things easier for those wards or services. Practitioners working on accreditation projects also discussed why there were differences between lengths of accreditation cycles, and that the CCQI should question these historical differences between projects instead of 'accepting them as a given'. More 'cross-pollination between CCQI programmes' was thought to be needed by practitioners.

#### **9.4.6 Additional staff efforts**

A frequently discussed theme across two focus groups was that accreditation practitioners had sometimes tailored the programmes for 'bad' or 'dangerous' wards to help them achieve accreditation:

“There are some of our services, I don't know if it's the same because we focus on them because they are quite bad so we actually do a lot more quality improvement work with them, than interventions and sharing of information because we have to sort this ward out because it's dangerous or whatever.” [FG 5b]

This theme of tailoring the programmes was echoed by other practitioners in other focus groups reported examples of providing extra support to help services which resulted in beneficial outcomes.

## 9.5 Outer Contexts

The most frequently mentioned outer context seen to influence the programmes were the type of programme, healthcare contexts of whether services were privately or NHS-provided and economic pressures faced by nearly all services.

### 9.5.1 Type of programme

Practitioners across all focus groups discussed the differences between accreditation and peer review networks. The 'carrot and stick' metaphor was used on two separate occasions, whereby peer review networks were usually considered to be the carrot, and accreditation acted as a stick. The interview practitioner considered accreditation to be both a carrot and a stick, given that if a service fails to meet accreditation standards then the consequences could be severe and result in targeted CQC involvement with the possibility to be shut down. The practitioner also described the advertising impact of accreditation:

“CQC will hear about that as well, and then they're less likely to inspect, and also you can put that on your newsletters that this ward's been accredited.” [Interview 1a]

Peer review networks were overall considered to be more supportive with practitioners as they occurred more frequently and practitioners felt with each cycle, the improvements are visible. Practitioners across two focus groups felt that accreditation was less supportive than peer review networks as there is perhaps more pressure to show evidence of meeting standards which can make the programmes quite restrictive. Practitioners from another separate focus group felt there was less networking present in accreditation than in peer review networks, as there is a higher focus on meeting standards than sharing best practice.

One practitioner eluded the difference between peer review networks and accreditation could also be due to the differing amounts of data that are requested during self-review. Some CCQI accreditation programmes requested additional questionnaires from service users and clinicians to enable more

accurate judgements to be made. Another practitioner felt that different CCQI projects required differing rates of preparation for accreditation, therefore the meaning of being CCQI-accredited as not equal across different projects.

Some practitioners who had experienced working on both peer review networks and accreditation programmes felt that the rate of change across time was different in both types of programmes. In accreditation, practitioners felt that there was a big jump in quality improvement the first year in terms of how people perform after joining, and then continue to improve, but at a much slower rate. It was felt that accreditation pushes very quick improvement in a short space of time, but then does not do much for the next few years. Once accredited, practitioners described that the CCQI trusts that the ward is compliant with key standards and it is left alone provided interim reports are received every year about how things have changed in the service. With most accreditation standards, practitioners felt if they were initially met, it is assumed that services will continue to meet them over the next two or so years. In contrast to peer review networks which were perceived to have a steadier rate of change.

“Whereas peer review is being part of a club and just kind of having that maintenance and almost like having your boiler serviced every year. It’s a bit like that, checking everything’s still OK.” [FG 2b]

One practitioner felt that accreditation did not assess the likeliness of services to improve their quality in the future.

“No because we don’t look for whether they’ll improve or what we do isn’t necessarily to improve it’s to get their accreditation or their certificate.” [FG 5b]

### **9.5.2 Healthcare system**

Practitioners discussed rivalry between NHS and private beds was a contentious issue which services were currently facing. This was felt to affect some of the peer review visits where hosts were from one type of healthcare

organisation, and reviewers were from another. In addition, it was perceived by practitioners that the method of making changes differed between NHS and private services, with changes in NHS services being a lot slower. However, a few practitioners felt, especially in children's services, that training and subject motivation was better in NHS services than in private services, which indicated to them that quality improvement would be better achieved in NHS than private services. Most practitioners agreed although this was an influential context, it did not affect how most services conducted themselves throughout the programmes.

### **9.5.3 Economy**

Practitioners discussed how budget cuts had both adversely and positively affected services. Some services were forced to drop out of the voluntary CCQI programmes as they could not afford the membership fees, and many services gave budget limitations as a reason for why they could not enforce recommendations from external peer reviews. However, many practitioners gave examples of how the economic pressures faced by some wards had forced them to come up with creative solutions, which helped them to bring about changes necessary to achieve quality improvement, and this was deemed impressive by the practitioners.

#### **9.5.4 Mandatory**

Legal implications were mentioned across two focus groups and one interview. In some cases, it was a legal requirement to join a CCQI programme or achieve accreditation in a certain area, which motivated services to succeed in programmes.

“...prison therapeutic communities can get closed, if they don't get accredited through the C of C process they will get closed if they don't improve within a certain amount of time afterwards so yeah definitely they are very motivated to pass.” [FG 3b]

Another aspect which was discussed in reference to being a barrier of change was a practitioner fear of litigation, especially in children's services.

## **9.6 Inner Contexts**

Membership length was important to most practitioners in all focus groups, with varying thoughts on the potential benefits and engagement levels of new and long-standing members to external peer review programmes. Non-financial resources, local context and joining influences were also considered to be important, as they were perceived to act as facilitators or barriers to achieving change through participation.

### **9.6.1 Membership length**

Membership length was discussed by many practitioners in all focus groups. There was a contrast in views, as some felt that early joiners could be self-selected, especially in voluntary programmes, thus had higher achievers. Whereas others thought that as membership increased, trust, understanding and familiarity of the programmes also increased, which in turn led to an increase in quality improvement and compliance with standards.

There was a large consensus among practitioners that long-standing members were less engaged, especially in mandatory programmes. This could be identified by not taking the programmes seriously and copying and pasting the same self-review information in each cycle without making the effort to undertake the programmes rigorously. Practitioners on some services highlighted that projects attempted to mitigate against this problem by reviewing the standards where possible as the presence of long-standing members was deemed as beneficial to newer members. One practitioner found in their experience that long-standing members welcomed more challenges than newly joined services.

### **9.6.2 Non-financial resources**

Non-financial resources were perceived by some practitioners as a key context and reason to why more changes could be observed in some services than others. Resources such as an old building structure were often quoted as a



hindrance for services joining the programmes, or perhaps being able to meet all the standards.

“But I’ve been to a few services that really stick out in terms of excellent examples of really good practice. They have a completely different approach to how they work, and they’re thinking outside the box. I’ve been to one service where their inpatient ward team is the same team as their home treatment team. So the patients will come into the ward, they’ll be treated on the ward and then the same staff will visit them in their home to facilitate their recovery. Do you know, that sounds really expensive, but, they’ve only got one building. They save all the costs of second building, they save all the other costs of another staff team, and actually it works better because you’ve got the continuity and the staff know the patients already, going back to their houses with them. It’s just brilliant, so simple and so brilliant. It probably wouldn’t work in a city like London because the turnover’s too high, but this is up North. It was a joy to review it. Not necessarily like that, but I wish more services had time to step outside the box, about the way they did things.”  
[Interview 1a]

In one focus group, there was discussion surrounding how participating in external peer review programmes can save resources, but there was not much consensus surrounding this.

### **9.6.3 Local context**

The theme ‘it works in different ways for different people’ was found in one focus group, regarding the suitability of different CCQI programmes for the nature of specific services or for different clinical service user groups. For example, practitioners discussed how the busy nature of acute wards with a shorter length of stay make engagement and communication difficult to facilitate, so the accreditation approach might be better compared to longer stay services.

“Whereas I know longer stay services, like the learning disability services actually, [accreditation] didn’t seem to work for them and it works on a more gentle slope, sort of a shallower learning curve. Just depends on the service.” [Interview 1a]

#### **9.6.4 Joining influences**

The decision to join was mentioned by all practitioners in all focus groups. There was extensive debate and discussion over the future success of wards and services if they were not involved in the joining decision (it was either mandatory or the decision was made by senior management from their organisation). Practitioners often experienced that services appeared to be more engaged with programmes if the decision to join was bottom-up, and came from nursing or ward staff instead of senior management, as they could understand how the programmes could be of benefit to them as oppose to being told they had to join something which would require additional time and effort. To mitigate against this, some CCQI projects obtained contact details of host contacts from the ward level, as a requirement of the joining process to help facilitate the success of the programmes.

The main reasons to join accreditation were perceived to be validation, as it often gave services a competitive edge or provided certification that they were meeting nationally recognised standards.

Furthermore, peer review networks and accreditation programmes also considered to provide an opportunity for geographically isolated and specialist services to gain help and support that they would not have had otherwise.

Motivation was a dominant context throughout the focus groups. Different motivations were discussed: the motivation to join, the motivation to carry out tasks necessary for the programmes such as completing the questionnaires, the motivations needed to enact change and implement recommendations and the motivation to continue to improve quality. It was often felt perceived that the prospect of being 'visited' could act as a motivating context if it was felt to be supportive, and not inspectorial or critical, as this had the opposite effect.

"I'd say one of the things that motivates them to do something is when they've gone to visit somewhere else and see what there is..."  
[FG 4c]

## 9.7 Readiness for Change

### 9.7.1 Indicators

All practitioners were asked if they were aware of any indicators of capacity to benefit from the programme. Nearly all practitioners felt that there were strong indicators well before the visit had taken place. Many practitioners felt that it was easy to tell during early stages, such as joining or self-review. Details such as who the host lead was and their communication approach with the CCQI team, levels of preparation, how the self-review was completed were strong indicators. If the host lead was a senior manager, practitioners felt in most cases they were too senior to be adequately engaged in the programmes and disseminate the necessary information to frontline staff, and conversely if ward managers or a nurse was the host contact, this indicated a higher capacity to benefit from the programme.

“Yeah if it comes from, for us if it comes from the partner or the head of the trust, services don’t know what’s going on. Then that’s it you’ve lost them because it’s coming from...” [FG 5d]

Demonstrating preparation before the visit, either during or before self-review was felt to be indicative of a higher capacity to benefit from the programme. Examples were discussed from practitioners who had experienced this with services preparing for accreditation, ensuring they were meeting the standards before joining to achieve successful accreditation.

Practitioners from one focus group highlighted that a clear capacity to benefit was dependent on who completed the self-review data. In some instances, they could observe ‘an argument’ in the self-review data, and these dynamics were indicative of how unlikely it would be that the ward would succeed through the programme. Often it was observed that missing, incomplete, poorly completed, or duplicate self-review data from previous years was indicative of a lack of engagement in the programme, and indirectly capacity to benefit from the programme. Some CCQI accreditation programmes had stricter rules and would cancel the peer review visit (but would still charge the service between £600-

800) if self-review data were not adequately completed. Practitioners agreed this was a good incentive for services to complete their self-review stage appropriately.

### **9.7.2 Leadership**

Leadership was another substantial readiness for change construct, mentioned across three focus groups.

“I think you do need a strong unit manager who’s really bought into the process, we had one with the PICU project who sort of organised the whole team and AIMS would be on the agenda for every team meeting and he’d get the whole team involved, so different people had different responsibilities in different areas and they were an excellent unit, and we went to see them so we could really see that they’d put the thought and the work into the standards and were thinking ‘we don’t meet this and how can we, evidence that we can meet it.’” [FG 4a]

A lot of practitioners coordinating CCQI programmes, perceived good leadership was such an essential prerequisite and their programmes would only allow a service to join if a suitable host contact who could fulfil an appropriate leadership role was identified. The theme of organisations signing up multiple wards or units in one go was highlighted. Where practitioners felt it was still necessary to identify a host contact for each ward or unit, they would specifically request this from host services:

“...it could just be one clinician in that service that joins them all up, it doesn’t mean the rest of the staff are on board with it at all. I mean you’d hope in an ideal world it’s a lovely, collective decision but it’s not always and we’ve found that, certainly on our work in the quality mark that in the first round of data collection we had all these heads of nursing and chief executives signing up their services and putting them forward and the wards were like ‘are you kidding me? We’re so short staffed, we can’t do this at the moment, we don’t have the time or the attention.’” [FG 3b]

Charge nurses or ward managers were deemed to be good host contacts in leadership positions as they spend more time with services users and other ward staff than senior managers.

### **9.7.3 Communication**

Communication was perceived to be a substantial readiness for change construct for several practitioners, as they felt lack of communication could really impede achieving success from the programme. More communication between host contacts and the practitioners earlier in the programme, perhaps before joining, during the joining process or during self-review signified interest and engagement in the programme and was indicative of their capacity to benefit from the programme. Communication between the host contact or the person who influenced joining and the remainder of the staff group on the ward was felt to be essential. Practitioners reported examples where a lack of communication between staff in a host service was identified through their behaviour during self-review such as sending duplicate information through to the CCQI team or failing to respond to requests for data during self-review. Although information was often relayed to host contacts, examples were highlighted where this information was not always communicated to the rest of the ward staff which was perceived to fragment the programme.

### **9.7.4 Understanding**

Understanding the programmes was mentioned by practitioners across three focus groups. A few practitioners experienced that a lack of understanding of the programmes could cause unnecessary anxiety and stress which would in turn act as a barrier to achieving success:

“So, in AIMS we have three different types of standards, ones a type one which you have to meet otherwise you can’t be accredited, and there’s type twos which you should be meeting and type three which are excellent if you’re meeting. But sometimes you go to somewhere and they kind of hadn’t grasped this, and they’re getting quite stressed because there’s this type three that they’re not being able to meet and that’s fine but they haven’t understood that part of the process...” [FG 4b]

“What I found a little bit frustrating was that it showed that they didn’t really understand the process, because some of the comments were ‘we’re putting this in place now so it’s met’ and ‘no if it’s not done then it’s not met’, as it has to be met at the point of peer review, and we tell them to agree to that at the start.” [Interview 1a]

It was felt that most non-senior staff did not have a thorough understanding of the programmes. One practitioner stated that this was more of an expectation, and explained how she mitigated it on visits:

“And often as part of the peer review we’ll meet with as many of the staff-team as we can. A lot of the time there’s been peer reviews and there’s been hardly anyone turn up to those meetings because they haven’t known what’s going to happen, and no-one’s told them what it’s for, and they’re sitting there really confused. So whenever I’m on a peer review I ask at the beginning of every meeting ‘Do you all know what I’m here for? Do you know what we’re here for? Do you know what the AIMS project is?’ And they normally just give a dummy-nod. I always kind of phrase it in a ‘don’t worry if you don’t know what we’re about, I’m here to explain to you,’ so for anyone that needs an explanation...” [Interview 1a]

Most practitioners agreed that understanding in services needed to be addressed and improved.

### **9.7.5 Engagement**

Engagement was felt to be a necessary readiness for change construct by nearly all practitioners to succeed through the programmes. Often some practitioners felt that lack of engagement was down to poor communication or leadership, as these themes were strongly linked, with some reporting examples of difficulties in engagement as they had not been informed correctly through communication channels or by the appropriate leader. Many practitioners perceived that quality improvement through the external peer review programmes could not take place without appropriate engagement.

“...I mean the wards that don’t do very well are the wards where the nurses don’t engage with patients for whatever reason, and quality engagement.” [FG 5b]

Some practitioners perceived if staff working in services were more engaged, they would be able to partake in more discussions which would help to facilitate quality improvement and a continued commitment.

One practitioner felt that staff working in services did not have the 'time' to be engaged in the programmes, due to the busy and complex nature of services.

#### **9.7.6 Connections to others**

Connections to others was seen to be an essential readiness for change construct across two focus groups. Practitioners highlighted how one unit can do well on a site but it does not always radiate from that unit to the rest of the hospital. It was felt to depend on the Trust or organisation, as some wards do not even communicate with their neighbouring ward. This was seen to be in stark contrast to other examples given to other wards who take good practice from one ward and one hospital to another, and they learn how to improve care. Practitioners in one focus group discussed possible methods they could use at the CCQI to facilitate more of this collaboration promotion as it was beneficial to the success of the programmes.

Many practitioners reported that services were 'clinically' isolated, especially some low secure or learning disability services, and the programmes enabled them to feel connected to other services and be part of a larger organisation.

## 9.8 Conclusions

Practitioners identified clear differences between when change occurred in accreditation and peer review networks. With services undergoing accreditation, they identified most changes taking place before the visit, with some services only joining after having undergone changes or identifying which changes to make. Some highlighted the difficulties this causes the CCQI when trying to measure change. Practitioners also identified that services participating in a peer review network mostly prepared for changes following feedback.

Networking was one of the most essential causal mechanisms within a programme factor by practitioners and was thought to hold the most value for those participating in external peer review. Training, standards, peer-aspect, CCQI organisation and additional staff efforts was also considered to be important.

The most frequently mentioned outer context seen to influence the programmes were the type of programme, healthcare contexts of whether services were privately or NHS-provided and economic pressures faced by nearly all services.

Membership length was discussed by many practitioners in all focus groups, with varying thoughts on the potential benefits and engagement levels of new and long-standing members to external peer review programmes. Non-financial resources, local context and joining influences were also considered to be important, as they were perceived to act as facilitators or barriers to achieving change through participation.

All practitioners were asked if they were aware of any indicators of capacity to benefit from the programme. Many practitioners felt that it was easy to tell during early stages, such as joining or self-review. Leadership was perceived as an essential prerequisite and practitioners detailed how this construct factored into the joining process, with regards to how they determined the host contact. Communication, understanding, engagement and connections to others were



also seen as key constructs of readiness for change which may influence how services could make best use of participation in external peer review.

Findings from focus groups were used to develop the semi-structured interview topic guide. Programme factors, causal mechanisms and contextual factors identified through focus groups with practitioners were used to develop and later test CMOCs. These were tested through semi-structured interviews with subjects working for services that participated in external peer review programmes.

## **Chapter 10 Results of Semi-Structured Interviews**

In this chapter, I will present the themes that were generated from framework analysis of the qualitative semi-structured interview data from subjects who belonged to services that were members of 3 CCQI external peer review programmes (low secure QNFMHS, medium secure QNFMHS, and AIMS PICU).

### **10.1 Service Characteristics**

As presented in Table 10.1, I sampled subjects from 18 services provided by NHS England, four by NHS Wales, one by NHS Scotland, and 12 private organisations. I conducted semi-structured interviews with subjects who worked across AIMS PICUs (12), combined MSU-LSU sites (5), MSUs (10), LSUs (7), and one LSU which had changed from medium to low secure. 17 services were from urban location, and 19 rural; 20 services were in the North, and 16 were situated in the South of the UK. 19 services were characterised as being small services (with less than 30 beds), and 13 were large services.

I included 12 services that had been members of a CCQI programme for over five years, and 24 that had been members for less than 5 years.

Subjects' tenure ranged between one and ten years. There was a large variety in terms of experience; some senior managers had experience of numerous external peer review programmes in both public and private sectors, and this often extended beyond those coordinated by the CCQI. This meant some of these subjects were in a unique position to be able to directly compare accreditation programmes and peer review networks, and often contrast these experiences with regulatory processes and more local quality improvement approaches. I have categorised the results from the Framework Analysis per the thesis aims in Table 10.2.

**Table 10.1 Purposive sampling of services**

<b>External Peer Review Programme</b>	<b>Private / NHS</b>	<b>Trust / Organisation</b>	<b>Urban / Rural</b>	<b>North / South</b>	<b>Size</b>	<b>Beds</b>	<b>Joining Year</b>	<b>Early / Late (post-2010)</b>
QNFMS Combined	Private	Alpha	U	N	8+ wards	164	2009	E
QNFMS LSU	Private	Partnerships in Care	R	N	1 ward	9	2012	L
QNFMS MSU	NHS England	Birmingham and Solihull Mental Health Foundation Trust	R	N	2 wards	30	2009	E
QNFMS MSU	NHS England	Nottinghamshire Healthcare NHS Trust	U	N	7 wards	90	2008	E
QNFMS LSU	NHS England	Devon Partnership Trust	R	S	1 ward	14	2012	L
QNFMS LSU	NHS England	Hertfordshire Partnership NHS Foundation Trust	R	S	1 ward	15	2012	L
QNFMS MSU	Private	Partnerships in Care	R	N	3 wards	64	2009	E
AIMS PICU	Private	Cambian	U	S	4+ wards	50	2012	L

QNFMS MSU	NHS Wales	Abertawe Br Morgannwg University Health Board	U	S	5 wards	64	2009	E
QNFMS Combined	Private	Riverside Healthcare Limited	U	N	6 wards	108	2009	E
QNFMS LSU	Private	Cygnnet	U	N	1 ward	15	2012	L
AIMS PICU	Private	Cygnnet	R	N	1 ward	14	2012	L
AIMS PICU	Private	Cygnnet	U	N	1 ward	15	22/02/2012(original) 01/08/2013 (re-started )	L
AIMS PICU	NHS England	East London NHS Foundation Trust	U	S	1 ward	10	2011	L
AIMS PICU	NHS England	2gether NHS Foundation Trust	U	S	1 ward	10	2010	L
AIMS PICU	Private	Cygnnet	U	S	1 ward	15	2014	L
QNFMS MSU	NHS England	Humber NHS Foundation Trust	R	N	2 wards	60	2008	E

AIMS PICU	NHS England	Norfolk and Suffolk Foundation Trust	U	S	1 ward	10	2012	L
QNFMS MSU	NHS England	Coventry and Warwickshire Partnership Trust	R	N	1 ward	15	2009	E
QNFMS LSU	Private	Partnerships in Care	R	N	1 ward	54	2012	L
QNFMS LSU	NHS England	Norfolk and Suffolk Foundation Trust	R	N	1 ward	12	2012	L
AIMS PICU	NHS England	Merseycare NHS Trust	R	N	1 ward	8	2011	L
QNFMS MSU	NHS England	Southern Health	R	S	5 wards	79	2006	E
QNFMS MSU	NHS Scotland	NHS Tayside	U	N	3 wards	32	2012	L
AIMS PICU	NHS England	Berkshire Healthcare NHS Foundation Trust	U	S	1 ward	14	2010	L
QNFMS LSU	NHS England	Southern Health	R	S	1 ward	28	2012	L
QNFMS LSU (was MSU)	Private	Partnerships in Care	R	S	2 wards	24	2012	L
AIMS PICU	NHS	Betsi Cadwaladr University Health Board	U	N	1 ward	8	2012	L

	Wales							
AIMS PICU	NHS England	Northumberland, Tyne and Wear NHS Foundation Trust	R	N	1 ward	14	2013	L
QNFMS Combined	Private	Partnerships in Care	R	S	4 wards	84	2010	L
AIMS PICU	NHS England	Worcestershire Health and Care NHS Trust	U	N	1 ward	9	2012	L
QNFMS Combined	Private	Priory Group	R	S	6+ wards	123	2009	E
AIMS PICU	NHS Wales	Betsi Cadwaladr University Health Board	U	N	1 ward	6	2012	L
QNFMS MSU	NHS Wales	Betsi Cadwaladr University Health Board	R	N	1 ward	25	2011	L
QNFMS MSU	NHS England	Devon Partnership Trust	R	S	1 ward	15	2007	E
QNFMS Combined	NHS England	West London Mental Health Trust	U	S	10+ wards	305	2007/2008	E

## 10.2 Overview of Themes

Table 10.2 provides an ordered overview of the themes generated from Framework Analysis. Although the frequency of subjects mentioning each specific theme throughout this Chapter is reported, this is not an indicator of salience.

**Table 10.2 SSI broad themes overview**

<b>Aim 1: Programme Factors</b>	<b>Aim 1: Causal Mechanisms</b>	<b>Aim 1: Outcomes</b>	<b>Aim 2: Outer Contexts</b>	<b>Aim 2: Inner Contexts</b>	<b>Aim 3: Readiness to change</b>
Before joining	Mechanisms within programme factors	Long-term changes	Type of external peer review	Length of membership	Awareness
Joining	Host team processes	Short-term changes	Healthcare system	Joining decision	Understanding
Self-review	Staff groups	Negative changes	Mandatory	Geographical location	Readiness for change
Visit		No change	Economy	Non-financial resources	Indicators of capacity to benefit
Verbal feedback		Perceptions		Local context	Barriers to change
Written feedback				Size	
Peer reviewing other services				Organisational culture	
Additional CCQI activities					

### **10.3 Programme Factors: When Change Happens**

Subjects were asked to reflect on their experiences of when most changes leading to quality improvement took place during external peer review programme factors.

#### **10.3.1 Before joining**

Senior and ward managers from five AIMS PICU member services reported most of their changes took place prior to joining. Subjects reported viewing standards and benchmarking themselves against essential standards prior to joining. One senior manager expressed they wished their team had done even more preparation before joining. Some had even formulated plans to ensure essential standards were met, or funding to make necessary changes had been secured prior to joining AIMS PICU.

“Unless the Trust are going to sign up to this and provide some resource for this we’re gonna fail.” [104FLAN, PICU Senior Manager]

#### **10.3.2 Joining**

Senior and ward managers from one MSU that joined the QNFMHS in 2011, and subjects from five member services of AIMS PICU reported that they had begun to implement changes immediately upon joining (prior to undergoing self-review).

#### **10.3.3 Self-review**

Subjects across all programmes reported starting to think about or undergo changes during the self-review stage. This was mostly mentioned by subjects participating in AIMS PICU; 34 of 80 examples cited were from AIMS PICUs, from 23 different subjects. This stage was cited more by senior and ward managers than frontline subjects.

“When we filled in the booklets at the beginning; when we looked at the standards, it gave us a bit of impetus to actually make changes and develop. [115SMMNW, MSU Senior Manager]



Many subjects spoke of identifying necessary changes during this stage, and reported that this was either followed up by acceptance that certain things could not be changed (such as structural resources) or initiating changes during self-review to meet the standards. The self-review stage was often reported as a quality assurance exercise to check that documents and policies were up to date and processes were in place that were required by the standards.

“When you go through the book you pick up things and you kind of think, “Oh, yeah we need to that or are we doing that and maybe we need to make sure we’ve got evidence to prove that we’re doing that... I can’t remember but I know working through the action plan or the booklet, we did put things in place because of the peer review, so we did... [I: Prior to the visit?] Prior to the visit, yeah, so we would have gone through them and what I did was I shared the booklet with the staff...” [46SMLP, LSU Ward Manager]

#### **10.3.4 Visit**

Although changes in relation to the visit were mentioned by subjects participating in all programmes, it was noticeably more by subjects participating in QNFMHS. Some subjects reported that changes because of the visit were attributed to being able to see the perspectives of peer review team.

“When you're actually going round with the team because you see it through their eyes, and because you've become quite complacent at times, because, say if it was on one ward that I used to be the manager of; I know that ward, I know it inside out, I know how it runs, I know what it looks like, but to look at it as an outsider then that's when you start to think, 'oh yeah, I can see why you'd think that now and I do need to change that,' yeah.” [24SMMP, MSU Senior Manager]

“It’s when they’re asking questions... But it's like prompting us to actually challenge people who make some of the decisions that influence our practice as well, actually, no you're right, why isn't it open?” [24SMMP, MSU Senior Manager]

Reviewers were reported to offer interesting insights that led to changes, providing an opportunity for people to reflect on their own services.

Some subjects partaking in AIMS PICU revealed a sense that visits were used as a method of validation of self-review information.

### **10.3.5 Verbal feedback**

Verbal feedback was mentioned across all programmes, by all staff groups; but mostly from senior and ward managers. Subjects reported that it was generally positive to have good practice highlighted, which provided external validation.

Subjects from low secure QNFMHS reported verbal feedback to be realistic, containing useful suggestions and ideas for change. Often verbal feedback highlighted issues services had not thought of, and in some instances also identified matters they were already aware of. Subjects reported verbal feedback had prepared them for the written feedback, and that it was easier to implement highlighted unmet standards as it was easier to seek permission or funding. The verbal feedback was identified by a few subjects as being able to provide weight to request resources and funding, but there was a much stronger sense of this from the written feedback stage. Subjects from low secure QNFMHS also found verbal feedback provided an opportunity for the peer review team to question why things were done in a certain way, enabling best practice to be shared between the host team and peer review team, providing further opportunities for change.

Some subjects partaking in the medium secure QNFMHS, reported that verbal feedback could be repetitive between cycles and in some cases, subjects reported recommendations not being useful, as they were around issues that could not really be resolved.

“I mean a lot of recommendations that come here some of them do repeat themselves. Some of them are to do with for example ligature points available ... and those haven't been changed, they are always flagged up.” [107SMMCP, LSU / MSU Combined Senior Manager]

Some subjects mentioned that they preferred verbal feedback to written feedback, as there was no room for misinterpretation. There were also two mismatched reports of verbal and written feedback.

Frontline responses were mixed, some subjects did not always receive verbal feedback, some felt proud of the external validation it offered to commend their hard work and efforts, and some who were present at the feedback sessions felt the discussions were solution focussed and this was helpful in bringing about change. One AIMS PICU senior manager felt that verbal feedback provided new ideas that enhanced and developed the service, which they deemed to be more valuable than the certificate itself.

### **10.3.6 Written feedback**

Written feedback was mentioned across all programmes, by all staff groups. Some subjects, more noticeably those partaking in QNFMHS, reported the most changes being attributed to written feedback.

“Probably during the visit and the biggest change would come when the report came through, yeah, cause they probably picked up on things that I may not have seen myself, yeah.” [17WMLCN, LSU Ward Manager]

“I think so; it gives an awareness of things like generally but I think like by having something written on paper sometimes it makes it more formal yeah? ...Because when it's written down on paper people are like “Oh right no that actually does need to happen, it's important like the other things that we do” do you know what I mean? So like some things people are impossible be like oh no that's not an issue, I can see people when they've read these things that need to be done, they're like “No actually it does need to be done, it's written down, it needs to be done, so this is important”” [97FLMCP, LSU / MSU Combined Frontline subject]

Action planning was the most frequently mentioned mechanism in this programme factor as an example of a structured process aiding change. Several subjects mentioned immediate changes following the written feedback, with some reports of prioritisation of changes.

However, discrepancies between written report and feedback were reported, with some subjects describing how the report did not always reflect the information fed back at the end of the visit. Subjects also voiced their disappointment of feedback or recommendations that they could not translate into changes. Frontline subjects reported that they did not always have access to written feedback. Even when they did, some reported reading the document to gain awareness but not being able to contribute to changes. One MSU frontline subject, highlighted that the vocabulary and wording of the feedback report could not always be easily understood by frontline staff.

### 10.3.7 Peer reviewing services

Per Table 10.3, 42 subjects had peer reviewed other services at the time of the interview, across the three CCQI programmes, with more reported peer reviewers from QNFMHS than from AIMS PICU.

**Table 10.3 Number of subjects who had peer reviewed other services**

	<b>AIMS</b>	<b>LSU</b>	<b>MSU</b>	<b>Combined LSU and MSU site</b>
<b>Frontline subject</b>	2	0	1	0
<b>Ward Manager</b>	1	6	3	4
<b>Senior Manager</b>	3	9	8	5
<b>Total</b>	6	15	12	9

Mostly reported by ward and senior managers from all programmes as the most valuable and useful programme factor, those involved described how it had helped them gain familiarity with concepts of external peer review.

“Yeah, I think being a reviewer, I understand the system better and that’s why I’d like more people to go but, obviously, there are limits...”

more people to be part of the review teams and get ideas and share do those networking elements.” [28SMMNW, MSU Senior Manager]

“That helped me because I did that before our last visit and I think that did help me... I suppose I've got an understanding of both sides of the fence really and I think, “Right, what are they looking at?” and, yeah we found things that needed improvement when we did our review which I had to put over. It's then seen how they put over their information to us as well, I suppose, yeah, it did help.”  
[13WMMN, MSU Ward Manager]

Subjects reported changes occurring through gaining new ideas, reviewing similar or related environments, such as the difference in public or private service provision, or a different level of security, as this was reported to be helpful through varying contexts. A few subjects even reported having picked things up from members of their peer review team.

“When our staff have gone to do peer reviews in other units, they have generally gone with the approach that not only are they going to look at other units but to try and pick up good practice and ideas for developing our own service. And on the two reviews, three reviews in fact that our service has been involved with, we've actually kept contact with members of that service afterwards, including a service user and gone back to them to ask for ideas or some practical things that we can develop here.” [79SMMNS, MSU Senior Manager]

### **10.3.8 Additional CCQI services**

The additional activities offered by the CCQI programmes, such as workshops, online discussion forums, annual forums, annual reports, meetings to review standards, carers events, training and newsletters were frequently cited as essential programme factors that brought about change. These programme factors were considered to add value and differentiate external peer review programmes offered by the CCQI from other quality improvement approaches. Some of these activities were described to sustain the networking element that took place over lunchtime sessions on peer review visits. A reasonable number of senior managers expressed that access to more additional CCQI services would enhance the programmes.

The online discussion forums were a free, useful and safe space for everyone to share ideas, and was the most frequently cited additional activity offered amongst subjects. Even those who did not necessarily contribute to the online discussions reported that just knowing they were there was a helpful resource. However, one subject expressed their concern for bad practice also being shared.

“I do and I don’t, because I think sometimes that... I think sometimes people just sort of take the attitude that there’s safety in numbers, and actually there’s safety in other people not doing things correctly, rather than... Well it’s like the... you know, I use the MSU, the discussion forum... Yeah. And you know, I haven’t contributed to it for ages, because it... I haven’t got the time really, but yeah, I often read other people’s postings, but... Somebody had put a question out there about, I don’t know, searching patients’ mail or something. I said, “Well, everybody knows that’s unlawful,” but people are sort of all joining and saying, “Well we do it this way. We do it that way.” And I just say well you’re all clubbing together, doing something that’s actually unlawful, but because other people are doing it...Yeah. You know, it’s the dubious practices like, you know...” [110WMMCP, LSU / MSU Combined Ward Manager]

But when other subjects were asked if they shared this view, it was not corroborated. Some MSU subjects expressed a need for a file-sharing platform. This was proposed as an add-on to the popular online discussion forum, but would allow searching of policies, documents and resources, which would benefit all and enable even more knowledge sharing.

Furthermore, some subjects made suggestions for the CCQI to link in with their local existing monthly forums or workshops, both relevant for those working in PICUs and forensic mental health, as they felt other services from outside their local area could benefit.

Newsletters were reported to offer benchmarks to work from. As contacts are offered it was seen to encourage fostering links between services and provided an opportunity for sharing of best practice.

“[The newsletter] is really, really useful. I think that seeing what other services are doing, you’ve got a contact email address, we use that a lot our security people, security manager “Oh I must speak to so and so about...” I like that. We’re going to be taking some things forward about developing technologies. I found that a particularly useful resource.” [14SMMN, MSU Senior Manager]

In a similar thread of benchmarking, one low secure senior manager found the annual report useful and reassuring to see where other services were also struggling:

“I know from looking at the end of cycle one, the annual report, for example, that lots of services are struggling round the same kind of areas really and sort of sharing solutions, that kind of thing.”  
[88SMLN, LSU Senior Manager]

Subject accounts suggested that most frontline staff were unaware or did not have access to these valuable additional CCQI services. Only one frontline subject partaking in the QNFMHS demonstrated awareness of the online forum. AIMS PICU senior and ward managers only mentioned training and the online forums.

Furthermore, subjects reported geographical barrier to accessing some of the additional CCQI services. As most of the workshops and activities are held in the South of the country, it was reported to be quite difficult (both time-consuming and costly) for some of the subjects from Northern services to attend.

Some subjects reported their involvement in additional CCQI services increased their familiarity with the process.

“[After being involved in a standards development workshop] I also knew when I was going to other sites what I was looking for and what I was aware of.” [99SMMCP, LSU / MSU Combined Senior Manager]

“The long-term benefits probably have been more so the workshops and things like that that follow on after the review process and not necessarily the actual review. It’s the other things that the Royal College offer.” [33SMLMCP, LSU / MSU Combined Senior Manager]



## **10.4 Mechanisms**

Table 10.4 presents the causal mechanisms that were divided into three main categories: mechanisms within programme factors, host team processes and staff groups.

**Table 10.4 All mechanisms described by subjects**

<b>Broad Mechanisms</b>	<b>Specific Theme</b>	<b>No. of subjects</b>
<b>Mechanisms within programme factors</b>	Sharing and learning	75
	Standards	35
	Evidencing	37
	Structured process	33
	Peer-aspect	25
	Reviewers	54
	Reflection	22
	Reassurance	18
	CCQI organisation	8
	Benchmarking	7
	Time	9
	Additional staff efforts	2
	Critical mass	1
<b>Team Processes</b>	Communication	41
	Consultation	27
	Teamwork	19
	Motivation	17
	Delegation	14
	Ownership	11
<b>Staff Groups</b>	Junior Involvement	45
	Senior Management	40
	Nurse Driven	2
	Clinician engagement	1

#### **10.4.1 Mechanisms within programme factors**

Mechanisms within programme factors referred to causal mechanisms that specifically occurred within the programme factors of external peer review. The sharing, learning and exchange of ideas was the most exceptionally reported and salient causal mechanism across all programmes and staff groups, as it was at the heart of many other themes. Sharing good practice is one of the ideals that underpins external peer review. Findings showed that CCQI external peer review programmes provide the opportunity for a learning environment whereby new members can pick ideas up from long-standing members, sharing resources and methods of practice with one another. Subjects frequently reported this theme as an integral part of the conversations and discussions during visits, indicating its importance as an essential causal mechanism.

Subjects from MSUs, combined services and AIMS PICU (mostly senior and ward managers) also demonstrated a great appreciation for CCQI standards acting as a causal mechanism, especially by those who confirmed a clear understanding of the programmes. Standards were used during self-review and visits. Subjects reported that as standards were evidence-based, which enabled services to benchmark themselves and offer reassurance of provision of correct care. Regular revision of the standards was seen to facilitate this mechanism, as it reflected the changing nature of the services themselves. Standards, especially those developed for specific services were reported to be important to subjects, as services could use them to provide evidence and rationale for decisions and prioritising resources. When some programmes began with lower, more achievable standards, these was perceived as highly beneficial for the subjects who mentioned it. Making standards more difficult to achieve once services had reached excellent scores was also seen to facilitate this mechanism, as subjects felt the bar for quality improvement was constantly being raised, making them aim higher. However, some standards were considered too high, and thus unachievable; which was reported to have negatively influenced levels of motivation in services. Furthermore, some managers on combined sites felt there ought to be clearer differences between LSU and MSU standards. This was

also occasionally expressed by LSU subjects that were visited by MSU members on their peer review team. Overall, a considerable number of subjects reported there were too many criteria for CCQI standards, giving an overview of the standards instead of going into depth, which they felt would be more helpful. One senior manager from an LSU MSU combined site felt one set of standards was needed as services currently face too many from the CQC, Department of Health and other regulatory bodies; whereas other subjects disagreed. Seven subjects called for more tailored standards to their services, and five subjects requested more focussed peer review visits, with one requesting visits specifically tailored to CQC outcomes.

Evidencing was a salient causal mechanism mentioned by subjects from all programmes and staff groups, but considerably more by subjects from AIMS PICU, especially ward managers. During the process of gathering evidence for self-review, services were prompted to review and check areas of compliance and perhaps make changes as necessary so they could then provide the necessary evidence.

“It was very much that our involvement was making sure that we had the evidence to actually support; making sure that we were doing, like, the CPAs were getting done and meeting all the standards for that. It was very much, rather than getting... doing the standing around what we were doing it for, it’s all very much, “Are we doing x, y and z? Yes, yes, yes. What aren’t we doing?”” [61FLLMCN, Combined LSU MSU Frontline subject]

However, in some cases, evidencing did not lead to any changes and it was just used as a mechanism to show what was being done already.

“No, no. I don’t think there was a lot of change as such around the unit. I think it was about evidencing what we did.” [74FLAN, PICU Frontline subject]

When tasks and activities were carried out in a structured process, this was a causal mechanism; mentioned across all programmes and staff groups, but mostly by those partaking in QNFMHS. Subjects reported using structured

processes for implementing change through reviewing standards, generating new ideas, disseminating feedback, and having specific meetings (such as appraisals and supervision) and forums. The use of quality improvement methodology such as driver diagrams was mentioned by an MSU ward manager. The theme of action plans was mentioned by 20 of the subjects across all programmes, but mostly by QNFMHS senior and ward managers. Action plans were a specific structured process, commonly used following feedback from a peer review visit to structure the implementation of recommendations. One subject outlined that although action plans were carried out by frontline staff, they were often developed by management. A senior manager partaking in the low secure QNFMHS specifically mentioned that action plans were only formulated around standards and not all recommendations; whereas a senior manager from a medium secure QNFMHS held a contrasting view that suggested all recommendations were incorporated into action plans.

Peer-aspect was mentioned across all programmes, and all staff groups; but most frequently amongst subjects partaking in medium secure QNFMHS, and slightly more by senior managers. Subjects reported 'peer endorsement', which they only received from peers working in similar services to themselves, as it was an essential causal mechanism, as it was felt that recognition and validation from peers held more weight and importance than regulatory bodies. The external viewpoint offered by peers coupled with direct external observation contributed to the importance of this mechanism, and was contrasted against other quality improvement approaches by subjects. Celebration of good practice was mentioned by subjects from all programmes, but most frequently by senior and ward managers. This was felt to be important to validate internal measures, and provide an opportunity for host services to show their practice to peers and receive their feedback. Subjects valued the access to expertise from peers that they might not have otherwise met. This was especially salient for newer member subjects, as they frequently cited the wealth of knowledge held by their longer-standing peers. Interviews uncovered an acknowledgement and deep appreciation for being able to access expertise from high scoring services. Linkages developed between professionals that promoted inter-organisational

and inter-professional collaboration through relationship building was highly remarkable. In some instances, subjects reported they had remained in contact with peers, and continued to share resources with them. Conversely, a small number of senior managers felt that the peer-aspect caused subjectivity, and it would be more efficient and standardised to have a small team or few small teams of people to conduct visits.

The theme of reviewers was mentioned across all programmes, and all staff groups, but considerably less from LSU subjects, and mostly by senior and ward managers. The composition of review teams, although variable, was perceived to be substantial. Different inputs from different professions, offering different multidisciplinary (MDT) perspectives was essential to bringing about change. However, issues such as competency and consistency between the peer reviewers was one of the most considerable issues influencing programmes, and frequently mentioned. Some subjects reported that having a member of the CCQI participate in the peer review team was also considered helpful to those who had experienced it before. This theme was only mentioned by four senior managers, mostly from member services to QNFMHS.

“They’ve always been really good and really proactive and really good at hosting and introducing people, and making the days run smoothly and pulling people together, so that’s always been really good. They’ve always been really knowledgeable and understand what they’re doing which I think helps the review team as well ‘cause in the early days, obviously, people hadn’t done many. You could tell people were quite anxious about coming and leading them and doing them.” [33SMLMCP, LSU / MSU Combined Senior Manager]

One senior manager mentioned CCQI attendance was not available in current AIMS PICU visits, and contrasted this against a low secure QNFMHS visit attended in a previous role, and how helpful they had found it. Service user involvement was also mentioned by subjects from MSUs, combined services and PICUs, but not at all by LSU subjects.

“I did a peer review with a very sort of proficient I suppose service user and I probably picked up more from him.” [81SMMNS, MSU Senior Manager]

Some ward managers and senior managers felt that service-user representatives should have a permanent role on peer review teams as they were felt to add high value to external peer review programmes. It was felt that more useful and realistic feedback was provided when a service user or carer representative participated in the peer review team.

“He was looking at what it would be like to be on that ward if he were a service user and things that I didn't even notice, like, he was commenting that the environment on one of the wards that we were looking round was quite stark. Whereas I looked at it and said, “Oh, this is nice and clean and tidy.” So, just a very lived experience. He was much more critical than we were, I must admit.” [60SMLMCN, LSU / MSU Combined Senior Manager]

The input of a lay person, service user, service user representative or carer was often reported as so essential, multiple subjects expressed their disappointment during visits where this view was not represented on the peer review team.

Subjects outlined how external peer review encouraged reflection of practice, and this causal mechanism was similarly represented from all staff groups and programmes.

“Just make people stop and reflect and that actually, you know, we're doing what we're here to do and you know, that external verification that you're doing it right as well helps.” [69FLLP, LSU Frontline subject]

This was reported to be vital, as being self-critical and self-reflective assisted services to start thinking about or making changes; especially as subjects reported it was easy to get caught up in day-to-day activities:

“Yes, I mean it was more of a reviewing and reflecting on our own practices as well. Because sometimes you don't think much about what you do on a daily basis. Until when you get time to sit down or

put it on your paper. For example, when they were looking at the security forms, how do we do security and then it also gives you that time for reflection to think you know is this a good practice, does it need to change we might have scored ourselves well but is it really good practice? So really it's also gave us that opportunity to reflect on our own practice." [122SMLCN, LSU / MSU Combined Senior Manager]

"I think from staff's point of view; I think it makes people reflect more. That's what I feel. I think, you know, we do a lot of things in set ways and we've maybe done numbers and numbers of years and then when we are having sort of peer reviews and things like this, it sometimes makes you question. It's a bit like having a student nurse on your ward really. You know, it makes you think, well actually, yeah, I've done that that way, but why have we always done it that way? Is there not another way we can do it and get the same outcome, you know?" [31SMMCP, LSU / MSU Combined Senior Manager]

Reassurance was reported by all staff groups mostly from QNFMHS. Subjects felt this mechanism provided an opportunity to check they were doing the right thing or along the right lines. Benchmarking, in contrast was only mentioned by AIMS PICU subjects from all staff groups. The opportunity to compare and benchmark against other services was especially mentioned by subjects who had experienced peer reviewing other services (see Table 10.3).

CCQI organisation was only mentioned by one senior manager from a combined service, and the remaining seven subjects were members of AIMS PICU, across all staff groups. Subjects detailed the valuable help received from the CCQI programme teams during preparation. In some specific cases, advice offered by the CCQI team were tailored to support the needs of the service, which was a causal mechanism.

Limited time was mentioned across all programmes and staff groups. Some subjects, especially from AIMS PICU, called for less time-consuming processes. However, other subjects felt like the visits had a rushed timeframe, and there was not enough time to cover all the topics they felt were necessary. Some senior managers from medium secure services belonging to peer review



networks suggested longer in-depth visits. Some frontline subjects echoed this and felt that it was important to have the opportunity to speak to more staff during the visit, and some subjects across all staff levels called for more frequent visits or follow-ups to promote the continuity of quality improvement. One senior manager from low secure advocated a smaller workbook so reviewers would have more time to speak to staff and service users, note observations and network, which was perceived as an essential causal mechanism.

Examples of additional staff efforts were only reported by one senior and one ward manager from AIMS PICU. This involved staff coming in to work on their days off, or taking on duties outside of their normal role, and was perceived as essential causal mechanism.

Critical mass implied that external peer review was more useful when more services participate, but was mentioned once by an MSU ward manager as a causal mechanism.

#### **10.4.2 Host team processes**

The most salient host team process mechanisms identified by subjects were consultation in QNFMHS (which was also linked to delegation, ownership and teamwork mostly in AIMS PICU) and communication. Motivation was also considered an important mechanism.

Consultation was mentioned by subjects across all programmes and staff groups, but more substantially by members of QNFMHS. This causal mechanism referred to consulting host team members. Subjects reported positive accounts of consultation leading to positive changes, such as when self-review standards were reviewed in consultation with the team, or action plans to implement changes were developed in consultation with the host team.

Delegation and ownership were related concepts, but were only identified from subjects undergoing AIMS PICU, from all staff groups. This occurred when tasks

were delegated to PICU staff giving them a feeling of ownership. Teamwork was mentioned by subjects from all staff groups and programmes (apart from low secure QNFMHS), but again most predominantly in AIMS PICU.

Communication was another salient host team mechanism, raised by subjects from all programmes and staff groups. Some frontline subjects and a ward manager belonging to QNFMHS expressed that feedback should be better communicated from visits. However, this mechanism was not mentioned by members of AIMS PICU. Motivation was mentioned by subjects from all programmes and staff groups, but it was expressed more substantially in accounts from AIMS PICU subjects.

#### **10.4.3 Staff groups**

Of all the staff groups, involvement of both junior and senior management were perceived as critical causal mechanisms. Nurse driven and clinician engagement were also seen as important mechanisms, but only slightly mentioned in comparison.

Junior involvement was mentioned by all staff groups and programmes, but most frequently amongst MSU subjects. Many QNFMHS subjects across all staff levels indicated more junior involvement was necessary, especially as healthcare assistants (HCAs) spend most time with service users. Quite a few frontline QNFMHS subjects expressed interest in becoming reviewers themselves, however, this was not raised among frontline subjects from AIMS PICU.

Senior management support was mentioned by subjects from all programmes, but mostly amongst AIMS PICU senior and ward managers. Management support was also reported to be a key indicator of capacity to benefit. Without management support, it was perceived very little change could be made, as support was instrumental in converting recommendations from the written report into changes. In addition, management support was felt to be instrumental to achieving accreditation. An AIMS PICU ward manager described

that it was in the best interests of senior management to facilitate changes necessary for external peer review as they had already invested in the programme.

A few senior managers participating in AIMS PICU described that the programme was very much nurse-driven, and that clinician engagement was often difficult at first. One senior manager reported how a consultant psychiatrist completely disregarded the programme at first, until leadership was shown by a younger consultant.

“Because of the... our consultant was slightly older and I think that he was dubious about things at first and then when he saw this young consultant who was very enthusiastic, he just... it just seemed to... you know, he was full of it, you know, after, but at first... Yes, he just suddenly thought, you know and I think that if he'd had continued with this attitude about, “Oh, I'm playing golf,” I would have said, “Well it's unfortunate, cause this is the Royal College, you know and...” but it... I think it's just a little bit of, you know, bravado at first, which was a shame, but he's fully embraced it and so have all of the other, you know, consultants and members of the MDT.”  
[102SMAN, PICU Senior Manager]

## 10.5 Changes

Table 10.5 presents the types of changes that were mentioned by subjects: long-term, short-term, negative, no changes and perceptions of changes. I have characterised long-term changes as a commitment towards institutionalised quality improvement beyond the impending peer review visit, and short-term changes as compliance with standards for an immediate peer review visit, per subject perceptions.

**Table 10.5 All changes described by subjects**

<b>Broad Changes</b>	<b>Specific Theme</b>	<b>No. of subjects</b>
<b>Long-term changes</b>	Examples of long-term changes	23
	Driver for change	22
	Continued quality improvement	37
<b>Short-term changes</b>	Short-term	23
	Rankings	13
	Achievement	12
	Advertisement	14
<b>Negative changes</b>	Deferred	6
	Failed	1
<b>No changes</b>	No change	18
	Do not get to see or hear about changes	12
	Change is not down to external peer review	4
	Does not work	3
<b>Perceptions</b>	Improvement suggestions	54
	Helpful or useful	41
	Want to be more involved	20

### **10.5.1 Long-term changes**

Long-term changes were categorised by examples of long-term changes implemented because of participating in external peer review, the programme acting as a driver for further change, and continued or sustained quality improvement. The majority of these themes were reported by services which joined their respective external peer review programmes in 2009.

Examples of long-term changes were mentioned by all staff groups from all programmes, but predominantly by MSU subjects. These were mostly focussed on, improvements for staff, security and a few improvements for service users. Examples included changes such as sustained improvements in staff atmosphere and 'happiness', sustained improvements in staff supervision, employing staff specifically to oversee security, increased information sharing within the host service, improved security training, culture of policy development and review, sustained improvement in the provision of service user activities and sustained improved work placements and employment opportunities for patients.

A change that was frequently mentioned, was the way in which the programmes promoted continued quality improvement thinking and acted as a driver for further change. This change was mentioned by all staff groups from all programmes except for the low secure QNFMHS, and most predominantly in AIMS PICU. Subjects reported using the momentum gained from external peer review to improve other areas of quality outside the standards.

Continued quality improvement after peer review cycles was mentioned by all staff groups from all programmes, but predominantly by AIMS PICU. Continuing from the momentum of peer review visits, subjects expressed the benefits they had found with continuing the in-house audit, action planning events and review activities in the interim periods in-between peer review visits. The momentum of external peer review also prompted the development of regional peer review networks, such as in the North Yorkshire & Humber Region of the UK reported carrying out similar peer reviews visits which were cited as being

helpful and supportive local activities. In addition, some services who remained in contact after meeting through peer review visits reported carrying out their own reciprocal peer review visits.

### **10.5.2 Short-term changes**

Examples of short-term changes were provided by subjects from all programmes, but mostly AIMS PICU, and more so from senior and ward managers. Advertisement was mostly mentioned by AIMS PICU senior and ward managers. In some accounts, advertisement was considered as a short-term outcome, as once the certification had been gained or the standard was met for the purposes of advertisement, there was no continual drive for further change or improvement. Rankings published in annual reports were mostly mentioned by QNFMHS senior managers as a way they advertised their services. Although the rankings were anonymised, subjects used their written feedback scores to calculate their rankings, and used this information to benchmark and advertise their services.

“At the end of the day I have to think what have I spent my money on and if I get an outcome, I think we were 7th out of 96 last year, that’s something I can sell in my sales meeting, that’s something I can use that’s worth three grand. Now if they’re going back and just saying, “Oh no the others who got below you didn’t like the fact there was rankings,” Oh well tough. They should have done better. It’s a competitive world and so if I can get a gold, silver, bronze or whatever that I can put on my website that I can show my commissioners, I’m therefore getting some value for money and benefit from CCQI.” [48SMLP, LSU Senior Manager]

### **10.5.3 Negative changes**

Failing accreditation was only mentioned by one AIMS PICU senior manager, and being deferred was only mentioned by AIMS PICU subjects from all staff groups. Lack of training, entrenched bad practice which did not meet standards, limited senior management involvement and unsuitable environments were reported as reasons for these negative changes. However, according to one ward manager, despite being deferred, evidence of short-term changes was required to achieve accreditation instead of long-term changes.

“Well we were deferred first time round actually because of a training issue... We didn’t complete it; we didn’t understand it... It was to do with CPA training... There wasn’t any CPA training available by the Trust, so I was not responsible for that and I should have realised and I all think we felt that bit. On another plus side is it’s sometimes a fairly easy win. Right, we’re just going to have to develop something. So let’s get it done and get everybody trained and that’s exactly what we did... Well, I’m still pushing to get CPA training all on the map and to put it frankly, I think they did the CPA training to get us through the AIMS but of course it doesn’t end there, does it? We’re peer reviewed and we have it in another probably, I don’t know, another eighteen months come October so we can’t just do that again. We can’t just oh let’s quickly through something together, so it is still on the table. I’d like to see it move a bit faster if the truth be told. In general running of the place it doesn’t make much difference like I said when nurses know what CPA is, you know, and they did receive the training anyway, but we need to make sure it’s ongoing. I’m honestly not satisfied that that’s quite happening as yet, but I think it will.” [106WMAN, PICU Ward Manager]

#### **10.5.4 No change**

No changes as a result of participation in external peer review was frequently mentioned by all staff groups from all programmes, but most frequently by LSU frontline subjects. Not getting to see or hear any changes was also mostly mentioned by frontline subjects participating in QNFMHS. Changes not attributed to external peer review was mentioned by all staff groups from all programmes, apart from low secure QNFMHS. In these reports, change was reported to have occurred anyway, driven by CQUIN targets or because of the organisational focus.

Senior and ward managers: three partaking in the low secure QNFMHS and one from a combined service outlined that they did not think external peer review worked.

“I don’t think it’s done a huge amount of good, but I don’t think it’s done a lot of damage either.” [110WMMCP, Combined LSU / MSU Ward Manager]

### **10.5.5 Perceptions**

Just under half of the subjects suggested improvements for their programmes, and these were mostly provided by senior managers. The most salient improvement suggestions were around more junior involvement needed, especially for them to attend peer review visits of other services, assistance with communication within host services and help with the continuity of quality improvement. The remaining suggestions have mostly been referenced under their respective themes, and were broadly under the following areas: consistency of peer reviewers, regularity or focus of review visits, transparency, tailored or improved standards, time, paperwork burden, training, lead reviewer drop-out rate, more networking opportunities requested, less geographical barriers to CCQI additional services, more and faster ways of sharing of best practice, unannounced elements, follow up on recommendations, more rigorous validation, consistency between verbal and written feedback, staff support during peer review visits, formal feedback dissemination, more service user involvement, and something in place for unachievable standards. Given the preference of peer review visits to regulatory inspections conducted by the CQC, some suggestions for unannounced elements to external peer review contrasted with the majority.

“The people doing it would have to sign up so if they turn up do it or might give them the choice, say, “Right, you can have an announced visit but you’re only going to get a certain ranking for that. If you agree to have an unannounced visit, you’ll get a higher ranking.”  
[48SMLP, LSU Senior Manager]

41 subjects reported the programmes were helpful or useful, with the majority being senior and ward managers. 20 subjects, mostly ward managers and frontline subjects expressed an interest to be more involved in the external peer review process.



## 10.6 Contexts

I have conceptualised the outer and inner contexts which facilitate or impede the ability of services to respond to the interventions delivered through CCQI external peer review programmes. I had accounted for some of these contexts in the purposive sampling frame: type of external peer review programme, healthcare system, mandatory, length of membership and geographical location. Therefore, the number of subjects for some of the themes in Table 10.6 are not applicable as they were generated using data from all participants.

**Table 10.6 All contexts described by subjects**

<b>Broad Contexts</b>	<b>Specific Theme</b>	<b>No. of subjects</b>
<b>Outer</b>	Type of external peer review programme	n/a
	Healthcare system	n/a
	Other quality improvement approaches	68
	Mandatory	n/a
	Finance	25
<b>Inner</b>	Length of membership	n/a
	Joining decision	38
	Geographical location	n/a
	Non-financial resources	39
	Local context	36
	Size	12
	Organisational structure	15
	Organisational culture	6

### 10.6.1 Outer contexts

The most salient outer contexts were the type of external peer review program and the influence of the healthcare system. Other quality improvement approaches, whether programmes were mandatory, and financial impacts were also considered to be important.

The main differences subjects mentioned between peer review networks and accreditation schemes was the obligation to make changes because of recommendations received through feedback stages. There was no penalty for non-compliance in peer review networks, contrasted with the obligation to comply with standards to achieve accreditation. Only one frontline subject could comment on the differences between the two, and they felt that if the peer review network was changed to an accreditation scheme or had a punitive element to enforce change it would detract from the sharing and learning aspect, which was felt to be essential by many subjects.

“It would be more of like an authoritative kind of thing, you know. I don’t think that would be the best idea... Because it’s meant to be more about learning, improving services. The CQC is inspector, you know. CQC, cause have you done this? Have you done that and followed the Government regulation? Yes. This peer review is all in the line of improving quality, which his still a government... yeah, standards, you know, but it’s more of information sharing amongst other establishments, you know.” [92FLLP, LSU Frontline subject]

A few subjects across both external peer review programmes recognised that staff efforts were validated through the certification awarded during accreditation, and this mechanism was not as substantial in peer review networks. However, many subjects, namely QNFMHS ward managers felt that the certificate made little difference to the outcome of quality improvement.

“Is it not just the same thing, just with a different title? Are we looking for accolades here in certificates or are we just looking at the quality of service that we deliver?” [67WMLP, LSU Ward Manager]

Some subjects who were members of AIMS PICU believed they were less likely to experience CQC inspections after having received an accreditation award.

Healthcare system context such as whether the organisations were provided through the public or private organisations influenced the perceived ability of services to respond to interventions delivered through external peer review. Several senior managers from private services commented on the 'increased scrutiny' from regulatory bodies faced by private organisations, and some mentioned the political context with regards to private services 'getting the leftover' or 'unwanted' cases from NHS services, or having to work to 'impress' NHS services.

“...in fact, if NHS is coming in here you want to impress them because they're potentially your referrers... You're always going to think about the business side. When you work in an independent healthcare provider you're always going to be thinking about business, it's drummed into your head that you think about business.” [10SMLP, LSU Senior Manager]

One senior manager working in a private PICU stated his preference for visiting other private PICUs; as he knew they would have similar, unique cases. Some senior managers had experience in both private and public services, and could contrast how this context influenced the ability to implement changes. They suggested that it was generally easier to make changes resulting in quality improvement in private services due to reduced bureaucracy. 'Red tape' was a frequently cited phrase by subjects in relation to making changes in public services. Furthermore, there was a perception amongst subjects that private services were better resourced than public services. This became evident when subjects were asked about their experiences of reviewing other services, staff from public services reported higher levels of resources in the private sector services they had reviewed. A few subjects from both AIMS PICU and QNFMHS mentioned competition between different private sector organisations; if they were being reviewed by a different private sector organisation this negatively impacted their experience and introduced a competitive element into external peer review. This theme was not shared by all subjects, but strongly felt by a

few. Competition was also raised in relation to peer review visits that occurred in close proximity to services. A few subjects mentioned that this made the visits feel more competitive than supportive. These views were mostly shared by ward managers and senior managers, and only a few frontline subjects.

The context of other quality improvement approaches was mentioned by subjects from all programmes. An MSU senior manager expressed an interest towards a peer review network exclusive to clinical directors, medical directors and associate directors, as it was thought to be beneficial in terms of strategic planning and developments for quality improvement. Most frontline subjects could not entirely distinguish between their respective external peer review programme and other quality improvement approaches their ward was involved in. Many subjects made comparisons and references to the CQC when providing their accounts and explanations. When compared to CQC visits the peer-aspect of reviewers were highlighted as being an instrumental mechanism to the programmes. CQC visits were also felt to be less supportive and more inspectorial by most subjects when compared to external peer review. However, some senior managers felt that CQC visits 'counted' for a lot more, especially given their regulatory powers, and thus placed more importance on them. Some PICU subjects had experienced visits through the National Association of Psychiatric Intensive Care & Low Secure Units (NAPICU). When comparing the visits, a senior manager outlined the external peer review visits were looking at different things:

"No they look at different things really. The NAPICU they came in to try and boost the morale and "oh the care plan for this patient is excellent" which is what AIMS put was you know they said that the actual care bits weren't that bad but they were giving us "you're doing really well; you're doing really well" trying to boost the staff. We work in this, we work with no staff and yet the experts from PICU care throughout the whole of the UK are saying well done. But they were looking at different things." [111SMANW, PICU Senior Manager]

Many senior managers stated that they were not involved in any other quality improvement programmes, thus made no comparisons.

Voluntary or compulsory participatory requirements were perceived as an essential context. Some subjects from accredited services referred to the legal implications involved, and even spoke of their fear of being closed if they failed to meet agreed standards of care. Publication of accreditation status was viewed by some of the subjects from AIMS PICU as a method of demonstrating a similar sense of accountability without the mandatory label. A few senior and ward managers belonging to the medium secure QNFMHS indicated that they would no longer participate if the mandatory requirement ceased, and this was echoed by other medium secure subjects who experienced less motivation to succeed due to the compulsory nature of the network. Although AIMS PICU was not compulsory scheme, there was a pressure expressed by some subjects to maintain their compliance with standards, especially if they had been accredited with excellence; however, this pressure was only reported in very few QNFMHS highly scoring services.

Finance was mentioned by subjects from all programmes. Subjects described how budget restrictions and financial restraints have hindered services involved in external peer review across the UK, and this context was often intertwined with many other themes. It was widely accepted throughout the interviews, that costs are an integral part of the programmes; from joining fees, to costs of travelling to peer reviews (which are absorbed by services), to the cost of implementing changes necessary to meet quality improvement standards recommended by the CCQI.

### **10.6.2 Inner contexts**

The most salient inner context was the length of membership. Joining decision, geographical location, non-financial resources, local context, size, structure and culture were also considered to be important.

Length of membership was the most salient inner context. Subjects provided many examples of behaviour between new and long-standing members, which were echoed by several accounts. Some long-standing medium secure services that had joined the peer review network quite early on felt the programmes

were less beneficial for them in recent years. Both senior managers and ward managers suggested less frequent peer review visits such as two-yearly could be more effective. This resonated with other MSU ward managers who expressed programmes should be 'less predictive'. Conversely, another MSU senior manager suggested that six month follow-ups would help to see if or how recommendations had been implemented. Alternatively, peer review teams could start with last visit's recommendations to aid continuity from one cycle to the next. Having more experience in external peer review programmes served as a facilitator of success for some subjects. Especially as some reported the more reviews they had led; the more improved their own host visits were. Some subjects felt increased familiarity with standards enabled them to gain a deeper understanding of their meanings. Repeated exposure benefitted these subjects as they became more comfortable with the programmes, which in turn helped them gain long-term success. There was a large consensus among subjects that long-standing members were less engaged, especially in mandatory programmes.

Joining decision awareness was mentioned by subjects, most whom were from AIMS PICU senior managers. NAPICU conferences were mostly referenced in relation to joining. Several subjects mentioned first hearing about the AIMS PICU programme and success stories of accredited PICUs at these conferences, and some even described an inadvertent peer pressure to join, despite the voluntary nature of the programme. Frontline subjects in general demonstrated very little awareness of joining decisions across both external peer review programmes. Subjects from low secure services provided a range of reasons for joining to having witnessed benefits from MSU membership they were linked to, pressure from commissioners, attending meetings, and one service reported an assumption that it would soon be compulsory for LSUs to join QNFMHS (although this is not presently the case).

Subjects reported that due to the geographical locations of some of the rural services, participating in an external peer review programme has greatly benefited them and alleviated feelings of isolation.

I will say that all the change has been positive. I can't see anything negative. We're... we're an isolated unit here really. I mean it's a lovely spot, but we... you know, we're... yeah, we're kind of isolated and it helps us to help people coming in and out, whether it's Quality Network, whether it's other organisations, just to, you know, share good practices. And the fact that we go elsewhere to look at medium secure units helps develop good practice. [116WMMNW, MSU Ward Manager]

However, geographical barriers also made it difficult to plan review visits and put reviewers forward for far away services given the high cost and difficult logistics of travelling to services that were large distances. Numerous senior managers who were interviewed in remote rural locations referred to this barrier, especially as they felt it hindered their ability to attend training sessions, workshops and networking opportunities that were based in London. Having no video-conferencing facilities for these sessions was an additional barrier mentioned by some senior managers.

Examples of non-financial resources were provided by subjects from all programmes and staff groups, the majority were senior managers. This was often mentioned as an important context to implementing change, and included resources such as the environment, structures (especially old buildings and having to retrofit medium secure sites to fit low security), facilities (i.e. IT, internet), staffing, and training. In addition, some who joined later also mentioned that delayed membership was due to a lack of non-financial resources. Structural resources were the most frequently mentioned as difficult to change, such as old buildings and positioning of wards. Some subjects expressed frustration at failure to comply with these standards, as they felt they had had little control over them.

Local context was mentioned by subjects from all programmes, by senior managers predominantly, but considerably less by those partaking in the low secure QNFMHS. Local context was reported as a substantial inner context and barrier by subjects from combined services who struggled with having LSUs inside medium secure sites. Most subjects who were members of AIMS PICU

mentioned the highly varying context of PICU. The nature of the service user questionnaires was raised by a few subjects and how by the time of the peer review visit it would be most likely that the service users who completed the questionnaires would no longer be present on the ward. They feared this would misinform the programmes.

Size was mentioned by subjects, mostly those partaking in the low secure and medium secure quality networks. Although this was only mentioned by a few subjects, there was a strong sense that changes were easier to implement in smaller services or smaller teams. However, two subjects also mentioned that their smaller teams struggled when staff members were sent on peer review visits. Organisational culture to promote and institutionalise quality as a successful context for change was only mentioned by some subjects, but across all programmes by all staff groups.

Well, I said before, the key thing is to embed the culture of quality, so when I spend that money so that people can just... it's a successful PICU... I think the point is to create a new culture where things are routinely quality driven ... If you have good staff, good well motivated staff, good knowledgeable skilled staff they would improve your quality. They would drive quality and good outcomes and if you have quality and good outcomes you will sell, so they work together. Quality is something that staff people have to deliver. So, if you don't have the right people you struggle with maintaining quality. If you don't have good quality, you struggle with selling.”  
[37SMAP, PICU Senior Manager]



## 10.7 Organisational readiness to change

To report the organisational readiness for change constructs, themes were split into five categories and presented in Table 10.7: constructs from the readiness for change tool ([Bobiak et al., 2009](#)), indicators of capacity to benefit from joining external peer review, barriers to change, awareness and understanding of external peer review.

**Table 10.7 All organisational readiness to change constructs described by subjects**

<b>Broad Organisational Readiness for Change Constructs</b>	<b>Specific Theme</b>	<b>No. of subjects</b>
<b>Awareness</b>		94
<b>Understanding</b>		46
<b>Readiness for change tool</b>	Connected with people in other practices	66
	Leadership vision	21
	Change is possible	16
	Effort to meet service users' expectations	16
	Taking time to think	4
	Other constructs	11
<b>Indicators of capacity to benefit</b>		60
<b>Barriers to change</b>		39

### 10.7.1 Awareness

Awareness of a subjects' respective external peer review programme was mentioned across all programmes and staff groups. This was the strongest theme relating to organisational readiness for change that arose throughout the interviews, with many subjects indicating that awareness was lacking in most frontline staff, and generally decreased in line with lessening seniority. Subjects

indicated if they were not present during the peer review visit, unless they were inquisitive, they would not otherwise be updated about the visit, related feedback or resultant action plans. When comparing awareness amongst subjects, frontline subjects participating in AIMS PICU appeared to be more aware than those belonging to QNFMHS, although this was still demonstrated in a few cases. Increased awareness through posters were mentioned by nine subjects, from all programmes and staff groups.

### **10.7.2 Understanding**

This theme was mentioned by subjects, from all programmes and staff groups but more frequently by those participating in QNFMHS than AIMS PICU. Understanding of the programmes varied between organisations and professional groups. Many frontline subjects lacked a basic understanding of the programmes, despite having heard the phrase 'peer review' before. Often frontline subjects participating in QNFMHS demonstrated a more limited understanding of external peer review than those participating in AIMS PICU. Unless specifically chosen to be part of staff discussion groups during visits or involved in the preparation during self-review, it was reported that despite an interest or willingness to participate, there was a strong lack of understanding. This was especially reported with regards to how external peer review impacted subjects directly, or what the benefits of these programmes were to them or their job roles. Amongst those frontline subjects who demonstrated a good understanding of the programmes, a sense of teamwork, empowerment and ownership of the programmes was reported. As one might expect, understanding was strongly related to involvement, increased involvement demonstrated by subjects, was linked to more understanding conveyed of external peer review and its benefits.

### **10.7.3 Readiness for change tool items**

Subjects from all staff groups and programmes described the construct of connecting with people in other practices as the most salient to organisational readiness to change. This theme was predominantly reported by ward and

senior managers, and was only reported by six frontline subjects. It was mentioned in relation to reviewing other practices, connecting with local services or attending additional CCQI activities (a key programme factor). Being connected with people in other practices was reported to have helped with organisational readiness for change, as the connections with others helped with sharing ideas and tackling similar issues, especially when the other practices were similar.

Leadership vision was expressed as the one of the most salient constructs of readiness for change mostly for AIMS PICU members, from all staff groups. In cases of successful accreditation experiences, leadership vision was described as instrumental for organisational readiness for change by being 'enthusiastic', a 'steering force', 'energetic' and 'motivational'.

Some subjects from all staff groups reported that change was possible in all programmes, but most reports came from services participating in the low secure QNFMHS. It was especially mentioned that the possibility of low cost changes was higher, as they were easier to implement. Changes that were of clinical priority or necessary for service user safety were perceived to be possible by subjects.

Some subjects from all programmes, but especially ward managers and frontline subjects reported examples of making efforts to meet service users' expectations.

A lot of compassion, yeah. We have a lot of compassion to the patients on the ward. It's basically just listening. Sometimes people don't always want to be told, "You're under the mental health act; you're on the ward for such and such;" they don't want to be told that. You've got to, like I said, with every patient, it's a very different approach. You have to watch, not watch, but you've got to watch what you're saying and address them correctly. It's just good manners at the end of the day and the staff are very respectful with that on here. [105FLAN, PICU Frontline subject]

Subjects felt taking time to think was an essential readiness for change construct, and many often found external peer reviews were the only opportunity they had for reflection through self-review and visit programme factors. This was also detailed by the mechanism of reflection (Table 10.4).

The remaining readiness for change constructs were less remarkable in the data. The theme of being flexible with change was mentioned by subjects from all programmes and staff groups, but most predominantly by senior managers. The theme of trying new things was only mentioned by subjects from QNFMHS, most predominantly by senior managers. The theme of being supported by people was mostly mentioned by frontline subjects from all programmes, and this theme was mentioned mostly in conjunction with the theme that leadership promotes an enjoyable environment. The theme of shared vision was mostly mentioned by AIMS PICU subjects. The theme of having similar opinions came up by all staff groups from programmes. This was mostly expressed through describing not agreeing with the opinions of the majority and the subsequent fears of reprimand or facing the possibility of facing peer humiliation. In general, subjects reported that it was fairly easy to propose changes, and whether they were always implemented was dependent on outer contexts.

The following constructs were only mentioned by less than two subjects: staff being friendly to service users, using community resources, adjust to routine in response to changes. The changes so fast and changes shot down theme were only mentioned by subjects undergoing AIMS PICU. It was not directly mentioned amongst any subjects that changes could not be implemented due to staff being so busy seeing service users. There was no mention of the healthcare system being helpful, stays on schedule, adapt to new standards even those forced, community resources helpful, incompatible requests to non-clinical staff, staff conflict or staff-clinician conflict.

#### **10.7.4 Indicators of capacity to benefit**

Indicators of capacity to benefit from participation were observed during visits such as preparation by the host service, engagement, openness, service users,

and general observations were reported. Subjects often described a 'feeling' they experienced, which was indicative of the host services' capacity to benefit from participating in external peer review.

The level of preparation undertaken was an indicator of a service or ward's capacity to benefit according to subjects in both accreditation programmes and peer review networks. Subjects felt that noticeable preparation indicated the host service was engaged with the process, took it more seriously; and thus, would be able to benefit more from participation.

The level of engagement, especially senior management engagement during the site visit was reported as an indicator of a service or ward's capacity to benefit from external peer review. Subjects perceived senior management presence or engagement demonstrated the likelihood that changes would follow the visit.

Subjects described openness in terms of transparency, honesty and freedom to walk around and speak to staff or service users during peer reviewing other services. However, subjects agreed that even when host services demonstrated transparency and openness, it did not necessarily mean they had met all standards, but it indicated their readiness to benefit from the external peer review programmes as they demonstrated a willingness to learn or change.

Observing service users' behaviour or feedback, for example their use of facilities and provisions, was viewed as an indicator of a services capacity to benefit from by some subjects.

"I suppose the fact that I've just flagged up that the service users didn't want talk to us could be an indicator of... you could take that both ways. I did at the time spin it on its head and say, "actually service users are more, most people in general not just service users, are more likely to want to talk to 'somebody' if they'd want to complain." You don't very often get people say, "Oh, I had a great holiday, I'm going to do a review, I had a great experience on it," it's usually people complain so you could say that, "actually maybe they had nothing to complain about so didn't see any purpose in talking to us." I suppose that was a factor but I suppose as a peer review team on the day, we were a bit, "not one service user wants to come talk to

us.” We did think that was quite unusual. Whether that’s an indicator... positive about the service, I’m not really sure.” [46SMLP, LSU Senior Manager]

“Yeah, it could just be their bad day or if they’re unwell at the time but they tend to be fairly honest, the patients, when they’re asked about things. I think you tend to get a good idea of what a service is like when you speak to the patients ‘cause they’ll tell you if the carer’s good or if they carer’s not good. They’re not going to hide anything away from you.” [99SMMCP, LSU and MSU Combined Service Senior Manager]

In some circumstances, subjects gave accounts of their general observations which indicated to them how a service might benefit from external peer review. These were largely based on their professional experience.

“Environment, I suppose I’m sensitive to the environment of a hospital because I very much try and stand in the shoes of a service user. I think, “How would I feel if I was detained in this hospital against my will? I didn’t want to be here,” and part of our group is to make a hospital stay as endurable as it can be for people that don’t even want to be here in the first place. So, I think environment is really important and one of the hospital’s I worked in had the most amazing environment, it was a beautiful, really big, open spacious area to be in. They had lots of facilities and it was very colourful and very bright and very lively. You could see that you could be quite enthused about living in that particular hospital and I suppose the one I done recently was a slightly more... it wasn’t a negative environment but it didn’t have that bright, buzzy feeling about it where you can, you know...” [46SMLP, LSU Senior Manager]

#### **10.7.5 Barriers to change**

Reluctance to change was mentioned by subjects, across all programmes and staff groups, but considerably more in the medium secure QNFMHS and AIMS PICU, which were longer standing programmes. The phrases ‘young staff’, ‘new staff’ and ‘new management’ were especially raised in the context of discussing reluctance to change. Senior management often used the terms ‘young’ or ‘new’ staff with descriptions of being ready for change, which suggested long-term staff demonstrated more resistance than newer staff members who were more amenable to change. The most common example reported, was when staff or

management were stuck or set in a specific mind-set that was difficult to change. This was often cited in relation to long-term staff:

“I think sometimes more experienced staff or possibly staff that have been qualified for longer were a bit reluctant to change.”  
[107SMMCP, LSU / MSU Combined Senior Manager]

In cases of ‘new management’, this was also cited frequently in conjunction with being ready for or flexible with change by other frontline staff.

Reassuring staff was a common theme that was described in conjunction with resistance to change. Mostly ward managers described how with explanation, meetings and perhaps direct supervision, resistance could be overcome:

“It’s about addressing it and sort of sitting with people and saying “Well you know, trying to change their frame of mind because yes, we are doing this, why wouldn’t you want to be recognised for something that you are actually doing?” And you find that with a lot of... with a few older members of the team, you know you have people who have been in post for quite a few years and they say “Oh we’ve seen it all before” it’s about changing why we’re getting an award. But on the other hand when you have got the followers already in there and I don’t think it’s hard work because they were going to join after they’ve seen everybody’s doing something, and they’re the only ones... so it’s about having conversations with people directly and once you’ve got that set of followers already moving along, they don’t want to be left behind, they want to join in.”  
[52WMAN, PICU Ward Manager]

Being apprehensive before a visit was frequently mentioned, mostly by ward and senior managers as a common experience before the site visit had taken place. Although this was reported across all programmes, it was mentioned considerably less from those partaking in the low secure QNFMHS.

Only eight subjects mentioned it to be explicitly difficult to make changes, across all programmes and staff groups. In some instances, this theme referred to general changes within their ward or service being difficult, thus acting as barrier for proposing new or additional changes.

“...there’s an expectation at times that things won’t change and therefore initiating change can be challenging because people will automatically think, “Oh, we won’t have enough staff to do this” or “There won’t be any money to do this.” So, it’s getting people over the initial inertia and when to make changes.” [60SMLMCN]



## 10.8 Conclusion

I could observe similar themes from each staff group. Frontline subjects approached interview themes with a more inquisitive nature, asking how they could increase their understanding or become more involved in external peer review; mentioning issues that impacted them directly such as receiving information about the programmes, and the dissemination of feedback. Ward managers gave accounts of their feelings towards the programmes, whether they personally found it easy or challenging, their anxieties and their 'journey'. There was a spectrum of involvement between the senior managers interviewed, some were quite removed from the programmes and others were much more involved. For those more involved, benchmarking and comparisons with other services, for both peer review networks and accreditation were perceived to be important. Nearly all the senior managers and some ward managers who were interviewed, mentioned the contexts of finance and non-financial resources quite frequently as a barrier that impeded services' ability to respond to recommendations delivered through external peer review. Peer review network senior managers were especially concerned with how to obtain additional value from the programmes. Some subjects, namely senior managers questioned whether the programmes offered value for money given all the costs involved, whereas other subjects acknowledged that the added value from networking, sharing resources and quality improvements far outweighed the cost.

Most changes for subjects participating in QNFMHS were reported as a result of written feedback; whereas subjects participating in AIMS PICU reported most changes occurring before the visit took place, either during self-review, upon joining, or for some services before joining. For those who had peer reviewed other services, the majority found this programme factor more valuable than reviews of their own ward or service. Peer reviewing other services was not seen as substantial to subjects participating in AIMS PICU as it was for QNFMHS subjects. Networking opportunities such as during peer review visits of their own services, peer reviewing other services and the additional activities offered

by the CCQI such as online forums, learning workshops and annual meetings helped people to feel connected with staff in other practices which was the most salient readiness for change construct. However, the accounts highlighted that many of these networking opportunities and additional activities could only be accessed by senior management, therefore it limited their benefits. Most of the views about additional CCQI services were from subjects belonging to MSUs, this could be because the medium secure QNFMHS is more longstanding, and it was suggested by subjects that it takes an increased length of membership to benefit or receive value from these additional services.

Sharing and learning was considered the most salient mechanism within a programme factor in external peer review, which was seen to underpin the process, and more opportunities for this were requested from subjects. Consultation was the most salient host team process for QNFMHS, which was linked to teamwork, delegation and ownership for AIMS PICU subjects. Junior involvement was also one of the most frequently mentioned causal mechanisms, and there was a recognition amongst subjects that not enough junior staff are currently involved in external peer review. This was considered a salient causal mechanism as junior staff were seen to spend the most time with service users. Senior management support was a substantial causal mechanism. This was linked to leadership vision as a salient readiness for change construct, and management presence was considered an important indicator of capacity to benefit. Senior managers mentioned the importance of senior management support the most, ward managers reported this theme too, but it was hardly mentioned by frontline subjects. Standards were expected to be an integral part of the programmes, and although they were recognised as being necessary, they were not identified as the most substantial.

Subjects from long-standing services provided examples of different changes than newer services. Certain types of changes such as continuity of change, and driver for change were only mentioned by the long-standing services. Most accreditation subjects said that process resulted in advertisement. Those partaking in low secure QNFMHS witnessed no change. This could have been

because they were all newly joined services, adding strength to the importance of length of membership as a salient inner context. As expected, the type of external peer review and healthcare system were important outer contexts. Implementing change and giving feedback was very different for accreditation than for peer review networks. As there appeared to be different pressures involved. Feedback in accreditation was tied in with a pass or fail judgement. The feedback mechanisms offered by peer review networks were often considered causal mechanisms for negotiation, providing services with power to negotiate for change and resources to facilitate further quality improvement. In this way, subjects demonstrated that the feedback mechanisms could support commissioning and resourcing. Feedback was also linked to the causal mechanism of communication. This was picked up by frontline subjects that 'did not hear or see changes' and reported that feedback from visits was not disseminated widely or specifically reached. Subjects participating in AIMS PICU did not mention this theme. This could be because if an accreditation is passed or failed, it is widely publicised throughout the service.

Differences in an obligation to make changes depended on the type of external peer review programme, but also if the programme had been made mandatory. An interesting informal pressure faced by some AIMS PICU participants was also reported despite the accreditation scheme remaining voluntary. Whether a service was publicly or privately funded influenced the perceived ability of participation in external peer review to bring about changes. Length of membership was the most salient inner context perceived to influence external peer review, with increased familiarity and repeated exposure seen as beneficial to gaining value and being able to make changes through external peer review. However, this was contrasted by some participants' views on less engagement and less value gained by long-standing members in recent years (despite the value they may add to programmes).

The lack of awareness of external peer review was picked up by both ward managers and senior managers, and was felt to be an important aspect of readiness for change. Subjects felt more strategies were needed for better

information accessibility and communication to all levels, namely frontline staff. Within reported awareness rates amongst subjects, frontline subjects belonging to AIMS PICU appeared to have a better understanding than those belonging to peer review networks, although this was still felt in a few cases. Amongst those frontline subjects who demonstrated a good understanding of the programmes, a sense of teamwork, empowerment and ownership of the programmes was felt. In this thread, as one would expect, awareness was strongly related to involvement. Understanding was also perceived to be linked to length of membership, as subjects from longer-standing services reported their understanding increased after having repeated exposure over a period of time. Understanding and increase in familiarity was also mentioned in relation to peer reviewing other services, and participating in the additional CCQI services. Experiencing apprehension before visits was a barrier commonly experienced by senior and ward managers. However, this was reported to be overcome through repeated exposure, increase in familiarity and peer reviewing other services as well. Reluctance to change appeared to be associated with the length of time staff had been working in services, and numerous examples of this were provided by subjects.

## **Chapter 11 Results of Cross-sectional Study**

The results chapter presents the findings from the Cross-sectional National Memory Clinics Audit study that aims to assess which readiness for change constructs influence quality improvement effect change in external peer review in community-based Memory Clinics.

### **11.1 Response Rate**

178 out of an estimated 214 memory clinics in England responded to the English National Memory Clinics Audit 2013, resulting in a response rate of 83%.

### **11.2 Service Characteristics**

Memory clinics from 63 NHS Trusts completed the survey, and one private memory clinic. Of these clinics, 100% completed the readiness for change items in the audit. 118 clinics were not MSNAP members at the time of the audit, and 60 clinics were members (this comprised of the following: 15 were accredited as excellent, 12 were accredited, 27 clinics were in review, and 6 were affiliate).

### 11.3 Summary of Readiness for Change Data

Firstly, I used descriptive statistics to visualise the data. Summary statistics are presented in Table 11.1 on the five Likert-type items that were included in the National Audit. This helped me to observe how respondents had answered the questions, and to see if there were any trends to inform further analyses.

**Table 11.1 Descriptive statistics for five readiness for change constructs**

Variable	SA	A	N	D	SD	Mean	Median	Mode
Leadership	75	86	13	3	1	4.45	4.00	4.00
Adapt new standards	82	83	10	6	0	4.43	4.00	5.00
Flexible	75	83	14	6	0	4.28	4.00	4.00
Try new things	117	57	4	0	0	4.72	5.00	5.00
Change so fast	30	49	39	51	9	3.33	3.00	2.00

SA = Strongly Agree, A = Agree, N = Neutral, D = Disagree, SD = Strongly Disagree

All the central measures of tendency for the construct 'leadership in this unit articulates a vision for the unit' indicated 'agree' (n=86).

Although the mean and median responses for 'this unit is typically able to adapt new standards or procedures, even those forced upon us' were 'agree', most of the responses (n=82) were 'strongly agree'. 83 respondents 'agreed' with this statement, and 10 remained neutral. No respondents 'strongly disagreed' and only three disagreed.

All central measures of tendency for the construct 'this unit tends to be flexible in dealing with change' were for the response 'agree' (n=83). 75 respondents strongly agreed with this statement, and 14 remained neutral. No respondents 'strongly disagreed' and only six 'disagreed'.

All central measures of tendency for the item 'this unit is generally willing to try new things to improve service user care' were for the response 'strongly agree'

(n=117). 57 respondents agreed with this statement and four remained neutral. No respondents ‘strongly disagreed’ or ‘disagreed’ with this statement.

For the construct ‘things change so fast in this unit it’s hard to keep up with what’s going on’ the mean and median responses were ‘neutral’ (n=39) and the most frequent response was ‘disagree’ (n=51). 49 respondents ‘agreed’ with the statement and 30 ‘strongly agreed’. Only nine respondents ‘strongly disagreed’.

All of the readiness for change items, apart from ‘Things change so fast in this unit it’s hard to keep up with what’s going on’ had interquartile ranges between the responses ‘agree’ and ‘strongly agree’. The interquartile range for ‘Things change so fast in this unit it’s hard to keep up with what’s going on’ was much broader, between ‘disagree’ and ‘strongly agree’.

**Table 11.2 Summary of outcomes of national memory clinics audit**

<b>Outcomes</b>	<b>Mean</b>	<b>Median</b>	<b>Mode</b>
Primary outcome (provision of services and treatments)	5.98	6.00	7.00
Secondary outcome (waiting times)	5.48	4.00	4.00

The primary outcome (presented in Table 11.2) was a composite score calculated by summing seven key components of quality of care (provision of home based assessments, access to specialist post-diagnostic counselling, initiation of anti-dementia medication, review of anti-dementia medication, access to cognitive stimulation therapy, access to education and support for carers, access to Life Story work). The most frequent response was that all memory clinics were providing all seven provisions. The mean and median responses were that memory clinics were providing six out of seven provisions. The interquartile range showed a similar spread, between six and seven provisions.

The secondary outcome waiting time (presented in Table 11.2) was characterised by the average number of weeks’ wait between receipt of referral and the person starting their assessment. These values ranged between one and

25 weeks. The mean waiting time was five weeks, with the mode and median being four weeks. The interquartile range showed a spread between three and seven weeks. Both the primary and secondary outcomes were not normally distributed so log transformation was necessary.



#### 11.4 The relationship between readiness to change constructs and the primary and secondary outcomes of the National Audit

**Table 11.3 Relationship between readiness to change and the provision of services and treatments in memory clinics**

Variable	Standard Deviation	Correlation Co-efficient	P-value
Leadership	0.56524	-0.09857	0.4537
Adapt to new standards	0.59280	-0.05325	0.6862
Flexible	0.66617	0.17430	0.1829
Try new things	0.49030	-0.07172	0.5860
Changing so fast	1.15958	-0.25318	0.0510

**Table 11.4 Relationship between readiness to change and waiting times in memory clinics**

Variable	Standard Deviation	Correlation Co-efficient	P-value
Leadership	0.56524	-0.17896	0.1713
Adapt to new standards	0.59280	0.01862	0.08877
Flexible	0.66617	-0.16382	0.2110
Try new things	0.49030	-0.07502	0.5689
Changing so fast	1.15958	0.12912	0.3255

As presented in Tables 11.3 and 11.4, both relationships between readiness to change and the primary and secondary outcomes were not statistically significant, thus indicating there was limited probability of observing the correlation coefficient or one more extreme under the null hypothesis.

Furthermore, the Pearson correlation coefficients were not at all close to +1 (indicative of a perfect positive correlation) or -1 (indicative of a perfect

negative correlation). As they were closer to 0, this is indicative of no correlation at all.

### **11.5 The differences between accredited and non-accredited memory clinics**

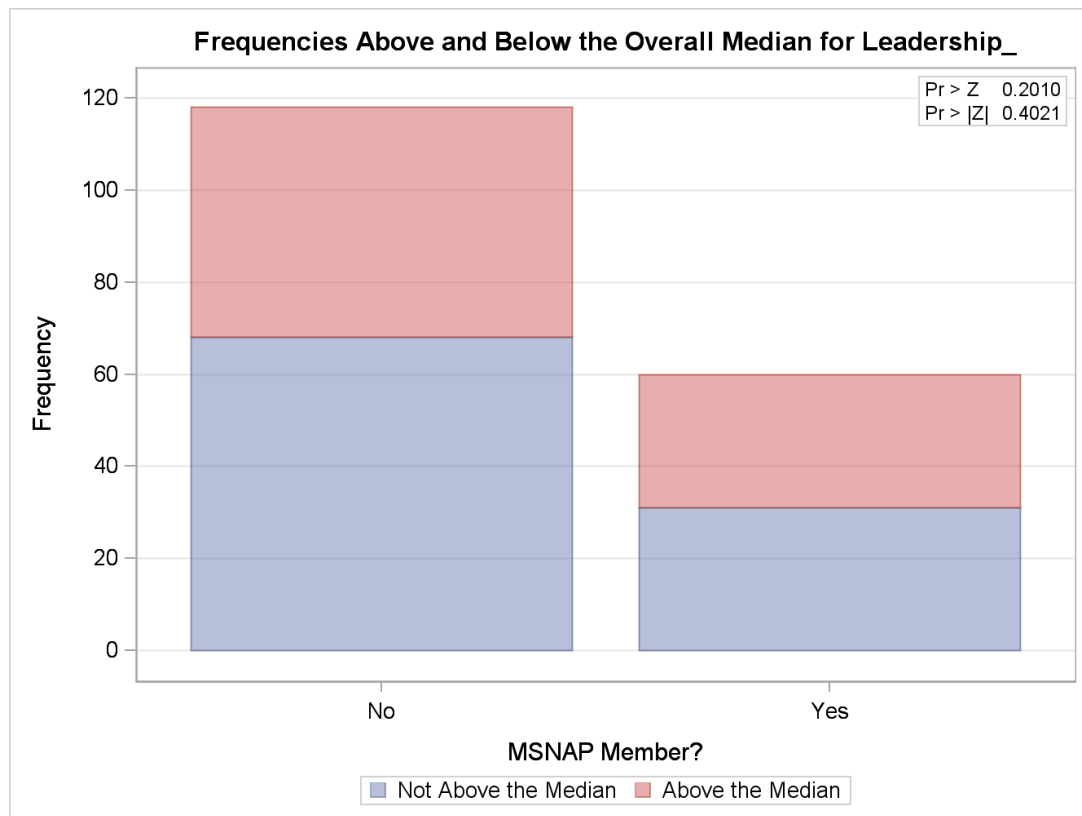
I used box plots to illustrate the median as the black horizontal line inside the box, and the inter-quartile range as the length of the box. The inter-quartile range indicates the 25<sup>th</sup> to 75<sup>th</sup> percentiles, as the range in which the central 25-75% (50%) of the data points lie. The whiskers represent the minimum and maximum values when they are within 1.5 times above or below the inter-quartile range.

I used histograms to show the frequency of measurements and the shape of the data, as they provide a visual judgement of whether the distribution approximates to a bell-shape. Histograms also helped to show whether there were any gaps in the data, whether there are any outlying values and how far any outlying values were from the remainder of the data.

For the leadership construct, presented in Figure 11.3, the results suggested that there is no statistically significant difference between the underlying distributions of leadership scores of accredited memory clinics and the leadership scores of memory clinics who have not been accredited ( $z = 1.0772$ ,  $p = 0.2814$ ).

The p-value of the Median Test was 0.4021. This did not provide sufficient evidence to reject the claim that there is no significant difference between leadership scores for accredited and non-accredited memory clinics.

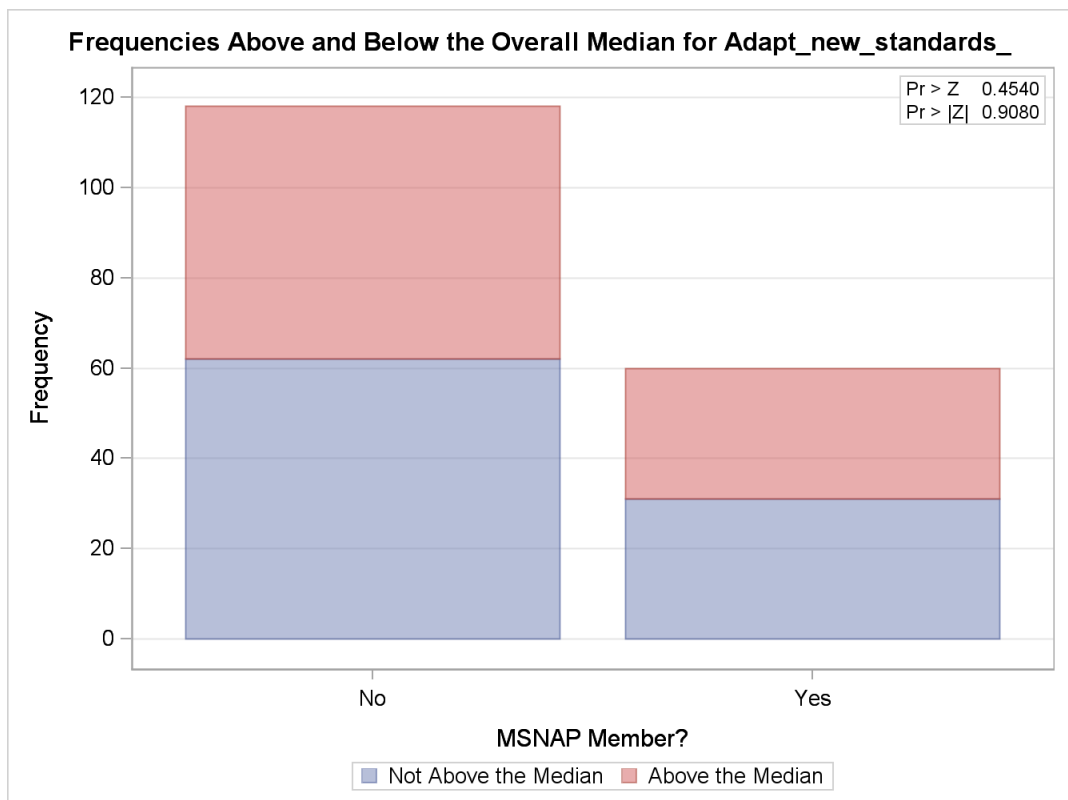
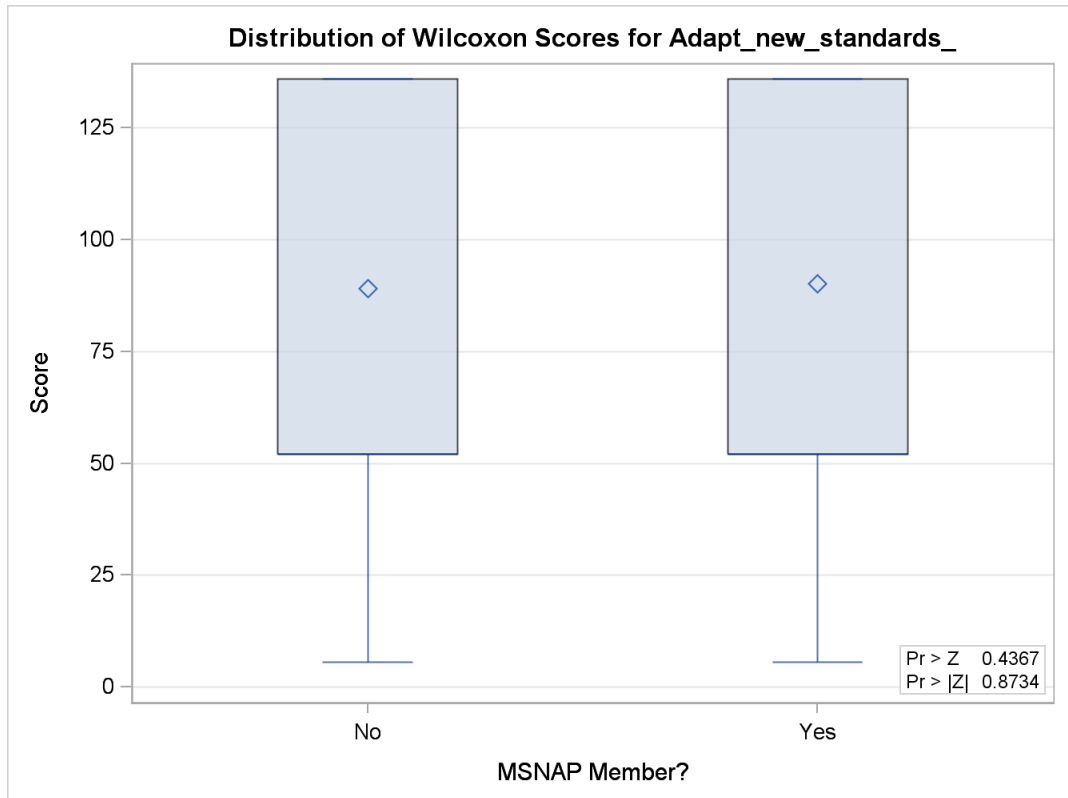
**Figure 11.1 Differences in leadership between accredited and non-accredited memory clinics**



For the adapt new standards construct, the results presented in Figure 11.4 suggested that there is no statistically significant difference between the underlying distributions of adapt new standards scores of accredited memory clinics and adapt new standards scores of memory clinics who have not been accredited ( $z = 0.1593$ ,  $p = 0.8734$ ).

The p-value of the Median Test was 0.9080. This did not provide sufficient evidence to reject the claim that there is no significant difference between adapt new standards scores for accredited and non-accredited memory clinics.

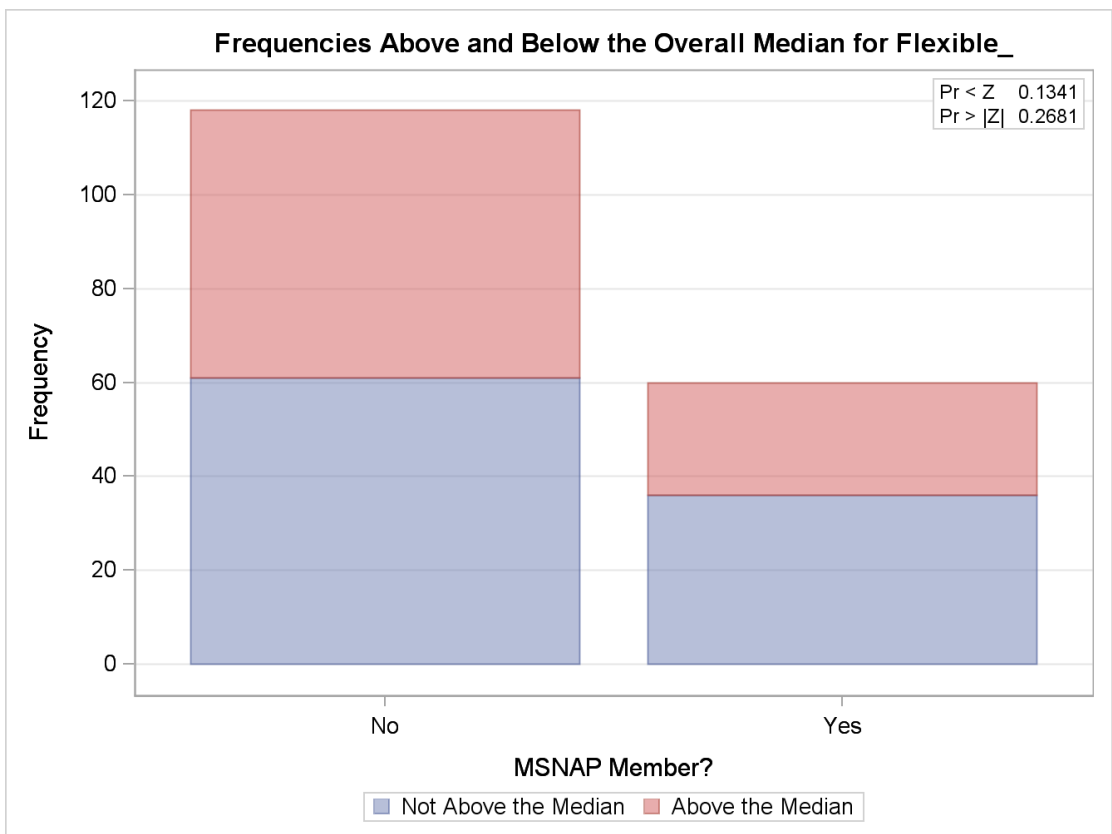
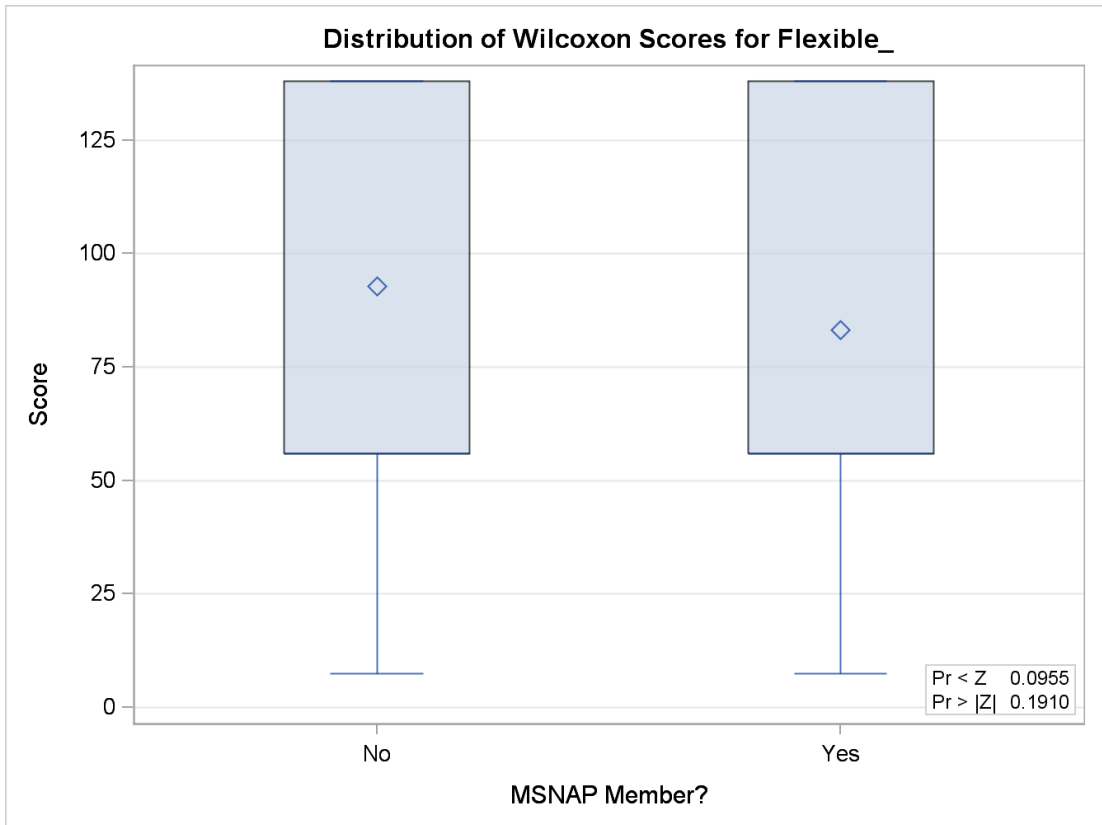
**Figure 11.2 Differences in adapting to new standards between accredited and non-accredited memory clinics**



The flexible construct presented in Figure 11.5 suggested that there was no statistically significant difference between the underlying distributions of flexible scores of accredited memory clinics and the flexible scores of memory clinics who have not been accredited ( $z = -1.3076$ ,  $p = 0.1910$ ).

The p-value of the Median Test was 0.2681. This did not provide sufficient evidence to reject the claim that there is no significant difference between flexible scores for accredited and non-accredited memory clinics.

**Figure 11.3 Differences in flexibility between accredited and non-accredited memory clinics**

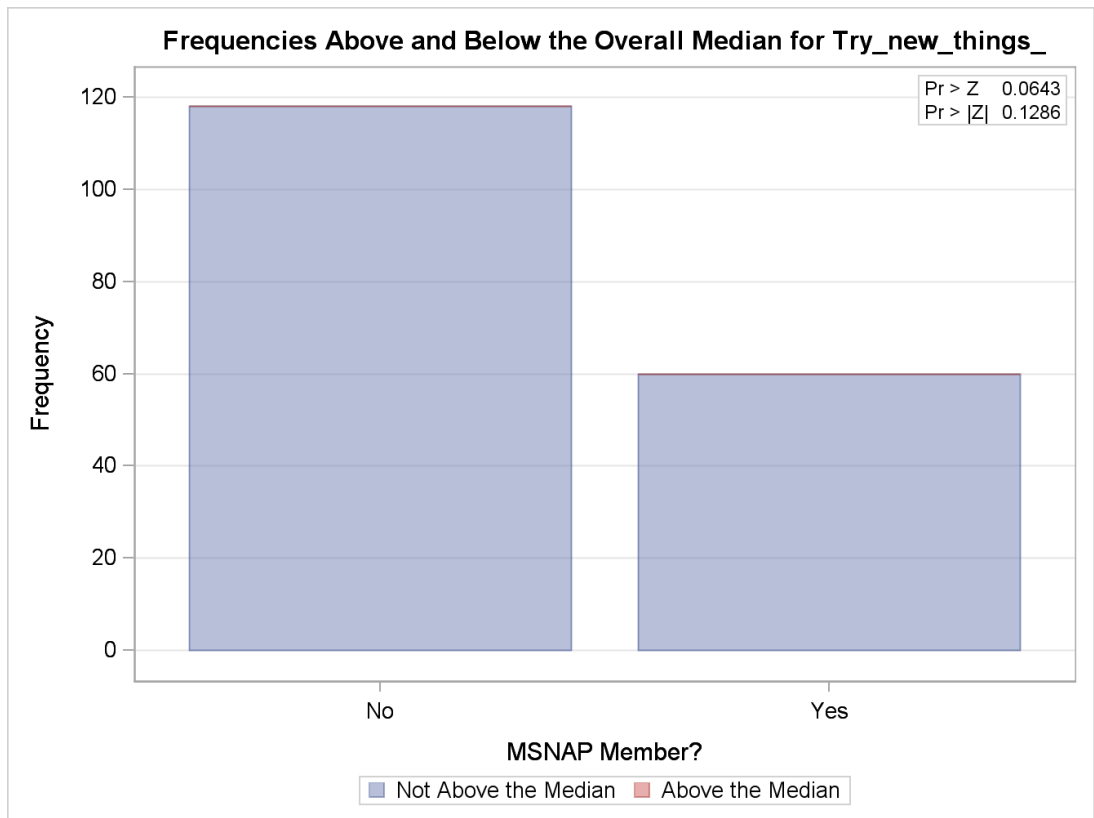
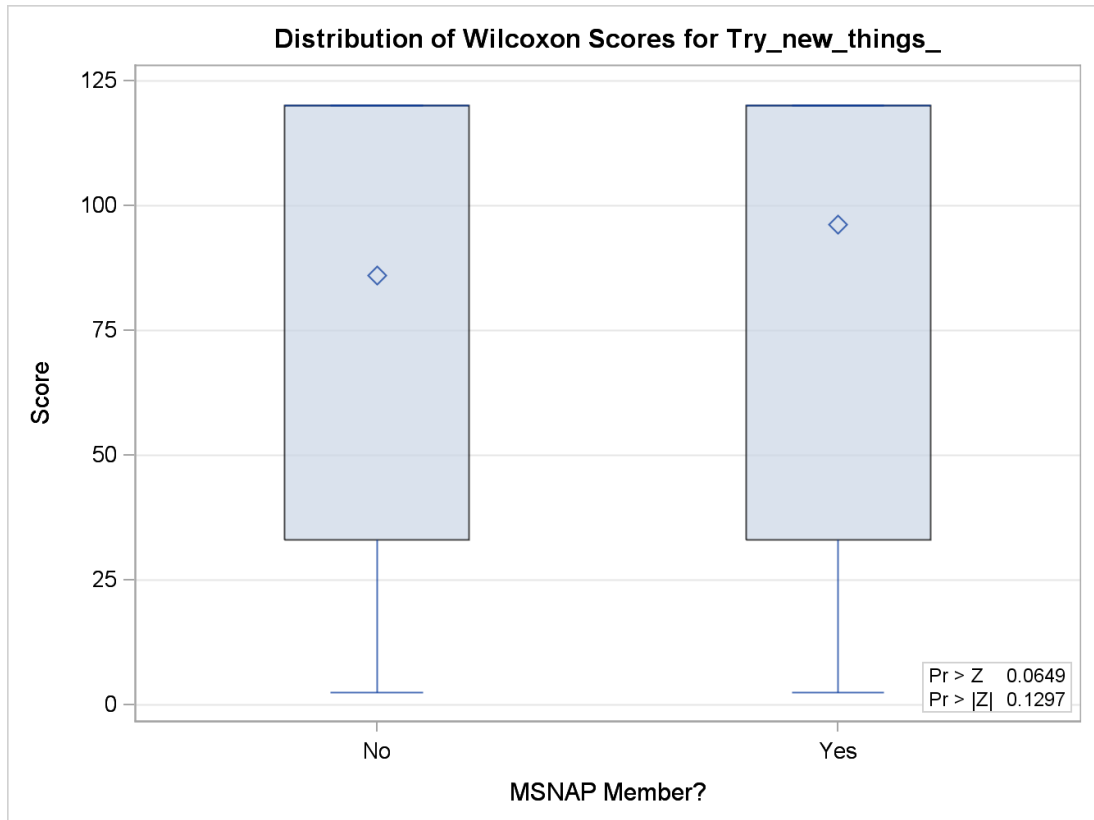


The try new things construct presented in Figure 11.6 suggested that there was no statistically significant difference between the underlying distributions of try new things scores of accredited memory clinics and the try new things scores of memory clinics who have not been accredited ( $z= 1.5152$ ,  $p = 0.1297$ ).

The p-value of the Median Test was 0.1286. This did not provide sufficient evidence to reject the claim that there is no significant difference between try new things scores for accredited and non-accredited memory clinics.



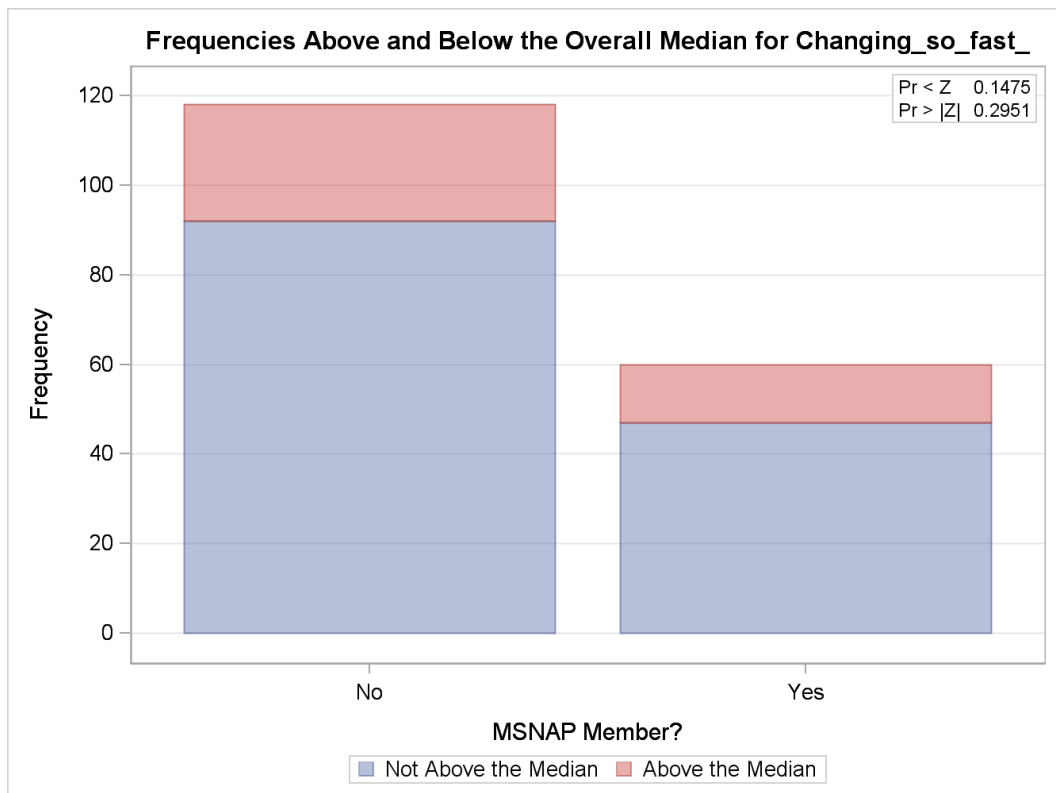
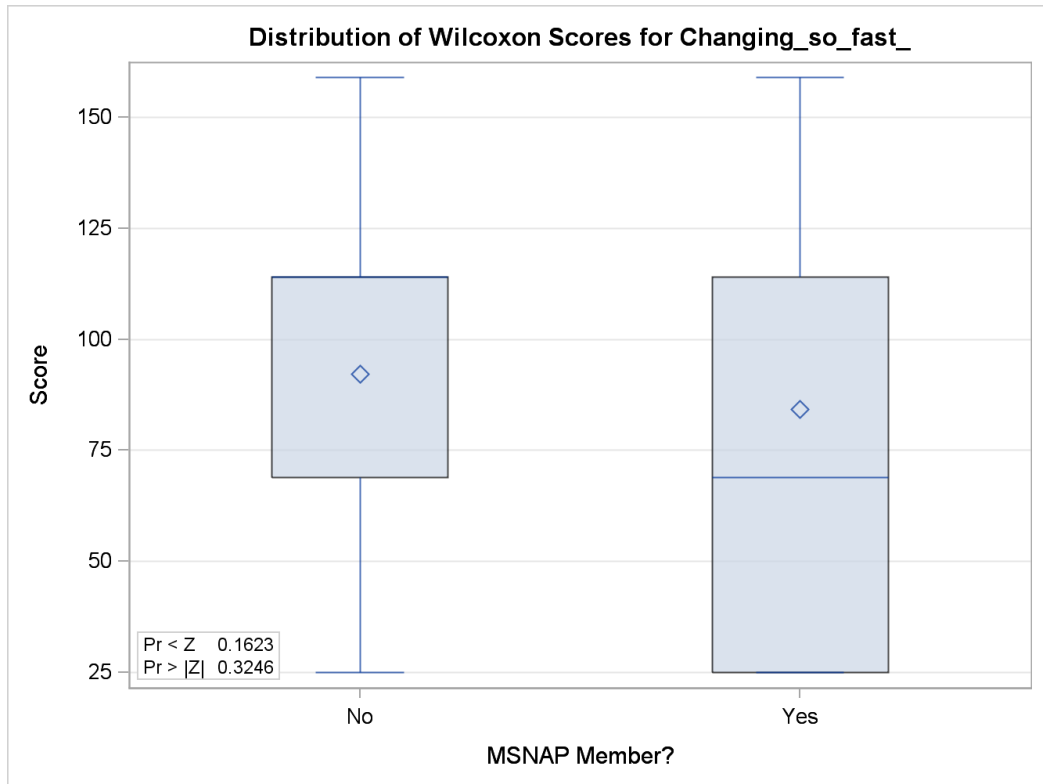
**Figure 11.4 Differences in trying new things between accredited and non-accredited memory clinics**



The results of the changing so fast construct presented in Figure 11.7 suggested there was no statistically significant difference between the underlying distributions of changing so fast scores of accredited memory clinics and the changing so fast scores of memory clinics who have not been accredited ( $z = -0.9851$ ,  $p = 0.3246$ ).

The p-value of the Median Test was 0.2951. This did not provide sufficient evidence to reject the claim that there is no significant difference between changing so fast scores for accredited and non-accredited memory clinics.

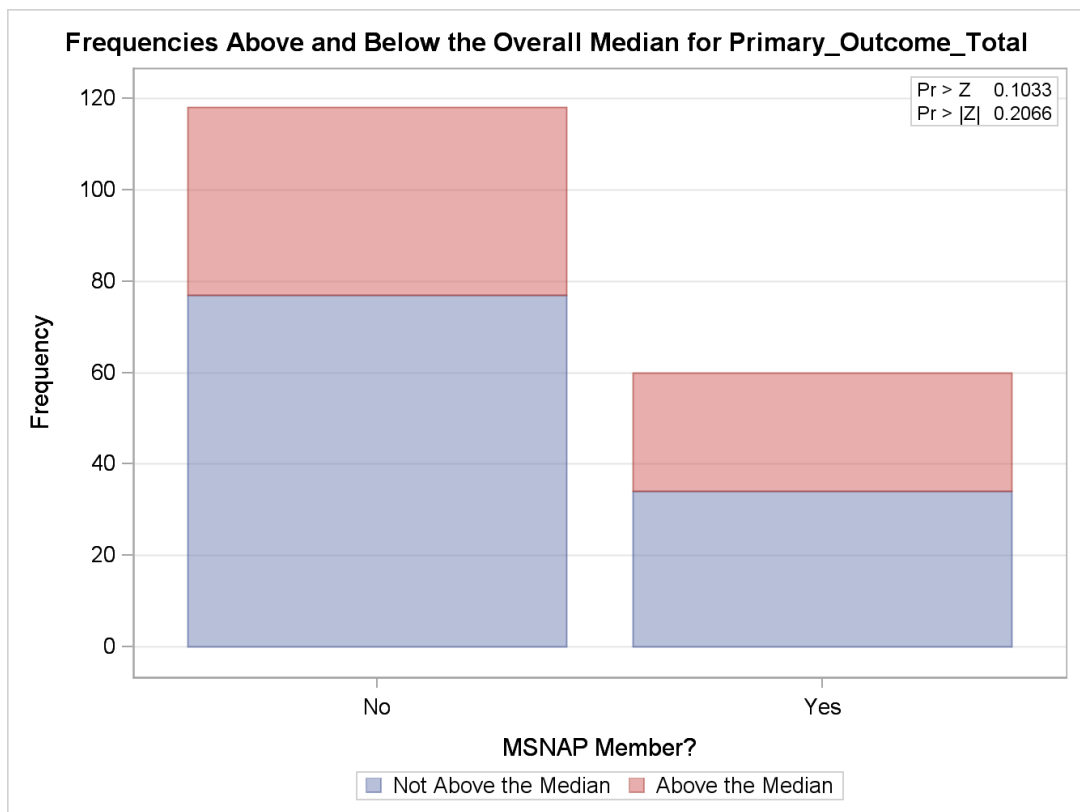
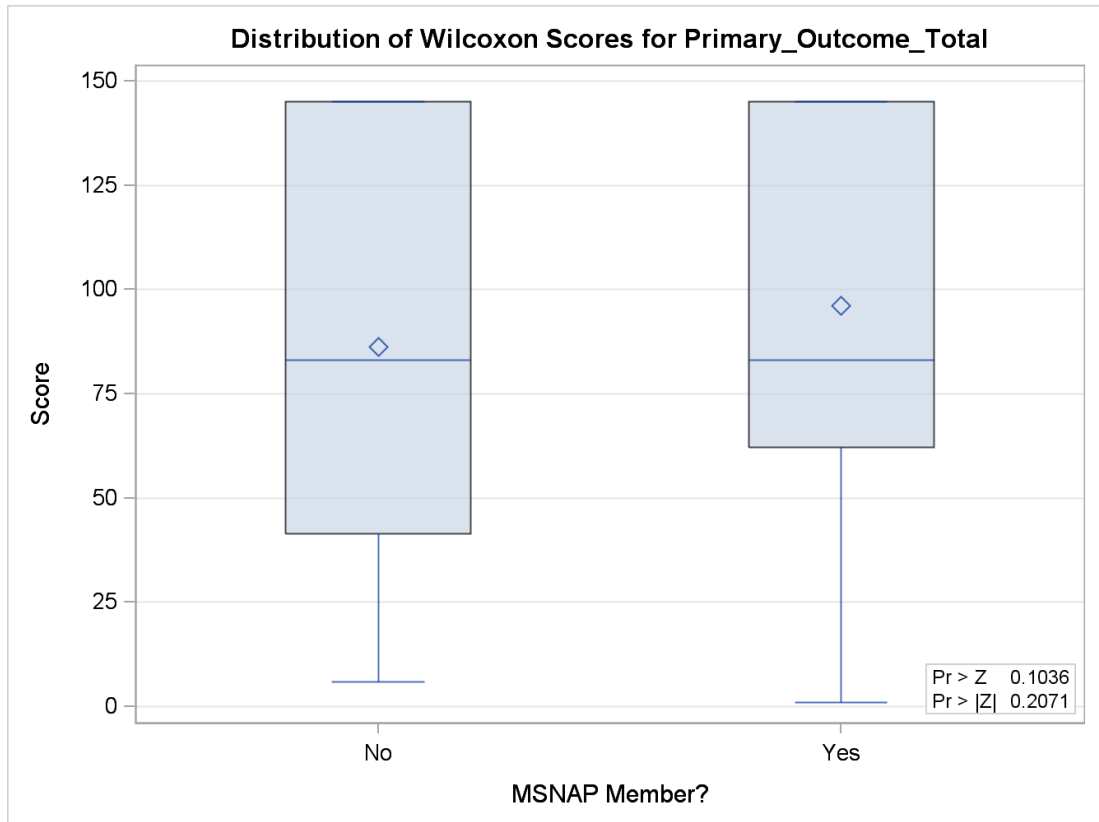
**Figure 11.5 Differences in changing so fast between accredited and non-accredited memory clinics**



For the primary outcome presented in Figure 11.8 the results suggested that there was no statistically significant difference between the underlying distributions of provision of facilities / access to services of accredited memory clinics and of memory clinics who have not been accredited ( $z= 1.2615$ ,  $p = 0.2071$ ).

The p-value of the Median Test was 0.2066. This did not provide sufficient evidence to reject the claim that there is no significant difference between changing so fast scores for accredited and non-accredited memory clinics.

**Figure 11.6 Differences in the provision of services and treatments between accredited and non-accredited memory clinics**



Based on the covariance estimates, I calculated the intraclass correlation as 0.4379, which is the portion of total variance that occurs between accredited and non-accredited memory clinics. The fit statistics were: AIC = 363, BIC = 366.1 (F-value = 0.65). As smaller is considered to be better, these large numbers indicated there was no correlation. The t-value and P>t are testing Null hypotheses that the estimate = 0.

## 11.6 Multivariate Model

### 11.6.1 Waiting time

Based on the covariance estimates for the secondary outcome of waiting time, the intraclass correlation was 0.3656, which is the portion of total variance that occurs between accredited and non-accredited memory clinics. Fit statistics: AIC = 221.3, BIC = 223.3

**Table 11.5 Observed significance levels for waiting time**

Effect	F Value	Pr > F
Leadership	0.50	0.6109
Adapt to new standards	0.08	0.9273
Flexible	1.17	0.3172
Try new things	1.41	0.2535
Changing so fast	0.96	0.4168
MSNAP member	0.03	0.8746
Provision of treatments and services (primary outcome)	0.91	0.3431
Maximum new patients (log)	0.01	0.9396
Specialist posts	2.60	0.1123
Education and support	1.47	0.2310
Planning changes	0.95	0.4233
Register interest	0.78	0.3813

Table 11.5 presents the observed significance levels of the hypothesis tests for each of the fixed effects, which were specified in the model statement. P-values for the test were calculated as the tail area from an F distribution. As these p-values were all above 0.05, the model yielded results which were not statistically significant.

### 11.6.2 Maximum number of new service users

Based on the covariance estimates for another outcome of the maximum number new service users that the service can assess per week, I calculated the intraclass correlation to be 0.3004, which is the portion of total variance that occurs between accredited and non-accredited memory clinics. Fit statistics: AIC = 210.2, BIC = 212.2

**Table 11.6 Observed significance levels for maximum new service users**

Effect	F Value	Pr > F
Leadership	1.74	0.1853
Adapt to new standards	0.12	0.8882
Flexible	0.66	0.5189
Try new things	1.41	0.2525
Changing so fast	0.04	0.9881
MSNAP member	0.69	0.4084
Provision of treatments and services (primary outcome)	0.03	0.8698
Maximum new patients (log)	0.22	0.6432
Specialist posts	4.40	0.0405
Education and support	0.00	0.9466
Planning changes	0.99	0.4023
Register interest	3.48	0.0674

Table 11.6 presents the observed significance levels for the hypothesis tests for each of the fixed effects, which were specified in the model statement. P-values for the test were computed as the tail area from an F distribution. As these p-values were all over 0.05, the model yielded results which were not statistically significant.



## **11.7 Conclusion**

Overall, the relationship between the five selected constructs of readiness for change were not statistically significantly associated with either the primary or secondary outcomes of quality. Whether memory clinics were accredited or not accredited also did not have a significant effect.

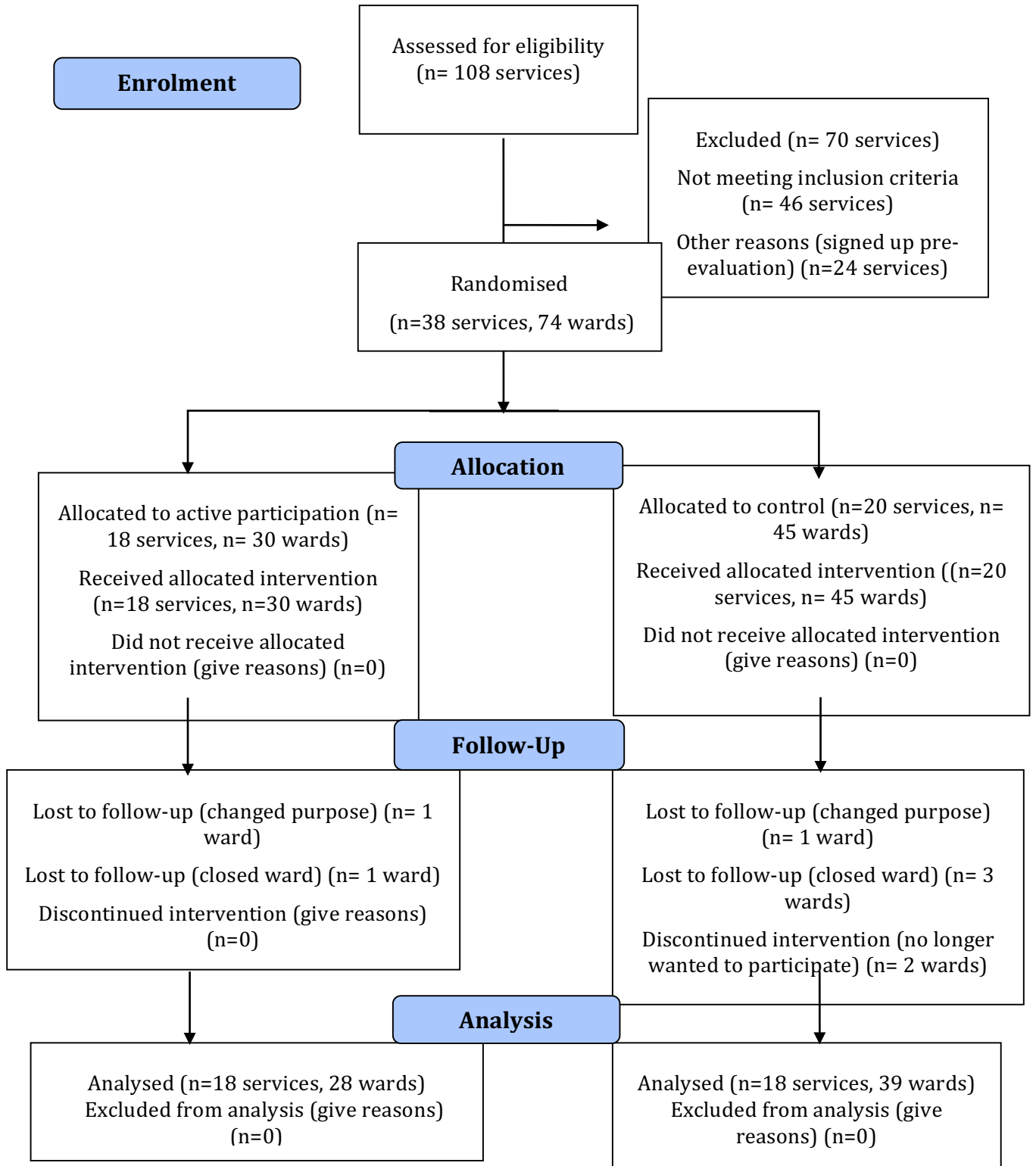
For the multivariate model, there did not appear to be a significant difference between accredited and non-accredited memory clinics for the secondary outcome of waiting time or another selected outcome of the maximum number of new service users that the service can assess in a week.

## **Chapter 12 Results of Longitudinal Study**

This results chapter presents the findings from the Readiness for Change Longitudinal Study. The study assesses how organisational readiness for change scores can influence the outcomes of service development.

The CONSORT (Consolidated Standards for Reporting Trials Statement) 2010 Statement ([Moher et al., 2010](#)) is presented in Figure 12.1 and consists of a 25-item checklist and respondent flow diagram ([Schulz et al., 2010](#)), which displays the progress of all respondents through the evaluation.

**Figure 12.1 CONSORT flow-chart showing participation in the eLSU trial among eligible services (Schulz et al., 2010)**



## 12.1 Response Rate

Readiness for change checklists were completed by ward or service managers at low secure hospitals. In total, completed readiness for change checklists were collected from 101 respondents (from both baseline and follow-up) from 33 different low secure services.

QELS checklists were also completed from respondents from 33 different low secure services, once at baseline and at follow-up.

## 12.2 Relationship between baseline readiness for change, and baseline QELS checklist

**Figure 12.2 Relationship between baseline readiness for change and baseline in QELS checklist score**

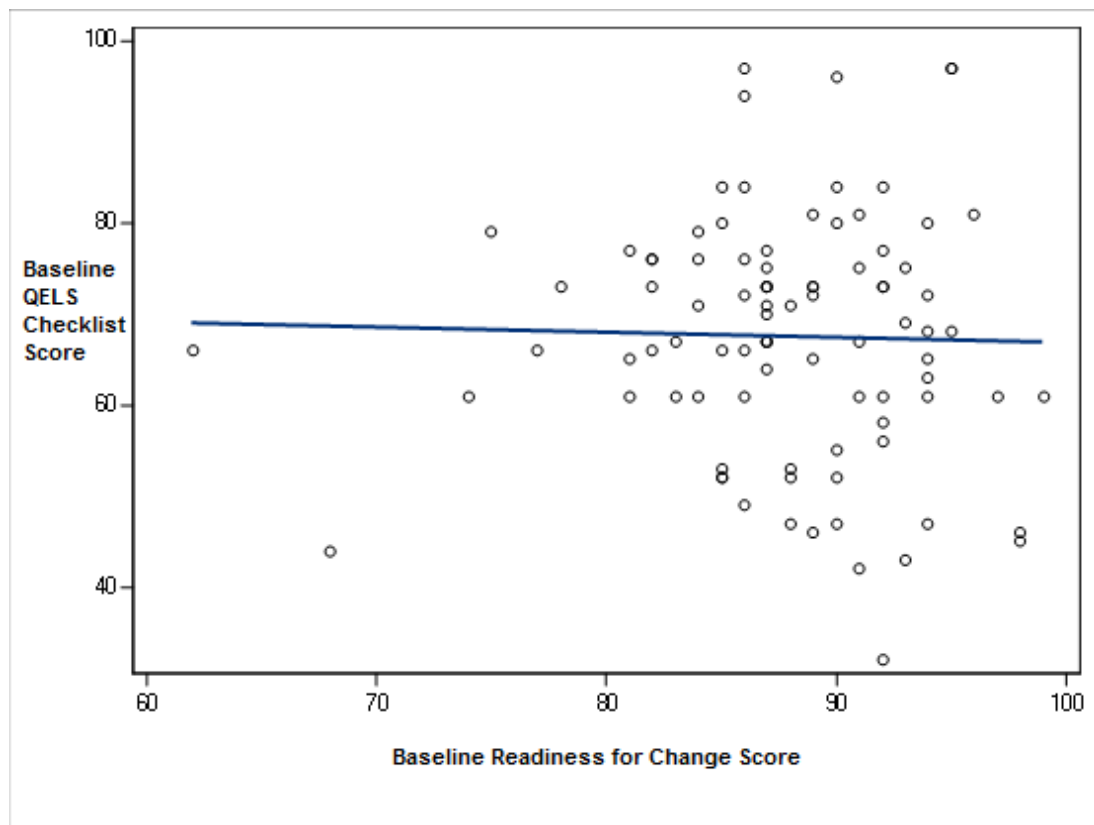
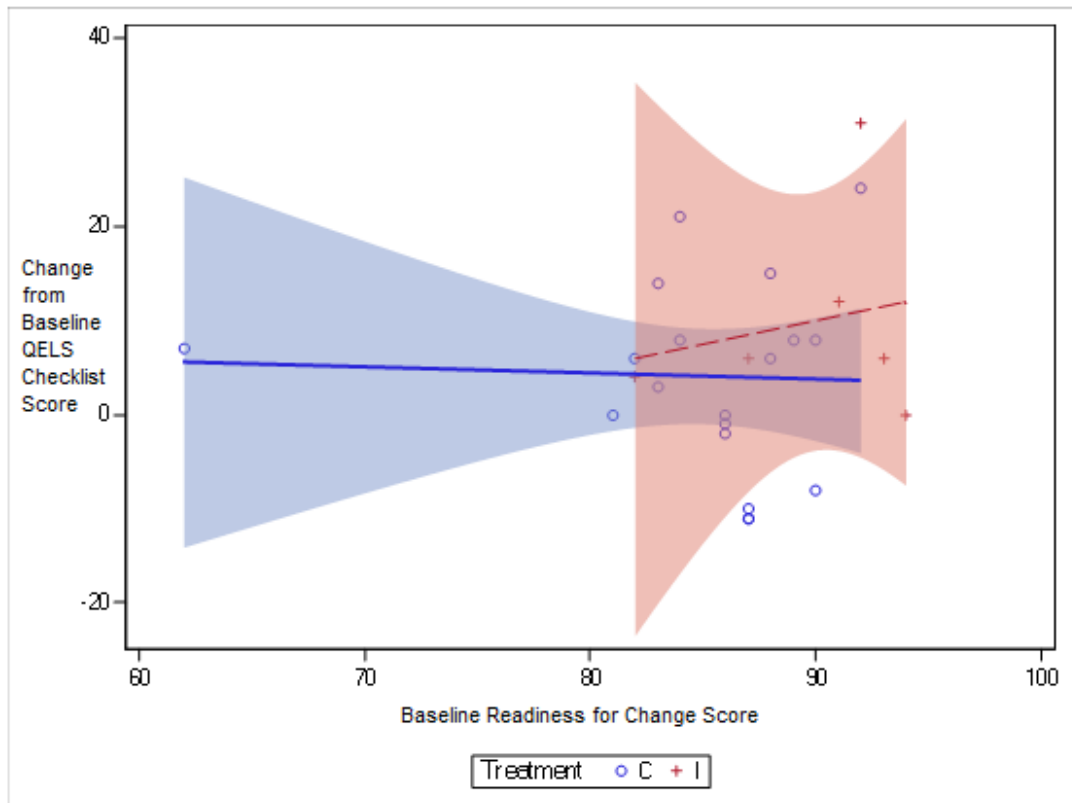


Figure 12.2 illustrates the majority of all (both control group and intervention group) wards had high readiness for change scores, but there was no observed relationship at this early stage.

### 12.3 Relationship between baseline readiness for change and change from baseline in QELS checklist score

Figure 12.3 Relationship between baseline readiness for change and change from baseline in QELS checklist score



There was no linear relationship in Figure 12.3 for the control arm. There appeared to be some grouping around the centre for the intervention arm. Fit statistics were AIC = 434.5, BIC = 439.1. The super-imposed coloured areas show 'uncertainty around the slope', at a 95% CI of the slope. These are the confidence intervals that show for both the control and intervention groups, there is a lot of room for error within the regression line that has been fitted.

**Table 12.1 Covariance parameter estimates between control and intervention LSUs**

	<b>Covariance Parameter Estimates</b>			
<b>Covariance Parameter</b>	<b>Group</b>	<b>Estimate</b>	<b>Lower</b>	<b>Upper</b>
<b>Intercept</b>		73.0840	41.3424	162.76
<b>Residual</b>	Control	29.3320	18.0087	56.0690
<b>Residual</b>	Intervention	11.9147	4.2835	97.7560

As presented in Table 12.1: based on this output, I calculated the Intraclass Correlation Coefficient (ICC) for the control group as:

$$ICC = 73.0840 / (73.0840 + 29.3320) = 7.1360$$

Based on this output, I calculated the ICC for the control group as:

$$ICC = 73.0840 / (73.0840 + 11.9147) = 0.8598$$

This indicates that 85% of the variability in achieving an increased QELS score is accounted for by intervention services, leaving 15% of the variability to be accounted for by comparison wards, which provides support for using a two-level model. The large proportion of the variability explained by intervention services emphasises the importance of accounting for the hierarchical nature of the data.

## 12.4 Relationship between baseline readiness for change and change from baseline in QELS checklist score (accounting for the same respondent)

Figure 12.4 Relationship between baseline readiness for change and change from baseline in QELS checklist score (accounting for the same respondent)

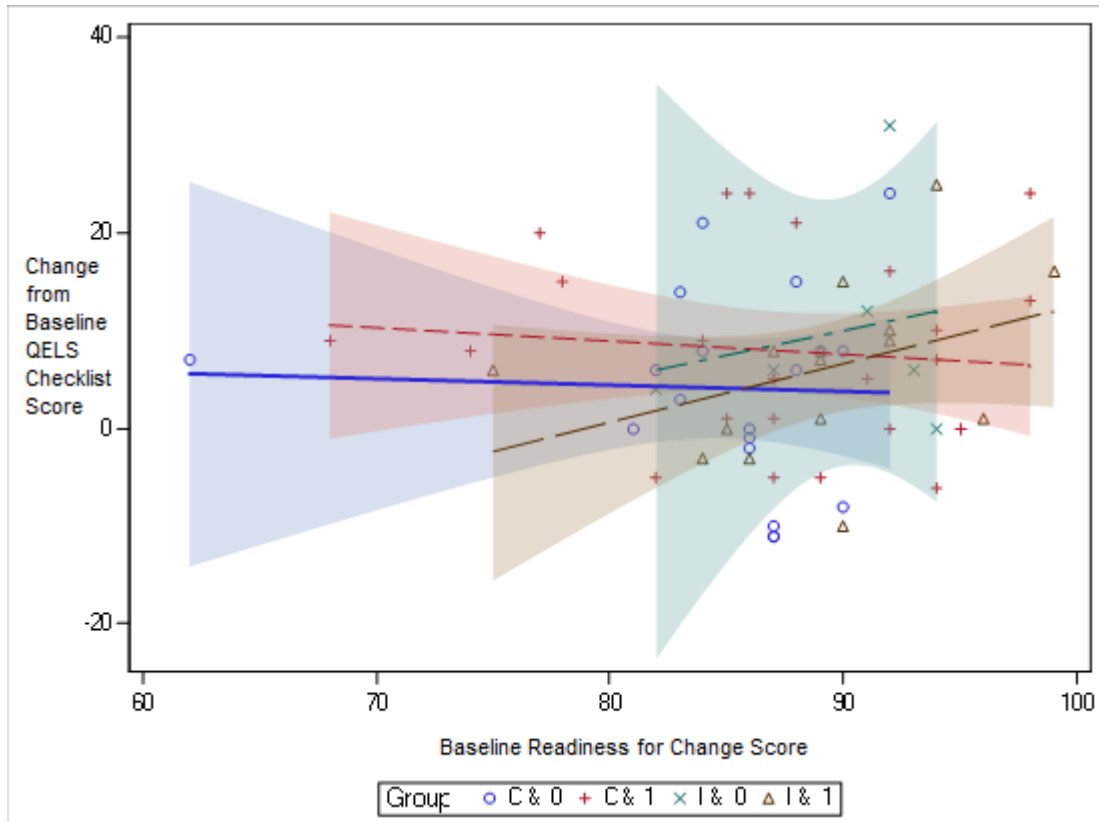


Figure 12.4 shows the relationship between baseline readiness for change and change from baseline in the QELS checklist score accounting for random effects (when the same respondent (indicated by group 1 on Figure 12.4) completed the QELS checklist during baseline and follow-up), illustrating differences between control (C) and intervention (I) arms. The point estimates show that there is more variability between two different respondents than with the same respondents. As before, there is a lot of variability, but Figure 12.4 signals towards the hypothesised relationships. Fit statistics: AIC = 440, BIC = 447.6.

**Table 12.2 Covariance parameter estimates between control and intervention groups accounting for the same respondent**

<b>Covariance Parameter Estimates</b>			
<b>Group</b>	<b>Estimate</b>	<b>Lower</b>	<b>Upper</b>
	75.8363	41.8932	177.13
Control, different respondents	61.9977	32.4367	162.52
Control, same respondents	14.3607	7.1430	42.4758
Intervention, different respondents	69.1179	12.3516	422946
Intervention, same respondents	9.8200	3.6170	73.0206

Referring to Table 12.2 and Table 12.3, there appeared to be a relationship between the same respondents as the p-value was smaller for same respondents.

**Table 12.3 Observed significance levels from two-level organisational linear model predicting quality improvement (QELS checklist) scores based on readiness for change (n=33 services)**

<b>Effect</b>	<b>F Value</b>	<b>Prob &gt; [F]</b>
QELS checklist change from baseline	8.87	0.0048
Readiness for change baseline	0.94	0.3453
Intervention	0.04	0.8348
Same respondent	0.01	0.9110

I went on to investigate if services in the active arm of the trial demonstrated greater readiness for change for change than those in the control arm, as this



was indicated in the qualitative findings, however this association was not present.

## **12.5 Conclusions**

Findings assessing the relationship of readiness for change and QELS checklist scores which was the primary outcome of the eLSU study (compliance to environmental standards) were inconclusive. While trends were observed in those with higher baseline readiness for change scores exhibiting greater differences in the change from baseline of the primary outcome; these differences were not statistically significant. The results suggested that baseline organisational readiness for change may influence the ability of a service to make use of an external peer review programme, but new research is needed to further test this hypothesis.

## Chapter 13 Discussion

In this chapter, I place the results in reference to the systematic literature review and the wider healthcare context. I present the conclusions and associated recommendations I have drawn, set within an appropriate recognition of the overall methodological considerations. I have recognised and, where possible, specified recommendations for practice, policy and future research.

Realist evaluation proposes that identification of programme theories should precede testing and refining those theories ([Pawson, 2002](#)), ([Pawson and Tilley, 1997](#)). I attempted to extensively develop CMOCs through undertaking a systematic literature review, facilitating focus groups with practitioners who coordinate external peer review programmes, and testing out constructs of organisational readiness for change in cross-sectional and longitudinal quantitative studies. During the process of testing the CMOCs with subjects (staff who work in member services) through semi-structured interviews, new concepts were identified which were not initially formulated as part of the MRT. This often necessitated the need to revisit the initial MRT to accommodate these new concepts. This meant that the process of developing, testing and refining the MRT was not entirely sequential but rather an iterative one. Although, this proved time consuming and complex initially, it nevertheless offered more depth to the analysis.

### **13.1 Summary of Main Findings**

Overall, the studies uncovered many positive findings, some of which were in agreement with the extant literature.

Practitioners and subjects were asked when the most changes took place during external peer review. Responses differed between accreditation programmes and peer review networks. Subjects participating in AIMS PICU confirmed the most change during the preparation stage either before joining, before or during the self-review stages. Subjects participating in QNFMHS also confirmed the majority of change occurred during feedback stages, in particular, following their receipt of the written report. These were initially suggested by practitioners.

Causal mechanisms were categorised into mechanisms within the programme factors, host team processes and staff groups. The most substantial of these causal mechanisms was sharing with and learning from peers, which took place during the peer review visit, verbal feedback, peer reviewing other services, and through the range of additional activities offered by the CCQI (i.e. online forum, newsletter, annual forum, training, workshops). Peer reviewing other services was viewed as the most valued programme factor in external peer review, it was closely linked with the most essential causal mechanism of sharing and learning as these were the two most frequently cross-referenced themes within the interview data. It was interesting that the programme factors that were most valued, and where most change occurred were different.

Both the importance of senior management support and junior staff involvement were equally highlighted as a causal mechanism of staff groups that need to be involved to achieve success through external peer review.

A variety of changes were reported to have taken place through the accounts of subjects, and these were broadly categorised as long-term changes, short-term changes, negative changes, no change, and perceptions of change. These changes were linked to the overall end goal of external peer review, which is a

measurable improvement in percentage compliance with CCQI standards, and this is how quality improvement is defined in these external peer review programmes.

The most influential outer context was the type of external peer review programme, the broader healthcare system in which services were operating in and other quality improvement programmes they were participating in. Inner contexts were influential at the organisational level, and the most salient was length of membership.

Readiness for change was a complex context to assess, as it had both organisational and team-level components ([Weiner, 2009](#)). Upon quantitative assessment of both a cross-sectional study of memory clinics, and a longitudinal study of LSUs there was no significant relationship observed to suggest that readiness for change could influence the ability of either memory clinics or LSUS to bring about service improvements through an external peer review programme. However, the qualitative data contradicted this finding, and indicated that readiness for change was a multi-faceted contextual factor that influenced the ability of services to be able to benefit from external peer review and make changes. In particular, readiness for change constructs such as awareness, understanding and being connected to other units were salient. Subjects reported that being connected with other units was also closely linked to the most essential causal mechanism of sharing and learning to effect change.

## 13.2 Refined CMOCs

Employing realist evaluation helped to unravel some of the nuances and inherent contradictions and dilemmas brought about by external peer review. During the process of developing and refining the CMOCs, it was often clear that an alternative CMOC could be generated. Other studies have often reported similar challenges ([Tolson et al., 2005](#)), ([Byng et al., 2005](#)). However, Rycroft-Malone et al. ([2010](#)) indicated that realism, the philosophical base of realist evaluation, makes provision for two or more mechanisms to operate concurrently. Six main CMOCs were developed from this research, and are displayed in Table 13.1. CMOCs are the final stage of the realist evaluation. Throughout data collection and analysis, I developed the most salient CMOCs relating to which mechanisms, in which contexts result in long-term changes necessary for quality improvement were developed.

**Table 13.1 Realist matrix of context-mechanism-outcome configurations**

<b>Causal Mechanism</b>	<b>Outer Context</b>	<b>Inner Context</b>	<b>Outcome</b>
Sharing and learning	External peer review	Any length of membership	Short-term and long-term changes
Consultation	Peer review network	Any length of membership	Long-term change
Ownership and delegation	Accreditation	Any length of membership	Short-term and long-term changes
(Lack of) communication	Peer review network	Any length of membership	No change
Action planning	External peer review	Any length of membership	Short-term and long-term changes
Senior management / junior involvement balance	External peer review	Any length of membership	Long-term change

Sharing and learning was considered the most important casual mechanism of change, in both peer review networks and accreditation schemes, irrespective of length of membership. Initially mentioned in focus groups with practitioners, this was further developed through interviews with subjects. It was found to lead to short-term changes when materials were borrowed or shared, but could also lead to long-term changes when relationships between peers who had met through external peer review were maintained and developed to continue to share experiences and resources. In some instances, learning was extended to reciprocal visits that were set up outside of formal external peer review to further improvements in quality. This can also be seen in the extant literature where staff have been stimulated to collaborate with different sites following participation in accreditation ([Touati and Pomey, 2009](#)).

Consultation was a salient mechanism that had been reported by subjects participating in peer review networks from both new and long-standing services that led to long-term changes. These changes were reported to have been brought about through consultation in self-review, and in implementing change. In the case of the latter, this was a powerful mechanism as consultation of frontline staff helped to reduce resistance to change.

Ownership and delegation were key mechanisms that had been reported by subjects participating in accreditation schemes from both new and long-standing services that led to both short-term and long-term changes. Short-term changes such as achievement and advertisement were reported in relation to these mechanisms, and long-term changes such as a drive to continue quality improvement outside of the accreditation standards were also reported.

Communication was a key mechanism to facilitate success for those participating in peer review networks, in both new and long-standing services. Initially perceived as an aspect of readiness for change by practitioners, many frontline subjects requested opportunities for improved communication to help their understanding (instrumental to readiness for change), and thus secure their future engagement and involvement in the programmes. When

communication was reported to have lacked, this coincided with reports of lack of engagement and understanding, which may explain why more frontline subjects than any other staff groups witnessed no changes following participation in a peer review network.

The use of structured processes, especially action planning was seen to facilitate achieving success in external peer review, irrespective of length of membership. Action planning, or similar structured processes such as grading standards based on a traffic light system can facilitate these stages was viewed as a causal mechanism that facilitated implementation of short-term and long-term changes.

The balance of senior management presence and junior involvement presence was identified as a key causal mechanism in external peer review, irrespective of length of membership to bring about long-term changes and commitment to quality improvement. Leadership and senior management presence was also an important construct of readiness for change, initially raised in focus groups by practitioners, and further developed by subjects. It was found to be linked to power for negotiating resources necessary to implement change, articulating and communicating a shared vision as well as acting as a change champion in some cases. However, striking the correct balance seemed to be important as practitioners expressed if there was too much involvement from a senior manager during the early stages of the programme, it would decrease the capacity of the service to benefit from the programme, as frontline staff might be less engaged. For this reason, practitioners outlined that they preferred to assign the role of a lead contact to a ward manager or charge nurse as oppose to a senior manager. Junior involvement was seen to be important as most of the changes were reported to be implemented by junior staff, and they also spent the most time with service users.

### **13.3 Aim 1: To explore stakeholder views of the causal mechanisms of external peer review programmes**

Most of the changes in accreditation schemes were reported to have occurred before the visit: before joining or during self-review. This was also previously identified in the extant literature where the preparatory period before the visit comprised of the most work and change ([Duckett, 1983](#)). The most important changes implemented during the accreditation cycle had been identified during self-assessment ([Pomey et al., 2010](#)), or in preparation for accreditation ([Pomey et al., 2004](#)). Practitioners reported if most changes were made during these early stages, it is difficult for the CCQI to measure these changes.

Most of the change in peer review networks were reported to take place following the written feedback stage, and there was frequent reference to using this programme factor as a tool to negotiate for more resources. This was observed in previous studies of peer review networks ([Roberts et al., 2010](#)) ([Page and Harrison, 1995](#)) as well as accreditation ([Baskind et al., 2010](#)) ([Duckett, 1983](#)).

As part of the CCQI process, practitioners outlined how member services are required to select staff to volunteer as peer reviewers during the joining process; however, there is no formal process to enforce these volunteers to undertake reviews once a service has joined. From the qualitative data, the personalities and behaviours of reviewers is a substantial programme factor. This was also found across the extant literature ([Stebbing, 2011](#)) ([Greenfield et al., 2010](#)). However, there is currently no screening or revalidation process to ensure that peer reviewers have the qualities deemed necessary by the qualitative findings and extant literature. According to both practitioners and subjects, reprimanding 'bad reviewers' is solely down to other peer reviewers if they feel confident enough to speak up, or a mediator from the CCQI for peer review network visits. Furthermore, practitioners and subjects both noted that peer reviewer opportunities were mostly limited to more senior staff.



There are three main concepts that characterise how changes occur in external peer review programmes: **internal** causal mechanisms inside the host service, **benchmarking** undertaken by the host service internally and with peer reviewers externally, and **external** causal mechanisms because of peer interaction. Causal mechanisms undertaken internally by a host organisation, can include consultation, ownership, delegation, communication and motivation. These were reported to be undertaken prior to joining, during the joining stage, during the self-review stage or before the peer review visit takes place.

Benchmarking overlapped between internal and external concepts, as it was carried out both internally by a host service during self-review, as well as externally with peer reviewers during visits. This concept involved other causal mechanisms such as standards and reassurance (which was also closely linked to peer endorsement). Mostly for accreditation, and in some examples of peer review networks, evidencing was used as a causal mechanism to highlight that standards had been met.

The external concept was related to interactions with peers, which is where sharing and learning were reported to have occurred. Both practitioners and subjects frequently mentioned the conversations and discussions that came about during the visit as an integral part of programmes, thus this mechanism was the most essential causal mechanism of external peer review. Through learning and sharing, services who participated in external peer review reported to have gained tacit knowledge (based on experience of services) of implementation methods. This builds upon evidence from NCROP ([Rivas et al., 2008](#)) ([Roberts et al., 2010](#)) that external peer review provides a breeding ground for an exchange of ideas, a validating and reassuring experience, and a networking opportunity in which mentoring relationships can be formed. Previous literature suggested that external peer review programmes can create the opportunity for a learning environment where new members can learn from long-standing members, share resources, methods of best practice and implementation of standards with one another. This was confirmed by CMOC

development in this thesis. This could also take place during other programme factors such as peer reviewing other services and additional activities offered by the CCQI, which were also described in the literature ([Worrall, 2011](#)).

Evidence-based standards of care provided services with explicit knowledge of what to aim for in relation to service delivery ([Butterfield et al., 2012](#)) ([Touati and Pomey, 2009](#)). However, qualitative findings from practitioners and subjects indicated that standards were often open to interpretation, and sometimes did not fit the local context. By learning and sharing, especially through the feedback and peer interactions, subjects could develop tacit knowledge of how to implement standards. The way that peer reviewer recommendations were seen to drive changes in external peer review was also dominant in the previous literature ([Roberts et al., 2010](#)).

Peer endorsement was viewed a vital causal mechanism by practitioners and subjects, as it was felt that recognition and validation from peers held a lot more weight and importance than other regulatory bodies. Subjects mentioned this was the main difference between CCQI programmes and other quality improvement visits such as the CQC, as inspectors did not have the same background or expertise as peers. Gaining peer endorsement was also seen to reduce the impact of certain contexts such as geographical or clinical isolation, economy, healthcare system, lack of non-financial resources and local context, as reassurance could be sought from both peers and standards. According to quality improvement expert Grol ([1994](#)) participants who may have worked for many years in isolation and without feedback on their style of working will each have developed their own style of working and own guidelines for good practice; thus could benefit from learning, sharing and peer endorsement through external peer review.

Peers also included service user representatives which were a crucial aspect of the peer review team by practitioners and subjects, as they offered additional insights. Practitioners highlighted this, and could provide background information as to why there are challenges in involving service user reviewers.

Peer reviewer issues, such as competency and consistency between the peer review site visits was one the most frequently cited, from the literature, practitioners and subjects.

Participants from the NCROP study reported that service user representatives demonstrated loyalties to their own service during peer review visits by over-commenting on their own service instead of the host service, or were seen to use the visit as an opportunity to find out more about what was available elsewhere ([Rivas et al., 2008](#)). However, this was not present in other studies ([O'Connor et al., 2007](#)), or the qualitative findings; instead, both practitioners and subjects considered the views service user peer reviewers as instrumental in bringing about change.

The research raised an interesting conflict between identifying senior management support as a necessary causal mechanism for change, as well as junior staff involvement as another, separate, causal mechanism for change. The qualitative findings revealed many junior staff would like to increase their involvement and understanding of the programmes, with some senior manager subjects in support of this. Amongst the suggestions by subjects for improvements of the programmes, having more junior staff such as staff nurses and HCAs involved, was the second most mentioned. This indicated all staff levels perceived junior involvement as a substantial priority. From an organisational perspective, as senior management support is instrumental in bringing about change, it makes sense for senior managers to currently be the most involved in external peer review, especially reviewing other services. However, many subjects felt that by limiting involvement to only senior management, programmes were missing the ideas and involvement of not only frontline staff, but because of this staff group spending the most time with service users, the service user voice as well. Management support was causal mechanism, and also a key indicator frequently mentioned by many subjects and present in the literature ([Pomey et al., 2004](#)), as it was felt that without management support very little change could be made. According to Devers et al., ([2004](#)) management and clinical leadership were substantial for external

peer review as resources and cooperation were more likely to follow. For staff working in peer review networks, management support was felt to be instrumental in converting recommendations from the written report into quality improvement changes, in addition, management support from accreditation was felt to be instrumental to achieving accreditation. This careful dynamic needs to be managed by sustaining engagement and relevance for both junior and senior staff, increasing communication and reiterating the benefits of participation to ensure success in external peer review.

Some of the previous literature suggested a lack of clinical involvement and tension between clinical and non-clinical staff was a barrier to success of external peer review. However, this concept was only slightly mentioned by practitioners or subjects. This may be due to the research context of inpatient mental health settings, in contrast to previous research which was mainly conducted in hospital settings. It is possible that clinical involvement is more integrated into these services compared to other health settings.

Subjects and practitioners expressed a broad range of views about changes as a result of external peer review programmes, from those that were convinced they had led to substantial changes in service delivery, to those that felt they had had limited or no impact. Mostly frontline staff reported no changes were attributed to participation of peer review networks. If frontline staff did not notice any changes and they spend the most time with service users, this calls into question the benefits which service users receive as a result of their services participating in external peer review. Similar findings were reported in the qualitative sub-study by Rivas et al. ([2012](#)). However, although staff reported that 'no changes' occurred as a result of participating in the peer review network, later 'generic changes' emerged through the data although staff did not initially classify them as changes. Changes that are not directly attributed to external peer review could also be due to many other causes. Maturation, whereby natural development could lead to quality improvement. Repeated testing can have an effect where knowledge or behaviour appears to improve due to familiarity with procedures. In addition, in many cases, services

in voluntary programmes might be expected to be more highly motivated to make changes in their behaviours.

One of the long-term changes due to external peer review was the impact that membership had on the ability of the service to promote other quality improvement programmes such as local audits and service evaluations. Continuing from the momentum of external peer review, subjects expressed the benefits they had found with continuing the audit and review activities in the interim periods in-between cycles.

### **13.4 Aim 2: To identify contexts which influence the ability of services to make use of their participation in external peer review programmes**

Contexts influencing the impact of membership of an external peer review programme included those operating at a national (outer) and organisational (inner) level. As expected, the types of programme (accreditation or peer review network) and healthcare system were the most influential at the outer-level. Length of membership was the most salient inner context.

#### **13.4.1 Outer contexts**

The most common outer contexts perceived to influence external peer review in this study were the type of programme and the healthcare system.

The main differences practitioners and subjects outlined between peer review networks and accreditation was the obligation to make changes. Whether external peer review programmes were mandatory or not did not have notable influence on their impact according to practitioners and subjects, although in some instances subjects from long-standing MSUs expressed their concern that whilst the medium secure QNFMHS was mandatory, it had not changed much since they had joined. When most subjects participating in QNFMHS mentioned the links to the CQC, Department of Health and other regulatory bodies, they spoke in fear of the programmes becoming too interlinked with these organisations. They expressed their concerns that if the external peer review process and outcome measures became too closely linked with regulation, it would reduce the learning and sharing mechanisms they valued, that were considered an essential causal mechanism for implementing change.

There was great debate about public and private provision of services, and the impact this may have on resources, and subsequent capacity to benefit. However, this was in contrast to evidence from previous literature that found no significant difference between Australian public and private organisations' accreditation ratings ([Braithwaite et al., 2010](#)).

Some senior managers had worked in both private and public services, and could contrast their services' abilities to implement changes. They suggested it was generally easier to make changes resulting in quality improvement in private services due to reduced bureaucracy. These reports were echoed by some practitioners. However, there were contrasting views from practitioners regarding this.

#### **13.4.2 Inner contexts**

These contexts were at the organisation level (both the CCQI as an organisation coordinating the external peer review programmes, and the member wards and services). Findings were centred on the main inner contexts: length of membership and nonfinancial resources.

The most frequently mentioned inner non-programme (contextual) factors was the length of membership, which was signalled by Roberts et al. ([2012](#)) that 1 or 3 years was not enough time to observe changes leading to quality improvement through external peer review. The qualitative data from this thesis demonstrates that some subjects feel that their understanding of the programmes has increased and apprehensiveness has decreased the longer their length of membership, this may mean the observed differences may take longer. This is congruent with a previous mixed methods study identified in the systematic literature review which suggested a climate of trust had to take root to maintain a capacity for reflection and bring about changes through accreditation ([Pomey et al., 2004](#)). Some subjects experienced benefits from repeated exposure as they became more comfortable with programmes, which in turn facilitated their success. However, some subjects from long-standing MSUs that had joined the QNFMHS early on felt the peer review network had become less beneficial for them. This suggests a possible plateau effect if external peer review programmes remain unchanged.

### **13.5 Aim 3: To assess whether 'readiness for change' influences the ability of services to make effective use of an external peer review programme**

This research highlights the importance of the readiness for change assessment, and despite the non-significant relationship found in the quantitative studies, several key constructs of readiness for change were identified by practitioners and subjects, providing a new body of evidence in this area.

Lack of awareness and understanding were common barriers to change mentioned by practitioners and subjects, especially in relation to frontline subjects. This was attributed to 'information not always being filtered down'. Several frontline subjects detailed how they did not receive feedback or information from peer review visits if they were not directly present on the day, and hardly ever from colleagues who had visited other services, unless this was communicated informally. Many frontline subjects requested opportunities during the interviews for improved communication (a substantial team process causal mechanism, and perceived readiness for change construct by practitioners) to help their understanding, and thus secure their future engagement and involvement in the programmes. This lack of engagement, understanding and communication may explain why more frontline subjects than any other groups saw no changes following external peer review. This is echoed by a case-study of accreditation conducted by Paccioni et al. ([2007](#)) in Canada which indicated that accreditation has little effect on the perceptions of those employees who are not directly involved in the programmes.

According to previous literature, I hypothesised that greater readiness for change in teams would lead to less resistance to change making it easier for services to implement change ([Weiner, 2009](#)). Qualitative data from practitioners and subjects with services providers highlighted the importance of readiness for change, especially the constructs of being connected to other units, being open and having a shared leadership vision. However, data from the quantitative analysis of memory clinics and low secure units were inconclusive. Being 'connected to other units' was the most substantial readiness for change



construct, which was not mentioned previously in the literature. It was also strongly associated with sharing and learning, which was the most substantial causal mechanism. By feeling connected to other units, learning and sharing through networking was reported to have resulted in belonging to a community, which many subjects did not have before joining external peer review. Through their membership, albeit some reported being more 'connected' than others by taking advantage of the opportunity to peer review other services and the additional services that the CCQI offers, subsequent changes were attributed to participation. However, the meaning of the community differed between peer review networks and accreditation programmes. Many peer review networks reported to have used the community to learn from one another and as an opportunity to benchmark themselves; whereas units taking part in accreditation saw the community as an 'exclusive club' which they wanted to be a part of that validated that their PICU had met high standards of care. Some subjects described wanting to enter (i.e. certificate) this 'exclusive club', and others demonstrated more of a drive to continually adhere to standards once they were part of the club, and possibly continue improvement.

Interviews with subjects enabled me to understand that readiness for change was not a precondition needed before embarking on external peer review. Practitioners and subjects suggested that by being open to change and welcoming external assessment from peers, services were less likely to be apprehensive before the visit as they had adopted a culture of openness and transparency. Through participation in external peer review programmes, some subjects reported that services that were already connected with people from other services were more likely to demonstrate greater readiness for change, and can work towards continued quality improvement instead of one-off short-term changes.

Having a shared leadership vision was reported as a substantial readiness for change construct by practitioners initially, this was further developed by subjects. However, it was not significantly associated with greater quality improvement when tested in the cross-sectional memory clinics audit dataset

or the longitudinal analysis of membership of the low secure network. Having a shared broad leadership vision of change was also considered to be an substantial aspect of the process of change by three managers of the NCROP qualitative sub-study ([Rivas et al., 2012](#)). Although this study took place in the context of UK respiratory services, it provides the most recent body of evidence of change in external peer review, and is the largest randomised controlled trial of external peer review to date. According to qualitative findings from NCROP, shared leadership vision, when communicated correctly can steer the team towards change and inspire ownership and motivation, which are substantial causal mechanisms of change. These findings were congruent with the results of this thesis.

Tenure and job role were interesting themes that were only slightly mentioned in the data, but had been referenced in extant literature. According to Duckett ([1983](#)), resistance from staff, especially amongst older staff and physicians, was perceived as a major challenge to the success of external peer review. In a similar thread, there was a sense in the findings of this thesis that younger or newer staff groups, or new management were more open to change, seeing change as possible or adapting, which are key constructs of readiness for change. When mentioning the theme of resistance to change, subjects also frequently mentioned staff who had worked in services for a long time.

When discussing indicators of capacity to benefit from membership of an external peer review programme, practitioners and subjects stated how necessary it was to have senior manager present on the unit, or attendance from senior management at key meetings or feedback sessions was an essential indicator of capacity to benefit from the programmes. This was echoed in the literature of accreditation of Endoscopy Units in the UK. In cases where senior management are not present, it is also considered an indication of poor organisational culture and support for the endoscopy service ([Stebbing, 2011](#)).

## **13.6 Methodological Considerations**

Before considering the implications of the findings of this research, I would like to highlight some of the methodological issues that are both substantive and relevant across this thesis. I will start by discussing the qualitative components of the study, followed by the quantitative components and finally examine the generalisability of the research.

### **13.6.1 Qualitative components of the study**

#### ***13.6.1.1 Strengths***

The range of groups I sampled across the UK, in both accreditation and peer review networks, with differing lengths of membership helped to provide further dimension to the overall thesis. A novel aspect of the qualitative studies was the use of stakeholders across different professional groups, including practitioners who coordinated external peer review at the CCQI; and senior managers, ward managers and frontline staff (subjects) who worked in services participating external peer review. Through purposive sampling, a good coverage of settings was achieved from the 36 different services I sampled from.

I conducted three pilot interviews between February and March 2014 with each type of service (low secure unit, medium secure unit and a PICU). This enabled me to test out the suitability of the topic guide, and the appropriateness of the timeframe allocated for interviews. I conducted the remainder of the interviews in a three-month period. I reached data saturation after undertaking 121 interviews, which also enabled me to capture data from dissenters.

The topic guides I formulated to use in focus group and semi-structured interview studies were grounded in the evidence from a robust systematic literature review. I used a developmental approach, whereby the findings of focus groups informed the topic guide for semi-structured interviews. In this way, CMOCs generated by staff coordinating external peer review programmes were tested through interviews with staff participating in external peer review programmes.

### **13.6.1.2 Limitations**

Despite the advantages of using focus groups, there were also disadvantages, which I took under consideration ([Oppenheim, 1992](#)). Initially, I found it challenging to find suitable dates as key CCQI staff (practitioners) are often out of the office to attend peer review visits. This may have reduced the number of practitioners who attended. I was challenged by time limitations as the focus groups took place during the lunch hour. This restricted the number of questions or topics that could be discussed. My facilitation style and approach potentially may have limited some practitioners' responses and engagement. One or more strong individuals might have tried to dominate the group, and this could have led to extreme views being debated, or group norms silencing some practitioners. There may have been personality conflicts and power struggles within the group which affected results. The groups' composition in terms of age, gender and culture may have affected the themes generated. A possible limitation is that there was only one male practitioner, so the views of this gender may have been underrepresented. Confidentiality concerns may have led to some practitioners not being completely open or fully engaged.

During analysis, I recognised that focus groups can overemphasise consensus and be dominated by either influential or opinionated participants. Thus, it could have been problematic to generalise from focus groups as practitioners were selected through non-random sampling which could have promoted more self-confident and articulate individuals to take part ([Morgan, 1997](#)). I considered the alternative approaches of using online questionnaires or undertaking semi-structured interviews with individual practitioners. However, the advantages of both approaches outweighed the concerns. As Walliman ([2005](#)) suggested, questionnaires would have only provided one perspective, would have required a longer period of time, and would have lacked the benefits provided by group interaction ([Oppenheim, 1992](#)). Considering these limitations, I decided focus groups were an appropriate method to use for eliciting practitioners' perspectives.

As there was a limited time and scope to conduct this study, four focus groups and one interview may not have been adequate to represent the views of this group of stakeholders. However, by the last focus groups I experienced a saturation of several key themes.

It could be a limitation to only gather evidence consistent with one theory, and analysing only in ways consistent with that theory. This prompted me to iteratively consider exceptions to patterns and other possible explanations during my analysis of qualitative data.

As lack of standardisation is inherent in semi-structured interviews, there is scope for concern to satisfy reliability to achieve similar results via another researcher. Saunders et al. ([2016](#)) suggests that these results are not necessarily intended to be repeatable, as they reflect the dynamic and complex reality of the situation at the point at which the interview was conducted. The extent to which bias is introduced into the interview may also present a threat to reliability ([Saunders et al., 2016](#)). In relation to any bias arising from my behaviour and conduct as an interviewer, I sought to negate this risk by seeking to present myself in a professional and credible manner.

I have been singularly responsible for undertaking all aspects of the qualitative studies and, as such, have been subject to the limitations of time and personal resources. Without this constraint, or alternatively by working in collaboration with other researchers, there may have been scope to collect further data, in particular, opportunities to sample from more external peer review programmes.

Not having observation data from focus groups or semi-structured interviews is an additional limitation, as this may have added to the triangulation of findings and provided a different perspective.

I was not able to link the data from semi-structured interviews to how services complied with standards during their respective external peer review

programme cycles. This was a limitation, as it would have enabled me to link mechanisms and contextual factors with reported outcomes, however this was not possible due to confidentiality agreements with the services (details of peer review network scores are only released to host services).

The question of what constitutes a mechanism remained a significant challenge ([Astbury and Leeuw, 2010](#)). There was also the issue of context. ([Barnes et al., 2003](#)) warn about the risk of interpreting context as a purely external; they argue that in open systems, context is much shaped by the actors as it constrains their activities. Byng et al. ([2008](#)) point to the possibility of multiple mechanisms acting at the same time, which again they claim is stressed by Bhaskar and ignored by Pawson and Tilley.

The face-to-face method of data collection was appropriate for gaining insight into programmes, but it was also open to social desirability and acquiescence bias. Some staff with implicit knowledge that their practices were being looked at could be likely to respond to questions in ways that suggested that they follow the process in a certain way. Although this threatens the validity of the CMOCs, findings of realist evaluation can help trigger new studies and provide useful transferable lessons for stakeholders interested in implementing similar programmes elsewhere ([Doi et al., 2015](#)), ([Pawson and Tilley, 1997](#)).

In conducting semi-structured interviews, there is an inherent threat of respondent bias. Despite my best efforts to both prepare for, and manage the interview process effectively, there remains the possibility that the responses provided an inaccurate reflection of subjects' experiences of the process of change in external peer review.

The subjects were pre-selected before I arrived so they could have been selected for their knowledge or involvement, or in some cases could have been primed prior to interviews. There was evidence of this in an interview with a frontline subject where I was explicitly informed that their knowledge and understanding stemmed from conversations with their line manager

immediately before the interview. I was only able to interview the services which were willing and responded to my initial contact e-mail, this could have introduced bias into the study and could have missed out services which had particularly negative experiences, or encountered difficulties in their external peer review programme.

Job titles differed between services, and given the variations in size, I found that different titles were sometimes assigned to staff with equivalent responsibilities. If I were to repeat the study, I would ensure I provided more explicit role descriptions of the categories of senior managers, ward managers and frontline staff when making initial contact with services. It is noteworthy that during the period I conducted interviews, the medium secure standards were being redeveloped in consultation with services. Some of the senior management subjects were directly involved in this, which may have affected their responses, especially when referring to standards. I was unable to seek further clarification in relation to the additional activities offered by the CCQI, such as the online forum, due to confidentiality limitations.

Time was a constraint, especially as the interviews were conducted whilst the participants were at work. As they often took place within the service environment, this could have increased anxiety or added pressure to their responses.

There were inherent constraints in sampling from existing peer review networks and accreditation programmes by having to agree systems for data collection that were perceived to not pose adverse impact on the programmes.

Whilst the work presented in this Thesis has focused on the process of change in external peer review, it is imperative not to lose sight of the fact that service users are at the heart of the healthcare system. None of the approaches used directly involved the views of service users and their families, essential to assessments of quality of care, this may have added a different dimension to the findings. There was no inclusion of the service user voice in focus groups or

interviews, as there was no literature to support that inpatient mental health service users have the appropriate understanding of the inner contexts (readiness for change constructs, programme factors, causal mechanisms, change outcomes) which influence the process of change.

Finally, this research has risked exposure to the preconceptions, personal and professional values and potential biases that I might have brought to the process. While every effort was made to negate these through the careful and systematic planning and execution of this study, it is appropriate to acknowledge that these in themselves represent a limitation to the qualitative components of this research.

### **13.6.2 Quantitative components of the study**

#### ***13.6.2.1 Strengths***

This analysis of readiness for change was the first assessment of this kind in external peer review programmes (quality networks and accreditation). As it was carried out in services that participated in the evaluation of low secure units' (eLSU) study before they had entered the programme, it is vital in shaping the peer review network over the coming years.

I piloted the readiness for change checklist in one LSU, which enabled myself and another researcher administering the tool to receive some feedback on the ease of completion and appropriateness of questions. It was found to be acceptable and easy to complete. I sought permission from the authors of the tool, to adapt the wording to increase suitability for a low secure forensic mental health environment. The multi-level model also adds a substantial dimension that would have been overlooked in a single-level approach.

#### ***13.6.2.2 Limitations***

Weiner ([2009](#)) considers 'organisational readiness for change' a critical precursor to the successful implementation of complex changes in healthcare settings, whilst arguing the concept has not been subject to extensive empirical



study, which meant limited guidance was available in the literature for me to consult when developing statistical models.

Self-administered questionnaire errors can be caused by many things: misunderstanding the question, not having the information needed to answer, and distorting answers to look good, which are only a few examples ([Fowler, 2013](#)). There is also no opportunity to probe respondents to elaborate an answer ([Bryman, 2012](#)). In addition, respondents can read all questions before answering the first question. When this occurs, none of the questions asked is truly independent of the others. This could have been mitigated against by using a web survey, where it is possible to ensure that respondents can view only a small number of questions at a time. However, resources during this study did not permit this. Partially answered questionnaires were more likely, because of a lack of supervision or prompting. It was also easier for respondents to actively decide not to answer a question when on their own compared to when being asked by an interviewer ([Bryman, 2012](#)). I took this into consideration when distributing the tool. Data can be affected by the characteristics of the respondents: some may not have taken the questionnaire seriously while others may not have accurately reported their beliefs and attitudes ([Boynton and Greenhalgh, 2004](#), [Boynton, 2004](#)). Questions could create ambiguity if they ask about more than one issue within the same phrase. Structured questionnaires may not be sufficiently comprehensive and not all answers may be easily accommodated. Some respondents could have been forced to choose inappropriate pre-coded answers that did not fully represent their views ([Bowling, 2014](#)).

As the National Audit of Memory Clinics had a constraint on the number of questions that could be included in the questionnaire, only five constructs could be selected from the 25-question tool ([Bobiak et al., 2009](#)). Based on the preliminary literature, five constructs of readiness for change were selected, but perhaps they were not the most sensitive of readiness to change, as only one of the constructs later emerged as salient in the qualitative findings. This could explain why the relationships observed in this study were not statistically

significant. Most of the questions in the National Memory Clinics Audit were mandatory (apart from the readiness for change items) and responders were obliged to enter a response before they could complete the questionnaires. One of the limitations this caused, was when responders did not want to enter the information required for the other audit questions they could enter '0', resulting in some instances of confusion for which responses were truly zero and which were incomplete responses. This could have impacted the primary and secondary outcomes of the National Audit. Some data was subsequently removed from the audit dataset where no clarification was given. As the data were routinely collected as part of a cross-sectional dataset, it was restrictive to the point in time the data were collected, there was inherent response bias, and I had no control over targeting specific respondents. The other questions in the dataset may have also influenced how the readiness for change constructs were answered, as these were the only Likert-scale options, and the results were skewed towards extreme responses.

Furthermore, there was no previous evidence of only using five of the constructs from the 25-question tool, so conducting analysis with only five constructs may have resulted in an incomplete assessment of readiness for change for the memory clinics.

Bryman ([2012](#)) suggested that you can never be sure whether the right person has answered questionnaires. As the questionnaires were intended for senior managers, it is possible that the task may have been delegated to someone else which could have affected the results.

There were some limitations of the 25-question readiness for change tool. Outcomes that may be linked to readiness for change may not have been achieved. As it was embedded as part of a randomised controlled evaluation, there was a limit on how many questions could be asked, which restricted the choice of validated tools that could be selected. The context of the ongoing evaluation, the time pressures faced during data collected, and only having data from one-year of follow up to assess change may have also resulted in

limitations. As indicated by qualitative results, the length of membership was a substantial inner context. As the data were collected without knowing respective allocation status to intervention groups, this led to uneven distributions between readiness for change scores between active and control arms of the evaluation. If I had been able to pre-select sites based on their organisational features, similar to the NCROP study ([Rivas et al., 2010](#)), this could have resulted in a more even distribution of longitudinal data. Also, the tool may not have been sensitive enough in the forensic mental health context.

There was missing data when assessing change from baseline QELS checklist scores as 3 wards had dropped out of the eLSU study by this point, and this may have affected the results, although it is unlikely.

A methodological limitation of the eLSU study was that there were not any accompanying notes during collecting data for the checklist. For further studies in this area, to ensure standardisation it could be good practice to include a guidance document to assist completion of the checklist, or perhaps a manual.

The analysis of multilevel data has been said to be problematic as those belonging to the same group are more likely to be similar than those from different groups. One disadvantage of SAS PROC MIXED is that it does not provide the specific efficiency for the nested random coefficients of the hierarchical linear model that is provided by dedicated multilevel programmes. I used model building, stepwise selection using AIC, for multiple imputation. The alternate, using as many predictors as possible may have been superior.

The threshold ceiling effect was a limitation of having multiple variables that were mostly scored on extremes of the Likert scales (strongly agree and strongly disagree). This may have posed limitations on effective measurement of the true effect of highly scoring respondents.

The sample size of the studies presented in this thesis are varied. In total, there were 178 MSNAP audit respondents, and 101 respondents for the readiness for

change checklist. Some may argue that the sample sizes for the quantitative studies are small, however when considering triangulation with the qualitative components, the study samples represent a reasonable proportion for generalisations to be made. This limitation highlights the resource and time commitments, which are inherent in the context of a thesis.

### **13.6.3 Generalisability to other programmes**

The research studies were conducted in the context of community and inpatient mental health settings in the UK regarding external peer review programmes delivered by the CCQI. Although this creates limitations concerning generalisability of my findings, there are also opportunities.

A key issue debated in the literature has always been whether an intervention such as external peer review which works in one setting can be transferred to another ([Øvretveit, 1997b](#)). Previous research did not provide sufficiently detailed descriptions of their context for transferability or translatability to be judged ([Øvretveit, 2002a](#)). Therefore, I sampled from both public and privately provided services in all my studies, which enables the findings to be generalised to both public and private healthcare contexts. The refined CMOCs of how change is achieved in external peer review through internal and external causal mechanisms can be applied to other external peer review programmes through theoretical generalisability ([Shaw, 2000](#)), as the programme factors for external peer review will be similar, both in the UK and globally. Inner contexts that influence teams and organisations are likely to be similar, as the extant literature indicates substantial inner contexts such as readiness for change, length of membership, non-financial resources, and organisational culture are salient in settings outside of inpatient mental health.

Peer reviewing other services was the most valuable programme factor of external peer review according to subjects, especially in peer review networks. This could be due to the geographical and clinical isolation faced by services, where it is relatively uncommon to visit outside services. This value may be found elsewhere, as it is a pertinent issue faced across multiple specialties in healthcare, exacerbated by geographical distances in the UK, which are also prominent in some other parts of the world.

## **13.7 Recommendations**

Several recommendations for practice, policy and further research have emerged from this study.

### **13.7.1 Directions for further research**

From a realist perspective, single evaluations cannot produce universally valid findings. Realist evaluation can help researchers find out in which specific conditions the programme works (or not) and how, and to refine the findings in a process of specification. This in turn leads to an accumulation of insights that help decision makers assess whether programmes that proved successful in one setting may be so (or not) in another setting and how ([Pawson and Tilley, 1997](#)), ([Marchal et al., 2012](#)). Thus, the findings from this thesis provide a signal that the context of readiness for change can be important to determining success or capacity to benefit through participation in external peer review. This calls for a larger scale examination of how organisational readiness for change influences the receptive context ([Pettigrew, 1992](#)) of external peer review programmes, with an increased sample size to mitigate against inherent limitations in study power. This research could include better matched sample populations, so that both qualitative and quantitative data could be collected from the same participants.

An under-researched area that should be focused upon in future research is the role that inpatients could potentially play in shaping and improving the quality of care they receive through external peer review. If the research is conducted in inpatient mental health, where there is a high turnaround of service users on wards such as PICUs, a tool which aggregates service users with similar symptoms such as 'Patients Like Me' could be used, or perhaps medication levels or activity uptake could be used as proxy indicators to ascertain rates of patient wellbeing in relation to recommendations made and changes implemented through external peer review. As service user representatives are already used in some external peer review programmes, this provides an additional facet that could be evaluated.

In a study of *visitatie* in the Netherlands, additional tailored quality improvement implementation support was offered to services participating in a peer review network ([Lombarts and Klazinga, 2003](#)). Present in the extant literature, a prominent suggestion for improvements raised by subjects was more support with implementation, further research could be conducted on assessing the impact and effectiveness of tailored or additional support in improving quality through external peer review.

Currently, there is a great amount of focus on process standards in external peer review, especially at the CCQI. When practitioners and subjects were discussing the appropriateness of standards, suggestions for more outcome standards were raised by both groups. This suggests an economic evaluation of using more outcome standards in external peer review could be useful, so costs can be measured, and the potential additional value to service users could be predicted, which could provide an evidence-base for increased use of outcome measures.

As ontology in this field was not interoperable, further attempts should be made to consult with the quality improvement research community to establish a common ontology framework to facilitate open data sharing and stewardship to continue research developments in the field of external peer review.

### **13.7.2 Implications for clinical practice**

The research findings suggest improvements can be made to the existing peer review networks and accreditation programmes at the CCQI: for those that commission and coordinate external peer review (coordinating bodies), and those that participate in external peer review (services). In relation to practice, these include a few specific actions and activities that may contribute to enhance the effectiveness of external peer review implementation, while those aimed at policy address external peer review in the wider context of the healthcare in the UK and beyond.

As such this study recommends:

## Coordinating Bodies

A balance must be obtained between the ambition of a perfect robust external peer review programme, and what is feasible and achievable. Although services may currently be working towards short-term changes to pass accreditation or undertake peer review visits as part of peer review networks; instead of investing in more measures of assessment of change; devising approaches to increase intrinsic motivation to work towards long-term changes would be a better use of resources. Equipping services with robust quality improvement methodology support from the IHI ([Davis et al., 2011](#)) or online resources such as QI4U ([CLAHRC NWL, 2016](#)) would support the training and understanding of quality improvement in host services. This would equip frontline staff and managers with further skills necessary to undertake and implement changes to improve the delivery of their services.

Evidence of regional networks have already been reported by subjects in the North Yorkshire & Humber Region of the UK. This could be formalised to create local QICs ([Institute for Healthcare Improvement, 2003](#)), whereby services who want to continue to improve outside of external peer review, can work towards improving a quality measure for a set period of time, use dashboards to compete and compare their improvement data and share best practice to collectively improve their quality in a single area ([Bate et al., 2002](#)).

Creating a change platform that takes advantage of social technologies that make large-scale collaboration easy and effective has the potential to drive deep change in external peer review ([Hamel and Zanini, 2014](#)). This could help to enhance the sense of community experienced through participation whilst also increasing access to learning and sharing where there are currently barriers to access (i.e. time, geographical location) ([Bevan, 2016](#)). If these opportunities are capitalised and developed, both newly joined member services who are not as familiar or comfortable with the programmes, or quite sure of how to use participation to make changes and longer-standing services who have found the programmes to reduce in effectiveness after many years of membership will be



able to reap enhanced benefits from external peer review through engaging with peers and collectively solving problems through virtual change platforms.

Being apprehensive before visits was a substantial finding in the qualitative interviews. Upon further analysis, this was linked with a feeling of fearing the initial unfamiliarity of the programmes, as many found it was difficult to get their head around at first and some also did not understand the meaning or value of the programmes until after they had completed their first cycle. This could suggest a preparatory visit, similar to the education surveys conducted in the Zambian accreditation programme to increase familiarity ([Bukonda et al., 2002](#)). The qualitative data suggested that greater the length of participation in external peer review, led to participants having greater understanding of the meaning and value of membership, and some stated this led to them becoming more comfortable as their sense of familiarity increased. It was unclear if the feelings of apprehension affected participants' readiness for change, but it was suggested that these feelings acted as a barrier to change. This could be tackled through training, tailored support before the visit, or increased peer support between new and long-standing members who could provide reassurance help bridge the gaps of knowledge and understanding.

## **Services**

The challenges of bringing about improvement in quality in real-world clinical environments should not be underestimated, as has been reported by work supported by the Health Foundation ([2013](#)) ([Gabbay et al., 2014](#)). These challenges include finding the time for both managers and clinical staff to review working practices and to test and implement improvements as well as providing these staff with training in methods and measurement. There was a wide range of changes witnessed by subjects, from short-term to long-term success. The institutionalisation of quality improvement requires a long-term commitment and engagement, which can be brought about by causal mechanisms such as increased understanding, shared leadership vision, strong senior management, junior management involvement and engagement and the

use of structured processes to implement and work towards changes. If used to the best of their ability, services have reported that these commitments and use of causal mechanisms can result in an embedded culture of quality improvement within their wards and services. If services can invest in and promote these causal mechanisms and help to foster a culture of continued long-term commitment it may help them achieve long-term success through sustainability of the changes made.

By involving more junior staff in external peer review, organisations can use external peer review as an opportunity to develop and invest in their junior staff. From an organisational perspective, this would need to be linked to training and could help staff retention issues which are similarly faced by inpatient mental health services across the UK.

Upon joining a peer review networks and accreditation programmes, if baseline readiness for change is assessed it could promote earlier thinking and planning of changes. Qualitative data shows staff in accreditation programmes who had started working towards the standards prior to even joining had managed to create a culture of change readiness on their ward and revealed the presence of many vital readiness for change constructs at this stage. Depending on how wards score on their baseline assessments, services may become more aware of areas of organisational improvement capability that they might need to work on prior to participating in a quality improvement programme ([Woodhead, 2016](#)).

Force-field analysis is based on Kurt Lewin's Field Theory ([1939](#)) and is a method for listing, discussing, and assessing the various forces for and against a proposed change. If services were to use methodology such as this during implementation of recommendations following external peer review, it could help them to look at the larger picture by analysing all the forces that could potentially impact on the change and weigh up the pros and cons prior to implementation. Having identified these forces, services can then develop strategies to reduce the impact of the opposing forces (restraining forces) and

strengthen the supporting forces (driving forces), which may support more successful change.

To improve communication within the host service and dissemination from external peer review, services could formalise feedback structures and include these in business MDT meetings.

## **Wider implications**

Against a backdrop of unwanted variation and pressures to deliver better value healthcare, there is a rising demand for a clear and coherent strategy for quality improvement both in the UK, and across the world ([Appleby et al., 2011](#)).

Networks have emerged as a recent government strategy for moving health research into action by creating clusters that break down disciplinary and geographic boundaries ([French et al., 2009](#)). However, communication structures alone are unlikely to be successful for knowledge transfer across specialised domains ([Swan, 2007](#)), ([Carlile, 2004](#)) without additional mechanisms to support the transfer of practice and process knowledge ([Blackler et al., 2000](#)). CMOCs from this thesis can be used to enhance the ability of networks to bring about change.

Denmark have outlined in their National Quality Programme for Health 2015-2018 ([2015](#)) a decision to replace current accreditation schemes with teams of clinical experts. Those who have experienced particularly good results will assist departments where results are not as good, with practical advice and guidance on best practices. The aim would be for outgoing quality teams to support local quality improvements, without introducing new requirements for external procedures or rules that might deprive motivation and ownership. Instead the idea would be act as an additional resource to rapidly improve quality. Findings from this thesis on causal mechanisms are congruent with a model centred around learning and sharing best practice, and have highlighted the salience of ownership.

As varying quality improvement approaches and perhaps external peer review are introduced into healthcare systems in coming years; knowing that changes are underpinned by causal mechanisms such as learning and sharing will help coordinating bodies, healthcare providers and healthcare professionals to maximise the benefits they receive for embarking on quality improvement.

### **13.8 Final Remarks**

I have presented a range of conclusions drawn from this study. It has offered several recommendations for practice, policy and further research in the area, while at the same time acknowledging the limitations inherent within the research itself. Finally, this chapter has served to highlight the contribution that this study has made on many fronts to knowledge and understanding of the process of change in external peer review programmes. In a future where more services might be provided by voluntary and private sector organisations, competition between services could increase and reduce the opportunities for the vital mechanism of learning and sharing. External peer review could become even more substantial in providing opportunities for sharing and learning to take place.

## References

- ACADEMY OF MEDICAL ROYAL COLLEGES. 2015. *Royal Colleges - Academy of Medical Royal Colleges* [Online]. Available: <http://www.aomrc.org.uk/Royal-Colleges/> [Accessed 18th October 2015 2015].
- ACHS 2002. The EQUiP guide: a framework to improve quality and safety in healthcare. Sydney: Australian Council on Healthcare Standards.
- AIMOLA, L., JASIM, S., TRIPATHI, N., HOLDER, S., QUIRK, A. & CRAWFORD, M. J. 2016. Quality of low secure services in the UK: development and use of the Quality of Environment In Low Secure Services (QELS) checklist. *The Journal of Forensic Psychiatry & Psychology*.
- ALKHENIZAN, A. & SHAW, C. 2011. Impact of Accreditation on the Quality of Healthcare Services: a Systematic Review of the Literature. *Ann Saudi Med*, 31, 407 - 416.
- ALKHENIZAN, A. & SHAW, C. 2012. The attitude of health care professionals towards accreditation: A systematic review of the literature. *J Family Community Med*, 19, 74 - 80.
- ALTMAN, D. G. 1991. *Practical statistics for medical research*.
- APPLEBY, J., RALEIGH, V., FROSINI, F., BEVAN, G., GAO, H. & LYSCOM, T. 2011. Variations in healthcare: the good, the bad and the inexplicable. London: The King's Fund.
- APPLETON, S. 2012. Defining mental health services: promoting effective commissioning and supporting QIPP. NHS Confederation. Mental Health Network: NHS Confederation.
- ARMENAKIS, A. A. & BEDIAN, A. G. 1999. Organizational change: a review of theory and research in the 1990s. *Journal of Management*, 25, 293-315.
- ARNOLD, M. E. 2015. Connecting the Dots: Improving Extension Program Planning with Program Umbrella Dots. *Journal of Human Sciences and Extension*, 3.
- ASHKANASY, N. M., WILDEROM, C. P. M. & PETERSON, M. F. 2000. *Handbook of Organizational Culture & Climate*, Thousand Oaks, SAGE Publications.
- ASTBURY, B. & LEEUW, F. 2010. Unpacking black boxes: mechanisms and theory building in evaluation *American Journal of Evaluation*, 31, 363-81.
- ATKINSON, S., INGHAM, J., CHESHIRE, M. & WENT, S. 2010. Defining quality and quality improvement. *Clinical Medicine*, 10, 537-9.
- BALKIZAS, D. 1995. Peer accreditation of a development unit. *Nursing Standard*, 9, 25 - 27.
- BARBOUR, R. S. 2001. Checklists for improving rigour in qualitative research: a case of the tail wagging the dog. *BMJ*, 322, 1115-1117.
- BARBOUR, R. S. & KITZINGER, J. 1999. *Developing focus group research, politics theory and practice*, London, SAGE Publications.
- BARNES, M., MATKA, E. & SULLIVAN, H. 2003. Evidence, understanding and complexity: evaluation in non-linear systems. *Evaluation*, 9, 265-84.
- BARNES, T. R. E. & PATON, C. 2011. Improving prescribing practice in psychiatry: the experience of the Prescribing Observatory for Mental Health (POMH-UK). *International Review of Psychiatry*, 23, 328 - 335.

- BARNES, T. R. E. & PATON, C. 2012. Role of the Prescribing Observatory for Mental Health. *The British Journal of Psychiatry*, 201, 428-429.
- BASKIND, R., KORDOWICZ, M. & CHAPLIN, R. 2010. How does an accreditation programme drive improvement on acute inpatient mental health wards? An exploration of members' views. *Journal of Mental Health*, 19, 405-411.
- BATE, P. 2014. Context is everything. *Perspectives on Context*. London: The Health Foundation
- BATE, P., ROBERT, G. & MCLEOD, H. 2002. Report on the Breakthrough Collaborative approach to quality and service improvement within four regions of the NHS. University of Birmingham.
- BELL, B. A., ENE, M., SMILEY, W. & SCHOENEBERGER, J. A. 2013. A Multilevel Model Primer Using SAS® PROC MIXED. *SAS Global Forum 2013*.
- BERWICK, D. 1989. Continuous improvement as an ideal in health care. *N Eng J Med*, 320, 53-6.
- BERWICK, D. M. 1990. Peer Review and Quality Management: Are They Compatible? *QRB*, 246-251.
- BEVAN, H. 2016. Improvement programmes are yesterday's news: it's the era of the improvement platform. NHS Horizons.
- BEYER, M., GERLACH, F. M., FLIES, U. & GROL, R. 2003. The development of quality circles / peer review groups as a method of quality improvement in Europe. *Family Practice*, 20.
- BHASKAR, R. 1978. *A Realist Theory of Science*, Hassocks, Harvester Press.
- BHASKAR, R. 1989. *Reclaiming reality: a critical introduction to contemporary philosophy*, London, Verso.
- BHASKAR, R., ARCHER, M., COLLIER, A., LAWSON, T. & NORRIE, A. 1998. *Critical Realism: Essential Readings*, Routledge.
- BLACKLER, F., CRUMP, N. & MCDONALD, S. 2000. Organizing processes in complex activity networks. *Organization*, 7, 277-300.
- BLAIKIE, N. 2003. *Analyzing Quantitative Data: From Description to Explanation*, SAGE Publications Ltd.
- BLAISE, P. & KEGELS, G. 2004. A realistic approach to the evaluation of the quality management movement in health care systems: a comparison between European and African contexts based on Mintzberg's organizational models. *International Journal of Health Planning and Management*, 19, 69-93.
- BOADEN, R., HARVEY, G., MOXHAM, C. & PROUDLOVE, N. 2008. Quality Improvement: Theory and Practice in Healthcare. *NHS Institute for Innovation and Improvement*. Coventry: National Library for Health, NHS Institute for Innovation and Improvement.
- BOBIAK, S. N., ZYAZANSKI, S. J., RUHE, M. C., CARTER, C. A., RAGAN, B., FLOCKE, S. A., STANGE, K. C. & LITAKER, D. 2009. Measuring practice capacity for change: a tool for guiding quality improvement in primary care settings. *Q Manage Health Care*, 18, 278 - 284.
- BOHIGAS, L. & HEATON, C. 2000. Methods for external evaluation of health care institutions. *International Journal for Quality in Health Care*, 2, 23 - 238.
- BOONE, H. N. & BOONE, D. A. 2012. Analyzing Likert Data. *Journal of Extension*, 50.
- BOWLING, A. 2014. *Research Methods in Health: Investigating health and health services*, Berkshire, England, Open University Press.

- BOYATZIS, R. E. 1998. *Transforming Qualitative Information: Thematic Analysis and Code Development*, Cleveland, SAGE.
- BOYNTON, P. M. 2004. Administering, analysing, and reporting your questionnaire. *BMJ*, 328.
- BOYNTON, P. M. & GREENHALGH, T. 2004. Selecting, designing, and developing your questionnaire. *BMJ*, 328.
- BRADLEY, E. H., CURRY, L. A., RAMANADHAN, S., ROWE, L., NEMBARD, I. M. & KRUMHOLZ, H. M. 2009. Research in action: using positive deviance to improve quality of health care. *Implementation Science*, 4, 25.
- BRAIN, J., SCHOFIELD, J., GERRISH, K., MAWSON, S., MABBOTT, I., PATEL, D. & GERRISH, P. 2011. A Guide for Clinical Audit, Research and Service Review Healthcare Quality Improvement Partnership.
- BRAITHWAITE, J., GREENFIELD, D., WESTBROOK, J., PAWSEY, M., WESTBROOK, M., GIBBERD, R., NAYLOR, J., NATHAN, S., ROBINSON, M., RUNCIMAN, B., JACKSON, M., TRAVAGLIA, J., JOHNSTON, B., YEN, D., MCDONALD, H., LOW, L., REDMAN, S., JOHNSON, B., CORBETT, A., HENNESSY, D., CLARK, J. & LANCASTER, J. 2010. Health service accreditation as a predictor of organisational performance: a blinded, random, stratified study. *Qual Saf Health Care*, 19, 14 - 21.
- BRAITHWAITE, J., SHAW, C. D., MOLDOVAN, M., GREENFIELD, D., HINCHCLIFF, R., MUMFORD, V., KRISTENSEN, M. B., WESTBROOK, J., NICKLIN, W., FORTUNE, T. & WHITTAKER, S. 2012. Comparison of health service accreditation programs in low- and middle-income countries with those in higher income countries: a cross-sectional study. *International Journal for Quality in Health Care*, 24.
- BRAITHWAITE, J., WESTBROOK, J., PAWSEY, M., GREENFIELD, D., NAYLOR, J., IEDEMA, R., RUNCIMAN, B., REDMAN, S., JORM, C., ROBINSON, M., NATHAN, S. & GIBBERD, R. 2006. A prospective, multi-method, multi-disciplinary, multi-level, collaborative, social-organisational design for researching health sector accreditation. *BMC Health Services Research*, 6.
- BRASURE, M., STENSLAND, J. & WELLEVER, A. 2000. Quality Oversight: Why Are Rural Hospitals Less Likely to be JCAHO Accredited? . *The Journal of Rural Health*, 16.
- BRAY, B. 2013. Stroke: the peer review scheme that is improving care. *Health Service Journal*, 123, 17 - 19.
- BRENNAN, T. A. & BERWICK, D. M. 1996. *New rules: regulation, markets and the quality of American health care*, San Fransisco, Jossey Bass.
- BRIAN, H., COOK, S., TAYLOR, D., FREEMAN, L., MUNDY, T. & KILLASPY, H. 2015. Occupational therapists as change agents in multidisciplinary teams. *British Journal of Occupational Therapy*, 78, 547-555.
- BRIGGS, T. 2015. A national review of adult elective orthopaedic services in England: Getting It Right First Time. British Orthopaedic Association.
- BRITTEN, N. 1995. Qualitative Research: Qualitative interviews in Medical Research. *BMJ*, 311, 215-253.
- BRITTEN, N., CAMPBELL, R., POPE, C., DONOVAN, J., MORGAN, M. & PILL, R. 2002. Using meta ethnography to synthesise qualitative research: A worked example. *Journal of Health Services Research and Policy*, 7, 209-215.



- BROER, T., NIEBOER, A. P. & BAL, R. A. 2010. Opening the black box of quality improvement collaboratives: an Actor-Network theory approach. *BMC Health Services Research*, 10.
- BROOK, R. H., MCGLYNN, E. A. & CLEARY, P. D. 1996. MEASURING QUALITY OF CARE. *New England Journal of Medicine*.
- BRYMAN, A. 2012. *Social Research Methods*, Oxford, Oxford University Press.
- BUKONDA, N., TAVROW, P., ABDALLAH, H., HOFFNER, K. & TEMBO, J. 2002. Implementing a national hospital accreditation program: the Zambian experience. *International Journal for Quality in Health Care*, 14, 7-16.
- BURGOYNE, J., WILLIAMS, S. & WALMSLEY, J. 2009. Evaluation of the Shared Leadership for Change Award Scheme. London: The Health Foundation.
- BURNETT, S., BENN, J., PINTO, A., PARAND, A., ISKANDER, S. & VINCENT, C. 2010. Organisational readiness: exploring the preconditions for success in organisation-wide patient safety improvement programmes. *Qual Saf Health Care*, 19, 313 - 317.
- BURNETT, S., VINCENT, C., MOORTHY, K. & HANNA, G. 2007. Report for the National Cancer Action Team: Evaluation of the National Cancer Peer Review Programme 2004 - 2007. *safe quality care*.
- BURRELL, G. & MORGAN, G. 1979. *Sociological paradigms and organisational analysis : elements of the sociology of corporate life*, London, Heinemann Educational.
- BUTTERFIELD, R., MCCORMICK, B., ANDERSON, R., BALL, J., WHITE, J. & ELEFThERIADES, C. 2012. Quality of NHS care and external pathway peer review. Oxford: Centre for Health Services Economics & Organisation.
- BYNG, R., NORMAN, I. & REDFERN, S. 2005. Using realistic evaluation to evaluate a practice-level intervention to improve primary healthcare for patients with long-term mental illness. *Evaluation*, 11, 69-93.
- BYNG, R., NORMAN, I., REDFERN, S. & JONES, R. 2008. Exposing key functions of a complex intervention for shared care in mental health: case study of a process evaluation. *BMC Health Services Research*, 8.
- CALDER, B. J. 1977. Focus Groups and the Nature of Qualitative Marketing Research. *Journal of Marketing Research*, 14, 353-364.
- CALDWELL, D. F., CHATMAN, J., O'REILLY III, C. A., ORMISTON, M. & LAPIZ, M. 2008. Implementing strategic change in a health care system: the importance of leadership and change readiness. *Health Care Manage Rev*, 33, 124 - 133.
- CALDWELL, S. E. M. & MAYS, N. 2012. Studying policy implementation using a macro, meso and micro frame analysis: the case of the Collaboration for Leadership in Applied Health Research & Care (CLAHRC) programme nationally and in North West London. *Health Research Policy and Systems*, 32.
- CARE QUALITY COMMISSION 2015. How CQC regulates specialist mental health services: Provider Handbook. CQC.
- CARLILE, P. R. 2004. Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries. *Organ Sci*, 15, 555-568.

- CARMINES & ZELLER 1979. *Reliability and validity assessment*, London, SAGE.
- CASEBEER, A., POPP, J. & SCOTT, C. 2009. Positively deviant networks: what are they and why do we need them? *Journal of Health Organization and Management*, 23, 610 - 626.
- CHALMERS, I. & ALTMAN, D. G. 1995. *Systematic reviews*, London, BMJ Publishing Group.
- CHAO, S. Advancing Quality Improvement Research: Challenges and Opportunities - Workshop Summary Forum on the Science of Health Care Quality Improvement and Implementation, 2007 Washington, D.C.: Institute of Medicine.
- CHINN, S. 1991. Scale, parametric methods, and transformations. *Thorax*, 46.
- CHISWICK, D. 1992. Reed report on mentally disordered offenders. *BMJ*, 305.
- CLAHRC NWL. 2016. *QI4U* [Online]. CLAHRC NWL. Available: <http://qi4u.ocbmedia.com/login> [Accessed 28th April 2016 2016].
- CLARK, A. M., WHELAN, H. K., BARBOUR, R. & MACINTYRE, P. D. 2005. A realist study of the mechanisms of cardiac rehabilitation. *Journal of Advanced Nursing*, 52, 362-71.
- CLARKE, J. 2011. What is a systematic review? *Evidence-based Nursing*, 14, 64.
- CLASON, D. L. & DORMODY, T. J. 1994. Analyzing Data Measured By Individual Likert-Type Items. *Journal of Agricultural Education*, 35, 31-35.
- CLEGG, S. & HARDY, C. 1996. Introduction: organisations, organisation and organising. In: CLEGG, S., HARDY, C. & NORD, W. (eds.) *Handbook of Organization Studies*. London: SAGE.
- COETSEE, C. & STANZ, K. 2007. 'Barriers-to-change' in a governmental service delivery type organisation. *South Africa Journal of Industrial Psychology*, 33, 76-83.
- CONNELLY, J. 2007. Evaluating complex public health interventions: theory, methods and scope of realist enquiry. *Journal of Evaluation in Clinical Practice*, 13, 935-41.
- CRABTREE, B. F. & MILLER, W. L. 1999. *Doing Qualitative Research*, Thousand Oaks, CA, SAGE.
- CRAIG, P., DIEPPE, P., MACINTYRE, S., MICHIE, S., NAZARETH, I. & PETTICREW, M. 2008. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*, 337.
- CREAN, P. M., STOKES, M. A., WILLIAMSON, C. & HATCH, D. J. 2003. Quality in paediatric anaesthesia: a pilot study of interdepartmental peer review. *Anaesthesia*, 58, 543-548.
- CRESWELL, J. W. 2009. *Research Design: Qualitative, Quantative, and Mixed Methods Approaches*, London, SAGE Publications.
- CRESWELL, J. W. 2013. *Qualitative inquiry & research design: choosing among five approaches*, Thousand Oaks, CA, SAGE Publications.
- CRESWELL, J. W. 2014. *Research Design: Qualitative, Quantative, and Mixed Methods Approaches*, London, SAGE Publications.
- CRESWELL, J. W. & PLANO CLARK, V. L. 2011. *Designing and conducting mixed methods research*.
- CRETIN, S., SHORTELL, S. M. & KEELER, E. B. 2004. An evaluation of collaborative interventions to improve chronic illness care, framework and study design. *Evaluation Review*, 28 - 51.

- CUNNINGHAM, F. C., RANMUTHUGALA, G., PLUMB, J., GEORGIU, A., WESTBROOK, J. I. & BRAITHWAITE, J. 2011. Health professional networks as a vector for improving healthcare quality and safety: a systematic review. *BMJ Quality & Safety*.
- DARZI, A. 2008. High quality care for all: NHS Next Stage Review final report. London: Department of Health.
- DAVIES, S. & KILLASPY, H. 2015. Rehabilitation in hospital settings. *Enabling recovery: The principles and practice of rehabilitation psychiatry*. 2nd ed. London: Royal College of Psychiatrists.
- DAVIS, M. V., CANNON, M. M., STONE, D. O., WOOD, B. W., REED, J. & BAKER, E. L. 2011. Informing the National Public Health Accreditation Movement: Lessons From North Carolina's Accredited Local Health Departments. *American Journal of Public Health*, 101.
- DAWSON, P. 2003. *Understanding Organizational Change: The Contemporary Experience of People at Work*, London, SAGE Publications.
- DEETZ, S. 2008. Resistance: Would Struggle by Any Other Name Be as Sweet? *Management Communication Quarterly*, 21, 387-392.
- DELAMOTHE, T. & SMITH, R. 1996. Redesigning the journal: having your say. *BMJ*, 312, 232-234.
- DENSCOMBE, M. 2007. *The good research guide: for small-scale social research projects*, Maidenhead, England, Open University Press.
- DENZIN, N. & LINCOLN, Y. 2008. *The landscape of qualitative research: Theories and issues*, Thousand Oaks, CA, SAGE Publications.
- DEPARTMENT OF CONSTITUTIONAL AFFAIRS 1998. Data Protection Act. London: Stationary Office.
- DEPARTMENT OF HEALTH 2002. National Minimum Standards for General Adult Services in Psychiatric Intensive Care Units (PICU) and Low Secure Environments. Department of Health.
- DEPARTMENT OF HEALTH. 2006. *No excuses. Embrace partnership now. Step towards change!* [Online]. Available: [http://webarchive.nationalarchives.gov.uk/20081108205945/dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_4137144](http://webarchive.nationalarchives.gov.uk/20081108205945/dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4137144) [Accessed 16th June 2016 2016].
- DEPARTMENT OF HEALTH 2007. Best Practice Guidance: Specification for Adult Medium-Secure Services Department of Health.
- DEPARTMENT OF HEALTH 2008. High Quality Care For All.
- DEPARTMENT OF HEALTH 2011. Liberating the NHS: Transparency in outcomes - a framework for the NHS. Department of Health.
- DEPARTMENT OF HEALTH 2012a. Low Secure Services Good practice commissioning guide, consultation draft. Department of Health.
- DEPARTMENT OF HEALTH 2012b. No health without mental health: implementation framework.
- DEVERS, K. J., PHAM, H. H. & LIU, G. 2004. What is Driving Hospitals' Patient-Safety Efforts? *Health Affairs*, 23, 103-115.
- DIXON-WOODS, M., AGARWAL, S., YOUNG, B., JONES, D. & SUTTON, A. 2004. Integrative approaches to qualitative and quantitative evidence. London: Health Development Agency.
- DIXON-WOODS, M. & FITZPATRICK, R. 2001. Qualitative research in systematic reviews. *BMJ*, 323, 765-766.

- DIXON-WOODS, M., MCNICOL, S. & MARTIN, G. 2012. Ten challenges in improving quality in healthcare: lessons from the Health Foundation's programme evaluations and relevant literature. *BMJ Qual Saf*, 21, 876 - 884.
- DOI, L., JEPSON, R. & CHEYNE, H. 2015. A realist evaluation of an antenatal programme to change drinking behaviour of pregnant women. *Midwifery*, 31, 965-972.
- DONABEDIAN, A. 1966. Evaluating the quality of medical care. *The Milbank Memorial Fund Quarterly*, 44, 166-203.
- DONABEDIAN, A. 1980. *The Definition of Quality and Approaches to its Assessment*, Ann Arbor, Michigan, Health Administration Press.
- DONABEDIAN, A. 1988. The quality of care: how can it be assessed? *JAMA*, 260, 1743-1748.
- DONABEDIAN, A. 1990. The seven pillars of quality. *Arch Pathol Lab Med*, 114, 1115-8.
- DONAHUE, K. T. & VAN OSTENBERG, P. 2000. Joint Commission International accreditation: relationship to four models of evaluation. *International Journal for Quality in Health Care*, 12, 243 - 246.
- DONNER, A. & KLAR, N. 2000. *Design and Analysis of Cluster Randomization Trials in Health Research*, London, Arnold Publishers.
- DOWNIE, R. & CALMAN, K. 1998. *Healthy respect: ethics in health care* Oxford, Oxford University Press.
- DOYLE, G. & GRAMPP, C. 2014. Accreditation as a Quality Tool in Public Sector Reform: The Fourth Stage of Convergence. *Working Paper*. Dublin: University College Dublin.
- DUCKERS, M. L., SPREEUWENBERG, P., WAGNER, C. & GROENEWEGEN, P. P. 2009a. Exploring the black box of quality improvement collaboratives: modelling relations between conditions, applied changes and outcomes. *Implement Sci*, 4, 74.
- DUCKERS, M. L. A., SPREEUWENBERG, P., WAGNER, C. & GROENEWEGEN, P. P. 2009b. Exploring the black box of quality improvement collaboratives: modelling relations between conditions, applied changes and outcomes. *Implementation Science*, 4.
- DUCKETT, S. J. 1983. Changing hospitals: the role of hospital accreditation. *Soc Sci Med.*, 17, 1573-9.
- EDWARDS, M. T. 2011. The Objective Impact of Clinical Peer Review on Hospital Quality and Safety. *American Journal of Medical Quality*, 26, 110-119.
- EL-JARDALI, F. 2007. Hospital Accrediation Policy in Lebanonon. *Lebanese Medical Journal*, 55.
- ELY, M., ANZUL, M., FRIEDMAN, T., GARNER, D. & STEINMETZ, A. C. 1991. *Doing qualitative research. Circles within circles.*, New York, Falmer Press.
- ERLANDSON, D. A., HARRIS, E. L., SKIPPER, B. L. & ALLEN, S. D. 1993. *Doing naturalistic inquiry: a guide to methods*, Newbury Park, CA, SAGE.
- ESTABROOKS, C. A., FIELD, P. A. & MORSE, J. M. 1994. Aggregating qualitative findings: an approach to theory development. *Qualitative Health Research*, 4, 503-511.
- EZZY, D. 2002. *Qualitative analysis practice and innovation*, London, Routledge.

- FERLIE, E., FITZGERALD, L., WOOD, M. & HAWKINS, C. 2005. THE NONSPREAD OF INNOVATIONS: THE MEDIATING ROLE OF PROFESSIONALS. *Academy of Management Journal*, 48, 117-134.
- FINLAY, L. 2002. Negotiating the swamp: the opportunity and challenge of reflexivity in research practice. *Qualitative Research*, 2, 209-230.
- FISHER, R. A. & BENNETT, J. H. 1990. *Statistical methods, experimental design, and scientific inference : A re-issue of Statistical methods for research workers, The design of experiments, and Statistical methods and scientific inference*, Oxford, Oxford University Press.
- FLEETWOOD, S. 2005. The ontology of organisation and management studies: A critical realist approach. *Organization*, 12, 197-222.
- FLODGRÉN, G., POMEY, M. P., TABER, S. A. & ECCLES, M. P. 2011. Effectiveness of external inspection of compliance with standards in improving healthcare organisation behaviour, healthcare professional behaviour or patient outcomes. *The Cochrane Collaboration*.
- FOURNIER, D. M. 2005. Logic of Evaluation-Working Logic. In: MATHISON, S. (ed.) *Encyclopedia of Evaluation*. London: SAGE Publications.
- FOWLER, F. J. 2013. *Survey Research Methods*, SAGE Publications.
- FRENCH, B., THOMAS, L. H., BAKER, P., BURTON, C. R., PENNINGTON, L. & RODDAM, H. 2009. What can management theories offer evidence-based practice? A comparative analysis of measurement tools for organisational context. *Implement Sci*, 4.
- FRENCH, W. & BELL, C. 1998. *Organization Development: Behavioural Science Interventions for Organization Improvement*, Englewood Cliffs, Prentice Hall.
- FRENCH, W. L., BELL, C. H. & ZAWACKI, R. A. 2004. *Organization Development and Transformation*, Singapore, McGraw-Hill.
- FUNNELL, S. C. & RODGERS, P. J. 2011. *Purposeful Program Theory: Effective Use of Theories of Change and Logic Models*.
- FURBER, C. 2010. Framework analysis: a method for analysing qualitative data. *African Journal of Midwifery and Women's Health*, 4.
- FURST, S. A. & CABLE, D. M. 2008. Employee Resistance to Organizational Change: Managerial Influence Tactics and Leader-Member Exchange. *Academy of Management Journal*, 93, 453-462.
- GABBAY, J., LE MAY, A., CONNELL, C. & KLEIN, J. H. 2014. Skilled for improvement? Learning communities and the skills needed to improve care: an evaluative service development. *Original Research*. London: The Health Foundation.
- GALBRAITH, S., DANIEL, J. A. & VISSSEL, B. 2010. A Study of Clustered Data and Approaches to Its Analysis. *The Journal of Neuroscience*, 30, 10601-10608.
- GALE, N. K., HEATH, G., CAMERON, E., RASHID, S. & REDWOOD, S. 2013. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. 13.
- GELSNE, C. & PESHKIN, A. 1992. *Becoming qualitative researchers: an introduction*, White Plains, NY, Longman.
- GIBBS, C. E. & CHEETHAM, P. S. 1988. Voluntary Review of Obstetric and Gynecological Services. *Quality Review Bulletin*, 14.

- GIDDINGS, L. S. & GRANT, B. M. 2007. A Trojan Horse for Positivism? A Critique of Mixed Methods Research. *Advances in Nursing Science*, 30.
- GLASER, B. G. & STRAUSS, A. L. 1970. Discovery of substantive theory. In: FILSTEAD, W. (ed.) *Qualitative methodology*. Chicago: Rand McNally.
- GOERTZ, G. & MAHONEY, J. 2012. *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences*, Princeton, NJ, Princeton University Press.
- GREEN, J. & THOROGOOD, N. 2009. *Qualitative Methods for Health Research*, London, SAGE Publications.
- GREENE, J. C. 2007. *Mixed methods in social inquiry*.
- GREENFIELD, D. & BRAITHWAITE, J. 2008. Health sector accreditation research: a systematic review. *International Journal for Quality in Health Care*, 20, 172 - 183.
- GREENFIELD, D., PAWSEY, M. & BRAITHWAITE, J. 2010. What motivates professionals to engage in the accreditation of healthcare organizations? *International Journal for Quality in Health Care*, 23, 8 - 14.
- GREENFIELD, D., PAWSEY, M. & BRAITHWAITE, J. 2012. A Peer-To-Peer Model to Improve Patient Safety: Harnessing Accreditation Programs. *American Journal of Medical Quality*, 27, 353 - 354.
- GREENHALGH, T., HUMPHREY, C., HUGHES, J., MACFARLANE, F., BUTLER, C. & PAWSON, R. 2009. How Do You Modernize a Health Service? A Realist Evaluation of Whole-Scale Transformation in London. *The Milbank Quarterly*, 87, 391-416.
- GREENHALGH, T. & PEACOCK, R. 2005. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ*, 331, 1064-1065.
- GRIMSHAW, J., MCAULEY, L. M., BERO, L. A., GRILLI, R., OXMAN, A. D., RAMSAY, C., VALE, L. & ZWARENSTEIN, M. 2003. Systematic reviews of the effectiveness of quality improvement strategies and programmes. *Qual Saf Health Care*, 12, 298 - 303.
- GROL, R. 1992. Implementing guidelines in general practice care. *Quality in Health Care*, 1, 184-91.
- GROL, R. 1994. Quality improvement by peer review in primary care: a practical guide. *Quality in Health Care*, 3, 147 - 152.
- GROL, R., BAKER, R. & MOSS, F. 2004. *Quality improvement research : understanding the science of change in health care*.
- GUEST, G., BUNCE, A. & JOHNSON, L. 2006. How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability. *Field Methods*, 18.
- GUO, P., EAST, L. & ARTHUR, A. 2015. Thinking outside the black box: The importance of context in understanding the impact of a preoperative education nursing intervention among Chinese cardiac patients. *Patient Education and Counselling*, 95, 365-370.
- GUSTAFSON, D. H., SAINFORT, F., EICHLER, M., ADAMS, L., BISOGNANO, M. & STEUDEL, H. 2003. Developing and testing a model to predict outcomes of organisational change. *Health Services Research*, 38, 751 - 766.
- HACKMAN, J. R. & WAGEMAN, R. 1995. Total quality management: empirical, conceptual and practical issues. *Administrative Science Quarterly*, 40, 309-342.

- HALLADAY, M. & BERO, L. 2000. Getting research into practice: implementing evidence-based practice in health care. *Publ Money Manag*, 20, 43-50.
- HAM, C., BERWICK, D. & DIXON, J. 2016. Improving quality in the English NHS: A strategy for action. The King's Fund.
- HAM, C., KIPPING, R. & MCLEOD, H. 2003. Redesigning work processes in health care: lessons from the National Health Service. *The Milbank Quarterly*, 81, 415-439.
- HAMEL, G. & ZANINI, M. 2014. *Build a change platform, not a change program* [Online]. McKinsey Insights & Publications. Available: <http://www.mckinsey.com/business-functions/organization/our-insights/build-a-change-platform-not-a-change-program> [Accessed 28th April 2016 2016].
- HAYES, A. F. 2006. A Primer on Multilevel Modeling. *Human Communication Research*, 32.
- HAYWARD, R. A., MCMAHON, L. F. & BERNARD, A. M. 1993. Evaluating the Care of General Medicine Inpatients: How Good Is Implicit Review? *Annals of Internal Medicine*, 118, 550-556.
- HEATON, C. 2000. External peer review in Europe: an overview from the ExPeRT Project. *International Journal for Quality in Health Care*, 12, 177 - 182.
- HECK, R. H. & THOMAS, S. L. 2000. *An introduction to multilevel modeling techniques*.
- HEIDEMANN, E. G. 2000. Moving to global standards for accreditation processes: the ExPeRT Project in a larger context. *International Journal for Quality in Health Care*, 2, 227 - 230.
- HEMMING, K., HAINES, T. P., CHILTON, P. J., GIRLING, A. J. & LILFORD, R. J. 2015. The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ*, 315.
- HENRY, G. T. 2005. Realist evaluation. In: MATHISON, S. (ed.) *Encyclopedia of Evaluation*. London: SAGE Publications.
- HIGGINS, J. P. T. & GREEN, S. 2011. *Cochrane Handbook for Systematic Reviews of Interventions: Version 5.1.0.*, The Cochrane Collaboration.
- HINCHCLIFF, R., GREENFIELD, D., MOLDOVAN, M., WESTBROOK, J. I., PAWSEY, M., MUMFORD, V. & BRAITHWAITE, J. 2012. Narrative synthesis of health service accreditation literature. *BMJ Quality & Safety*, 0, 1-13.
- HINCHCLIFF, R., GREENFIELD, D., WESTBROOK, J., PAWSEY, M., MUMFORD, V. & BRAITHWAITE, J. 2013. Stakeholder perspectives on implementing accreditation programs: a qualitative study of enabling factors. *BMC Health Services Research*, 13.
- HOLDEN, J. D. 2001. Hawthorne effects and research into professional practice. *Journal of Evaluation in Clinical Practice*, 7, 65-70.
- HOLT, D. T., ARMENAKIS, A. A., FEILD, H. S. & HARRIS, S. G. 2007. Readiness for Organizational Change : The Systematic Development of a Scale. *Journal of Applied Behavioral Science*, 43.
- HOLT, D. T., HELFRICH, C. D., HALL, C. G. & WEINER, B. J. 2008. Are you ready? How health professionals can comprehensively conceptualize readiness for change. *J Gen Intern Med*, 35, 50 - 55.

- HOVLID, E. & BUKVE, O. 2014. A qualitative study of contextual factors' impact on measures to reduce surgery cancellations. *BMC Health Services Research*, 14.
- HOX, J. J. 2010. *Multilevel Analysis: Techniques and Applications*, New York, NY, Routledge.
- HUDSON, D. W., HOLZMUELLER, C. G., PRONOVOST, P. J., GIANCI, S. J., PATE, Z. T., WAHR, J., HEITMILLER, E. S., THOMPSON, D. A., MARTINEZ, E. A., MARSTELLER, J. A., GURSES, A. P., LUBOMSKI, L. H., GOESCHEL, C. A. & PHAM, J. C. 2012. Toward Improving Patient Safety Through Voluntary Peer-to-Peer Assessment. *American Journal of Medical Quality*, 27.
- HURST, K. 1997. The nature and value of small and community hospital accreditation. *International Journal of Health Care Quality Assurance* 10, 94-106.
- INSTITUTE FOR HEALTHCARE IMPROVEMENT 2003. The Breakthrough Series: IHI's Collaborative Model for Achieving Breakthrough Improvement. *Innovation Series 2003*.
- INSTITUTE OF MEDICINE 2000. To err is human: Building a safer health system. Washington, DC: The National Academies Press.
- INSTITUTE OF MEDICINE 2001. *Crossing the Quality Chasm: A New Health System for the 21st Century*, Washington, D.C., National Academy Press.
- INSTITUTE OF MEDICINE 2003. Health professions education: A bridge to quality. Washington, DC: The National Academies Press.
- INSTITUTE OF MEDICINE 2011. The future of nursing: Leading change, advancing health. Washington, DC: The National Academies Press.
- IRVINE, D. & IRVINE, S. 1991. *Making sense of audit*, Oxford, Radcliffe Medical Press.
- JAAFARIPOOYAN, E., AGRIZZI, D. & AKBARI-HAGHIGHI, F. 2011. Healthcare accreditation systems: further perspectives on performance measures. *International Journal for Quality in Health Care*, 1-12.
- JAMES, L. R. 1982. Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology*, 67, 219-229.
- JAMIESON, S. 2004. Likert scales: how to (ab)use them. *MEDICAL EDUCATION*, 38.
- JAMTVEDT, G., YOUNG, J. M., KRISTOFFERSEN, D. T., O'BRIEN, M. A. & OXMAN, A. D. 2010. Audit and feedback: effects on professional practice and health care outcomes (Review). *The Cochrane Collaboration*.
- KAPLAN, H. C., BRADY, P. W., DRITZ, M. C., HOOPER, D. K., LINAM, W. M., FROEHLE, C. M. & MARGOLIS, P. 2010. The Influence of Context on Quality Improvement Success in Health Care: A Systematic Review of the Literature. *The Milbank Quarterly*, 88, 500-559.
- KEELER, E. B., RUBENSTEIN, L. V., KAHN, K. L., DRAPER, D., HARRISON, E. R., MCGINTY, M. J., ROGERS, W. H. & BROOK, R. H. 1992. Hospital characteristics and quality of care. *JAMA*, 268, 1709-14.
- KERN, M. 2002. Reducing Risk and Enhancing Value Through Accreditation. *Nursing Homes*, 34 - 37.
- KILSDONK, M., SIESLING, S., OTTER, R. & VAN HARTEN, W. 2015. Evaluating the impact of accreditation and external peer review. *International Journal of Health Care Quality Assurance*, 28.



- KILSDONK, M. J., VAN DIJK, B. A. C., OTTER, R. & VAN HARTEN, W. H. 2014. The impact of organisational external peer review on colorectal cancer treatment and survival in the Netherlands. *British Journal of Cancer*, 110, 850 - 858.
- KLAZINGA, N. 2000. Re-engineering trust: the adoption and adaptation of four models for external quality assurance of health care services in western European health care systems. *International Journal for Quality in Health Care*, 12, 183 - 189.
- KLAZINGA, N., FISCHER, C. & TEN ASBROEK, A. 2011. Health services research related to performance indicators and benchmarking in Europe. *Journal of Health Services Research & Policy*, 16, 38-47.
- KLAZINGA, N., LOMBARTS, K. & VAN EVERDINGEN, J. 1998. Quality management in medical specialties: the use of channels and dikes in improving health care in The Netherlands. *Jt Comm J Qual Improv*, 24, 240-50.
- KLEIN, K. J. & KOZLOWSKI, S. W. J. 2000. *Multilevel theory, research, and methods in organizations: Foundations, extensions and new directions*, Jossey-Bass.
- KMET, L. M., LEE, R. C. & COOK, L. S. 2004. STANDARD QUALITY ASSESSMENT CRITERIA FOR EVALUATING PRIMARY RESEARCH PAPERS FROM A VARIETY OF FIELDS. Alberta Heritage Foundation for Medical Research.
- KREIN, S. L., DAMSCHRODER, L. J., KOWALSKI, C. P., FORMAN, J., HOFER, T. P. & SAINT, S. 2010. The influence of organizational context on quality improvement and patient safety efforts in infection prevention: A multi-center qualitative study. *Social Science & Medicine*, 71, 1692-1701.
- KUMAR, J. 2014. *Research methodology: a step-by-step guide for beginners*, London, SAGE Publications.
- LEATHERMAN, S. & SUTHERLAND, K. 2003. The quest for quality: a midterm evaluation of the ten-year quality agenda guide. London: Nuffield Trust.
- LEDFORD, G. E., MOHRAM, S. A., MOHRMAN, A. M. & LAWLER, E. E. 1990. The phenomenon of large-scale organizational change. In: MOHRMAN, A. M., MOHRAM, S. A., LEDFORD, G. E., CUMMINGS, T. G. & LAWLER, E. E. (eds.) *Large-Scale Organizational Change*. San Fransisco Jossey-Bass.
- LEHMAN, W. E. K., GREENER, J. M. & SIMPSON, D. D. 2002. Assessing organizational readiness for change. *Journal of Substance Abuse Treatment*, 22, 197 - 209.
- LEMIEUX-CHARLES, L., GAULT, N., CHAMPAGNE, F., BARNSLEY, J., TRABUT, I., SICOTTE, C. & ZITNER, D. 2000. Use of mid-level indicators in determining organizational performance. *Hosp Q*, 3, 48-52.
- LEWIN, K. 1939. Field Theory and Experiment in Social Psychology: Concepts and Methods. *American Journal of Sociology*, 44, 868-896.
- LEWIN, K. 1951. *Field Theory in Social Science*, New York, Harper.
- LEWIS, G. & KILLASPY, H. 2014. Getting the measure of outcomes in clinical practice. *Advances in psychiatric treatment*, 20, 165-171.
- LIBERATI, A., ALTMAN, D. G., TETZLAFF, J., MULROW, C., GØTZSCHE, P. C., IOANNIDIS, J. P. A., CLARKE, M., DEVEREAUX, P. J., KLEIJNEN, J. & MOHER, D. 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*, 339.

- LIKERT, R. 1932. A Technique for the Measurement of Attitudes. *Archives of Psychology*, 150, 1-55.
- LINCOLN, Y. & GUBA, E. G. 1999. Establishing Trustworthiness. In: BRYMAN, A. & BURGESS, R. G. (eds.) *Qualitative Research*. London: SAGE Publications.
- LINCOLN, Y. S. & GUBA, E. G. 1985. *Naturalistic inquiry*, Beverley Hills, CA, SAGE.
- LINCOLN, Y. S. & GUBA, E. G. 1994. Competing paradigms in qualitative research. In: DENZIN, N. K. & LINCOLN, Y. S. (eds.) *Handbook of qualitative research*. Thousand Oaks, CA: SAGE.
- LING, T., SOPER, B., BUXTON, M., HANNEY, S., OORTWIJN, W., SCOGGINS, A. & STEEL, N. 2010. How Do You Get Clinicians Involved in Quality Improvement? An Evaluation of the Health Foundation's Engaging with Quality Initiative - a Programme of Work to Support Clinicians to Drive Forward Quality. London: The Health Foundation.
- LOHR, K. N. 1990. *Medicare: A Strategy for Quality Assurance*, Washington, DC, National Academy Press.
- LOMBARTS, M. J. M. H. & B., V. W. F. C. 2003. External Peer Review by Medical Specialists (Visitatie) in a Legal Perspective. *European Journal of Health Law*, 10, 43 - 51.
- LOMBARTS, M. J. M. H. & KLAZINGA, N. S. 2003. Supporting Dutch medical specialists with the implementation of visitatie recommendations: a descriptive evaluation of a 2-year project. *International Journal for Quality in Health Care*, 15, 119 - 129.
- LONGMAN, H. 2015. Hospital accreditation to end – in Denmark. *Paperwork out for hospitals*. GP Access UK.
- MACK, N., WOODSONG, C., MCQUEEN, K. M., GUEST, G. & NAMEY, E. 2005. *Qualitative research methods: a data collector's field guide*, Carolina, Family Health International.
- MAINZ, J. 2003. Defining and classifying clinical indicators for quality improvement. *International Journal for Quality in Health Care*, 15, 523–530.
- MALBY, B. & MERVYN, K. 2012. Summary of the literature to inform the Health Foundation questions. Centre for Innovation Management: University of Leeds.
- MARCHAL, B., DEDZO, M. & KEGELS, G. 2010. A realist evaluation of the management of a well-performing regional hospital in Ghana. *BMC Health Services Research*, 10.
- MARCHAL, B., VAN BELLE, S., VAN OLMEN, J., HOERÉE, T. & KEGELS, G. 2012. Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, 18, 192 - 212.
- MARK, M. M., HENRY, G. T. & JULNES, G. 2000. *Evaluation: An integrated framework for understanding, guiding, and improving policies and programs*, San Fransisco, Jossey-Bass.
- MARSHALL, C. & ROSSMAN, G. 2011. *Designing qualitative research*, Thousand Oaks, CA, SAGE.
- MARSHALL, M. & BAMBER, J. The role of context in improving the quality of patient care. 2011. The Health Foundation
- MASON, F., THOMAS, D. & WILSON, L. 2015. Quality assurance and clinical audit in psychiatric care. In: DICKENS, G., SUGARMAN, P. & PICCHIONI, M. (eds.) *Handbook of Secure Care*. London: RCPsych Publications.

- MASON, J. 2002. *Qualitative researching*, London, SAGE Publications.
- MAXWELL, J. & MITTAPALLI, K. 2010. Realism as a stance for mixed methods research. In: TEDDLIE, A. T. C. (ed.) *Handbook of Mixed Methods in Social and Behavioral Research*. 2nd ed. Thousand Oaks, CA: Sage Publications.
- MAXWELL, R. J. 1984. Quality assessment in health. *BMJ*, 288.
- MCEVOY, P. & RICHARDS, D. 2003. Critical realism: a way forward for evaluation research in nursing? *Journal of Advanced Nursing*, 43.
- MCEVOY, P. & RICHARDS, D. 2006. A critical realist rationale for using a combination of quantitative and qualitative methods. *Journal of Research in Nursing*, 11, 66-78.
- MCGURRIN, M. C. & HADLEY, T. R. 1991. Quality of Care and Accreditation Status of State Psychiatric Hospitals. *Hospital and Community Psychiatry*, 42.
- MCLOUGHLIN, V. & LEATHERMAN, S. 2003. Quality or financing: what drives design of the health care system. *Quality and Safety in Health Care*, 12, 136-142.
- MENTAL HEALTH STRATEGIES 2012. 2011/12 National Survey of Investment in Adult Mental Health Services.
- MERALI, F. 2005. NHS managers' commitment to a socially responsible role: the NHS managers' views of their core values and their public image. *Social Responsibility Journal*, 1, 38-46.
- MERTON, R. K. 1968. *Social Theory and Social Structure*, New York, The Free Press.
- MILES, M. B. & HUBERMAN, M. A. 1994. *An Expanded Sourcebook: Qualitative Data Analysis*, Thousand Oaks, California, SAGE Publications.
- MILES, M. B., HUBERMAN, M. A. & SALDAÑA, J. 2013. *Qualitative data analysis : a methods sourcebook*.
- MILLER, R. & REES, J. 2014. Mental health commissioning: master or subject of change? *Mental Health Review Journal*, 19, 145-155.
- MING-CHU, Y. & MENG-HSIU, L. 2014. Unlocking the black box: Exploring the link between perceive organizational support and resistance to change. *Asia Pacific Management Review*, 20, 177-183.
- MINTZBERG, H. 1973. *The Nature of Managerial Work*, New York, Harper & Row.
- MOHER, D., HOPEWELL, S., SCHULZ, K. F., MONTORI, V., GØTZSCHE, P. C., DEVEREAUX, P. J., ELBOURNE, D., EGGER, M. & ALTMAN, D. G. 2010. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ*.
- MORGAN, D. L. 1997. *Focus groups as qualitative research*, London, SAGE Publications.
- MORGAN, G. & SMIRCICH, L. 1980. The Case for Qualitative Research *The Academy of Management Review*, 5, 491-500.
- MORSE, J. M. 1991. Approaches to Qualitative-Quantitative Methodological Triangulation. *Nursing Research* 40, 120-123.
- MULLEN, P. E. 2000. Forensic Mental Health. *British Journal of Psychiatry*, 176, 307-311.
- NABITZ, U., KLAZINGA, N. & WALBURG, J. 2000. The EFQM excellence model: European and Dutch experiences with the EFQM approach in health care. *International Journal for Quality in Health Care*, 12.
- NEEDHAM, D. M., SINOPOLI, D. J., DINGLAS, V. D., BERENHOLTZ, S. M., KORUPOLU, R., WATSON, S. R., LUBOMSKI, L., GOESCHEL, C. &

- PRONOVOST, P. J. 2009. Improving data quality control in quality improvement projects. *Int J Qual Health Care*, 21, 145-50.
- NHS COMMISSIONING BOARD 2013. NHS Standard Contract for Medium and Low Secure Mental Health Services (Adults)
- NHS ENGLAND 2015. Guidance to support the introduction of access and waiting time standards for mental health services in 2015/16. In: MEDICAL DIRECTORATE, M. H. T. (ed.). London: NHS England.
- NICOLAY, C. R., PURKAYASTHA, S., GREENHALGH, A., BENN, J., CHATURVEDI, S., PHILLIPS, N. & DARZI, A. 2012. Systematic Review of the application of quality improvement methodologies from the manufacturing industry to surgical healthcare. *British Journal of Surgery*, 99.
- O'CONNOR, E., FORTUNE, T., DORAN, J. & BOLAND, R. 2007. Involving consumers in accreditation: the Irish experience. *International Journal for Quality in Health Care*, 19.
- O'NEILL, S. M., HEMPEL, S., LIM, Y.-W., DANZ, M. S., FOY, R., SUTTORP, M. J., SHEKELLE, P. G. & RUBENSTEIN, L. V. 2011. Identifying continuous quality improvement publications: what makes an improvement intervention 'CQI'? *BMJ Qual Saf*, 20, 1011-1019.
- OGRINC, G. & BATALDEN, P. 2009. Realist evaluation as a framework for the assessment of teaching about the improvement of care. *Journal of Nursing Education*, 48, 661-7.
- OLSEN, W. 2012. *Data Collection: Key Debates and Methods in Social Research*, London, SAGE Publications.
- OPPENHEIM, A. N. 1992. *Questionnaire Design, Interviewing and Attitude Measurement*, London, Pinter.
- OREG, S. & SVERDLIK, N. 2011. Ambivalence Toward Imposed Change: The Conflict Between Dispositional Resistance to Change and the Orientation Toward the Change Agent. *Journal of Applied Psychology*, 96, 337-349.
- ØVRETVEIT, J. 1997a. A comparison of hospital quality programmes: lessons for other services. *International Journal of Service Industry Management* 8, 220-235.
- ØVRETVEIT, J. 1997b. Would it work for us? Learning from quality improvement in Europe and beyond. *Jt Comm J Qual Improv.*, 23, 7-22.
- ØVRETVEIT, J. 2002a. Producing useful research about quality improvement. *International Journal for Health Care Quality Assurance*, 15, 294-302.
- ØVRETVEIT, J. 2002b. Quality collaboratives: lessons from research. *Quality and Safety in Health Care*, 11, 345-351.
- ØVRETVEIT, J. 2004. A framework for quality improvement translation: understanding the conditionality of interventions. *Joint Commission Journal on Quality and Safety*.
- ØVRETVEIT, J. 2009a. Does improving quality save money? A review of evidence of which improvements to quality reduce costs to health service providers. The Health Foundation.
- ØVRETVEIT, J. 2009b. *Leading evidence informed value improvement in health care*, Chichester, Kingsham Press.
- ØVRETVEIT, J. 2011a. How does context affect quality improvement. *The role of context in improving the quality of patient care*. The Health Foundation.

- ØVRETVEIT, J. 2011b. Understanding the conditions for improvement: research to discover which context influences affect improvement success. *BMJ Qual Saf*, 20.
- ØVRETVEIT, J. & GUSTAFSON, D. 2002. Evaluation of quality improvement programmes. *Qual Saf Health Care*, 11, 270 - 275.
- ØVRETVEIT, J. & GUSTAFSON, D. 2003. Improving the quality of health care: Using research to inform quality programmes. *BMJ*, 326.
- ØVRETVEIT, J. & GUSTAFSON, D. 2004. Evaluation of quality improvement programmes. In: GROL, R., BAKER, R. & MOSS, F. (eds.) *Quality Improvement Research: Understanding the science of health care*. London: BMJ Books.
- OWEN, J. M. 2005a. Organizational Change. In: MATHISON, S. (ed.) *Encyclopedia of Evaluation*. London: SAGE Publications.
- OWEN, J. M. 2005b. Organizational Culture. In: MATHISON, S. (ed.) *Encyclopedia of Evaluation*. London: SAGE Publications.
- PACCIONI, A., SICOTTE, C. & CHAMPAGNE, F. 2007. Accreditation: a cultural control strategy *International Journal of Health Care Quality Assurance in Health Care*, 21.
- PAGE, R. L. & HARRISON, B. D. W. 1995. Setting up interdepartmental peer review. *Journal of the Royal College of Physicians of London*, 29, 319-324.
- PAGE, R. L. & HARRISON, B. D. W. 1997. Interdepartmental Peer Review. *BMJ*, 314.
- PATTON, M. Q. 1980. *Qualitative evaluation methods*, Beverley Hills, CA, SAGE.
- PATTON, M. Q. 2012. *Essentials of Utilization-Focused Evaluation*, Saint Paul, MN, SAGE Publications.
- PATTON, M. Q. 2015. *Qualitative Research & Evaluation Methods*, London, SAGE Publications.
- PAWSON, R. 2002. Evidence-based policy: the promise of 'Realist Synthesis'. *Evaluation*, 8, 340-358.
- PAWSON, R. 2006. *Evidence-Based Policy A Realist Perspective*, SAGE Publications.
- PAWSON, R. & SRIDHARAN, S. 2009. Theory-driven evaluation of public health programmes. In: KILLORAN, A. & KELLY, M. P. (eds.) *Evidence-based Public Health: Effectiveness and efficiency*. Oxford University Press.
- PAWSON, R. & TILLEY, N. 1994. What works in evaluation research? *British Journal of Criminology*, 34, 291-306.
- PAWSON, R. & TILLEY, N. 1997. *Realistic Evaluation*, London, SAGE Publications.
- PAWSON, R. & TILLEY, N. 2005. Realistic Evaluation. In: MATHISON, S. (ed.) *Encyclopedia of Evaluation*. London: SAGE Publications.
- PERNEGER, T. 2006. Ten reasons to conduct a randomized study in quality improvement. *International Journal for Quality in Health Care*, 18.
- PETT, M. A. 1997. *Nonparametric Statistics in Health Care Research: Statistics for Small Samples and Unusual Distributions*, SAGE Publications.
- PETTIGREW, A. 1987. Context and action in the transformation of the firm. *Journal of Management Studies*, 24, 649-670.
- PETTIGREW, A. M. 1992. The character and significance of strategy process research. *Strategic Management Journal*, 13, 5-16.

- PIDERIT, S. 2000. Rethinking resistance and recognising ambivalence: a multidimensional view of attitudes toward an organizational change. *Academy of Management Review*, 25, 783-794.
- PIPER, H., HASSELL, A. B., ROWE, I. F., DELAMERE, J. & WEST MIDLANDS RHEUMATOLOGY SERVICE & TRAINING COMMITTEE 2006. Experience of six years of a regional peer review scheme in rheumatology. *Rheumatology*, 45, 1110-5.
- PLESK, P. E. & GREENHALGH, T. 2001. The challenge of complexity in health care. *BMJ*, 323, 625-8.
- POMEY, M.-P., CONTANDRIOPOULOS, A.-P., FRANCOIS, P. & BERTRAND, D. 2004. Accreditation: a tool for organizational change in hospitals? *International Journal for Health Care Quality Assurance*, 17, 113-124.
- POMEY, M.-P., FRANCOIS, P., CONTANDRIOPOULOS, A.-P., TOSH, A. & BERTRAND, D. 2005. Paradoxes of French accreditation. *Qual Saf Health Care*, 14, 51-55.
- POMEY, M.-P., LEMIEUX-CHARLES, L., CHAMPAGNE, F., ANGUS, D., SHABAH, A. & CONTANDRIOPOULOS, A.-P. 2010. Does accreditation stimulate change? A study of the impact of the accreditation process on Canadian healthcare organizations. *Implement Sci*, 31.
- POMMIER, J., GUÉVAL, M.-R. & JOURDAN, D. 2010. Evaluation of health promotion in schools: a realistic evaluation approach using mixed methods. *BMC Public Health*, 10.
- POPE, C. & MAYS, N. 2006. *Qualitative research in healthcare*, Cambridge, Blackwell.
- POPE, C., ROBERT, G., BATE, P., LE MAY, A. & GABBAY, J. 2006. Lost in translation: a multi-level case study of the metamorphosis of meanings and action in public sector organizational innovation. *Public Administration* 84, 59 - 79.
- PURVIS, G. P., JACOBS, D. & KAK, N. 2010. International Health Care Accreditation Models and Country Experiences: Introductory Report on Options for The Republic of South Africa. USAID.
- RALEIGH, V. S. & FOOT, C. 2010. Getting the measure of quality: opportunities and challenges. The King's Fund.
- RANDALL, S. 2013. Leading networks in healthcare: Learning about what works – the theory and the practice. The Health Foundation.
- RAPLEY, T. 2004. Interviews. In: SEALE, C., GOBO, G., GUBRIUM, J. F. & SILVERMAN, D. (eds.) *Qualitative research practice*. London: SAGE Publications.
- RAUDENBUSH, S. W. & BYRK, A. S. 1992. *Hierarchical linear models: Applications and data analysis methods*, SAGE Publications.
- REED, M. 2005. Reflections on the 'realist turn' in organization and management studies. *Journal of Management Studies*, 42, 1621-44.
- REERINK, E. 1990. Defining quality of care: Mission impossible? *International Journal for Quality in Health Care*, 2, 197 –202.
- REICHERTZ, J. 2014. Induction, deduction, abduction. In: FLICK, U. (ed.) *The SAGE Handbook of Qualitative Data Analysis*. London: SAGE Publications Ltd.
- RITCHIE, J. & LEWIS, J. 2003. *Qualitative research practice : a guide for social science students and researchers*.

- RIVAS, C., TAYLOR, S., ABBOTT, S., CLARKE, A., GRIFFITHS, C., ROBERTS, C. M. & STONE, R. 2010. Collaborative working within UK NHS secondary care and across sectors for COPD and the impact of peer review: qualitative findings from the UK National COPD Resources and Outcomes Project. *International Journal of Integrated Care*, 10.
- RIVAS, C., TAYLOR, S., ABBOTT, S., CLARKE, A., GRIFFITHS, C., ROBERTS, C. M. & STONE, R. 2012. Perceptions of changes in practice following peer review in the National Chronic Obstructive Pulmonary Disease Resources and Outcomes Project. *International Journal of Health Care*, 25, 91 - 105.
- RIVAS, C., TAYLOR, S., CLARKE, A. & GROUP, O. B. O. T. Q. S. S. S. 2008. Final Report on the Qualitative Sub-study of the National COPD Resources and Outcomes Project (NCROP). Barts and the London School of Medicine and Dentistry.
- ROBERT, G., BATE, P. & MENDEL, P. 2008. Organizing for quality: the improvement journeys of leading hospitals in Europe and the United States.
- ROBERT, G. B., ANDERSON, J. E., BURNETT, S. J., AASE, K., ANDERSSON-GARE, B., BAL, R., CALLTORP, J., NUNES, F., WEGGELAAR, A.-M., VINCENT, C. A., FULOP, N. J. & QUASER TEAM 2011. A longitudinal, multi-level comparative study of quality and safety in European hospitals: the QUASER study protocol. *BMC Health Services Research*, 11.
- ROBERTS, C. M., BUCKINGHAM, R. J., STONE, R. A., LOWE, D. & POTTER, J. M. 2012. A randomized trial of peer review: the UK National Chronic Obstructive Pulmonary Disease Resources and Outcomes Project: three-year evaluation. *Journal of Evaluation in Clinical Practice*, 18, 599 - 605.
- ROBERTS, C. M., STONE, R. A., BUCKINGHAM, R. J., PURSEY, N. A., HARRISON, B. D. W., LOWE, D. & POTTER, J. M. 2010. A randomised trial of peer review: the UK National Chronic Obstructive Pulmonary Disease Resources and Outcomes Project. *Clinical Medicine*, 10, 223-7.
- ROBSON, C. 2011. *Real World Research: a resource for users of social research methods in applied settings*.
- ROETHLISBERGER, F. J. & DICKSON, W. J. 1939. *Management and the worker*, Cambridge, MA, Harvard University Press.
- ROGERS, E. M. 1995. *Diffusion of Innovation*, New York, The Free Press.
- ROYAL COLLEGE OF PSYCHIATRISTS' CENTRE FOR QUALITY IMPROVEMENT (CCQI) 2014. Quality Network for Forensic Mental Health Services: Standards for Medium Secure Services RCPsych CCQI.
- RUBIN, H. J. & RUBIN, I. 2005. *Qualitative Interviewing: The Art of Hearing Data*, London, SAGE Publications.
- RUTHERFORD, M. & DUGGAN, S. 2008. Forensic mental health services: facts and figures on current provision. *The British Journal of Forensic Practice*, 10.
- RYCROFT-MALONE, J., FONTENLA, M., BICK, D. & SEERS, K. 2010. Realistic evaluation: the case of protocol-based care. *Implement Sci*, 5.
- SAKE, R. 2010. *Qualitative research: studying how things work*.
- SALMON, J. W., HEAVENS, J., LOMBARD, C., TAVROW, P., WITH FOREWORD BY HEIBY, J. R., AND COMMENTARIES BY WHITTAKER, S., MULLER, M., KEEGAN, M. & ROONEY, A. L. 2003. The Impact of Accreditation on the

- Quality of Hospital Care: KwaZulu-Natal Province, Republic of South Africa. *Operations Research Results*. Bethesda, MD.
- SANDELOWSKI, M. 1994. Notes on transcription. *Research in Nursing & Health*, 17, 311-314.
- SANDELOWSKI, M. 1995. Focus on qualitative methods: sample size in qualitative research. *Research in Nursing & Health*, 18, 179-183.
- SAUNDERS, M., LEWIS, P. & THORNHILL, A. 2016. *Research Methods for Business Students*, Prentice Hall.
- SAYER, A. 2009. *Realism and social science*, London, SAGE Publications.
- SCHEIN, E. 2004. *Organisational culture and leadership*, San Fransisco, Jossey-Bass.
- SCHOUTEN, L. M. T., HULSCHER, M. E. J. L., VAN EVERDINGEN, J. J. E., HUIJSMAN, R. & GROU, R. P. T. M. 2008. Evidence for the impact of quality improvement collaboratives: systematic review. *BMJ*, 336, 1 - 9.
- SCHULPEN, T. W. & LOMBARTS, K. M. 2007. Quality improvement of paediatric care in the Netherlands. *Arch Dis Child*, 92, 633-6.
- SCHULZ, K. F., ALTMAN, D. G. & MOHER, D. 2010. CONSORT 2010 Statement: Updated Guidelines for Reporting Parallel Group Randomized Trials. *Ann Intern Med*, 152, 726-732.
- SCOTT, T., MANNION, R., MARSHALL, M. & DAVIES, H. 2003. Does organisational culture influence health care performance? A review of the evidence. *J Health Serv Res Policy*, 8, 105-17.
- SCRIVEN, M. 1967. *The methodology of evaluation perspectives of curriculum evaluation*, Chicago, IL, Rand McNally.
- SCRIVENS, E. 1995. International trends in accreditation. *The International Journal of Health Planning and Management*, 10, 165-181.
- SCRIVENS, E. 1997. The impact of accreditation systems upon patient care. *J Clin Effect*, 2, 102-105.
- SCRIVENS, E. 1998. Widening the scope of accreditation - issues and challenges in community and primary care. *International Journal for Quality in Health Care*, 10, 191 - 197.
- SCRIVENS, E. J., KLEIN, R. & STEINER, A. 1995. Accreditation: what we can learn from the Anglophone model? *Health Policy*, 34, 193 - 204.
- SHAH, J., WORRALL, A., CRISP, H., SHAW, C. & WEBSTER, A. 2010. A Survey of UK Standards-based External Assessment Schemes.
- SHAW, C. 2001. External assessment of health care. *BMJ*, 322, 851 - 854.
- SHAW, C. 2004. The external assessment of health services. *World Hospitals and Health Services*, 40, 24 - 27.
- SHAW, C. D. 2000. External quality mechanisms for health care: summary of the ExPeRT project on visitatie, accreditation, EFQM and ISO assessment in European countries. *International Journal for Quality in Health Care*, 12.
- SHAW, C. D. 2003. Measuring against clinical standards. *Clinica Chimica Acta*, 333, 115 - 124.
- SHOUTEN, L. M. T., HULSCHER, M. E. J. L., AKKERMANS, R., VAN EVERDINGEN, J. J. E., GROU, R. P. T. M. & HUIJSMAN, R. 2008. Factors That Influence the Stroke Care Team's Effectiveness in Reducing the Length of Hospital Stay. *Stroke*, 39.
- SILVERMAN, D. 2005. *Doing qualitative research: a practical handbook*, London, SAGE.



- SILVERMAN, D. 2011. *Interpreting Qualitative Data*, London, SAGE Publications.
- SIM, J. 1998. Collecting and analyzing qualitative data: issues raised by the focus group. *Journal of Advanced Nursing*, 28, 345-352.
- SLINKER, B. K. & GLANTZ, S. A. 2008. Multiple linear regression: accounting for multiple simultaneous determinants of a continuous dependent variable. *Circulation*, 117, 1732-7.
- SNIJDERS, T. A. B. & BOSKER, R. J. 1999. *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, London, SAGE Publications.
- SPRANGERS, M. A. G. & SCHWARTZ, C. E. 1999. Integrating response shift into health-related quality of life research: a theoretical model. *Social Science & Medicine*, 48, 1507-15.
- STEBBING, J. F. 2011. Quality assurance of endoscopy units. *Best Practice & Research Clinical Obstetrics & Gynaecology*, 25, 361 - 370.
- STICKLEY, T. 2006. Should service user involvement be consigned to history? A critical realist perspective. *Journal of Psychiatric and Mental Health Nursing*, 13.
- SUGARMAN, P. & DICKENS, G. 2015. The evolution of secure and forensic mental healthcare. In: DICKENS, G., SUGARMAN, P. & PICCHIONI, M. (eds.) *Handbook of Secure Care*. London: RCPsych Publications.
- SUÑOL, R., VALLEJO, P., THOMPSON, A., LOMBARTS, M. J. M. H., SHAW, C. D. & KLAZINGA, N. 2009. Impact of quality strategies on hospital outputs. *Qual Saf Health Care*, 18, i62-i68.
- SVYANTEK, D. & BOTT, J. 2004. Organizational culture and organizational climate measures: an integrative review. In: THOMAS, J. (ed.) *Comprehensive handbook of psychological assessment: industrial and organizational assessment*. Hoboken (NJ): Wiley.
- SWALLOW, V., NEWTON, J. & VAN LOTTUM, C. 2002. How to manage and display qualitative data using 'Framework' and Microsoft® Excel. *Journal of Clinical Nursing*, 12, 610 - 612.
- SWAN, J. 2007. Managing knowledge for innovation: production, process and practice. In: MCINERNEY, C. R. & DAY, R. E. (eds.) *Knowledge Management: from knowledge objects to knowledge processes*. Berlin: Springer.
- SWEENEY, J. & HEATON, C. 2000. Interpretations and variations of ISO 9000 in acute health care. *Int J Quality Health Care*, 12, 203-9.
- TANON, A. A., CHAMPAGNE, F., CONTANDRIOPOULOS, A.-P., POMEY, M.-P., VADEBONCOEUR, A. & NGUYEN, H. 2010. Patient safety and systematic reviews: finding papers indexed in MEDLINE, EMBASE and CINAHL. *Quality and Safety in Health Care*, 19, 452-461.
- TASHAKKORI, A. & TEDDLIE, C. 2010. *SAGE Handbook of mixed methods in social and behavioural research*, SAGE Publications.
- TAUBMAN, M. A. 1989. Whither Peer Review? *J Dent Res*, 68, 1379-1380.
- TEDDLIE, C. & TASHAKKORI, A. 2009. *Foundations of mixed methods research integrating quantitative and qualitative approaches in the social and behavioural sciences*, SAGE Publications.
- THE HEALTH FOUNDATION 2011. Safer Patients Initiative Phase One: Mixed Method Evaluation of a Large-Scale Organisational Intervention to Improve Patient Safety in Four UK Hospitals. London: The Health Foundation

- THE HEALTH FOUNDATION 2013. Lining Up: How do improvement programmes work? Lessons from an ethnographic research study of interventions to reduce central line infections. *Learning Report*. London: The Health Foundation.
- THE HEALTH FOUNDATION 2014. Effective networks for improvement. The Health Foundation.
- THE NHS INFORMATION CENTRE (COMMUNITY & MENTAL HEALTH TEAM) 2011. Establishment of Memory Services - Results of a survey of Primary Care Trusts, final figures.
- THE ROYAL COLLEGE OF PSYCHIATRISTS 2013. English National Memory Clinics Audit Report. The Royal College of Psychiatrists.
- THE ROYAL COLLEGE OF PSYCHIATRISTS 2015. Second English National Memory Clinics Audit Report.
- THOMAS, L., CULLUM, N., MCCOLL, E., ROUSSEAU, N., SOUTTER, J. & STEEN, N. 2000. Guidelines in professions allied to medicine. *Cochrane Database Syst Rev*, 2.
- THURSTON, W. E., COVE, L. & MEADOWS, L. M. 2008. Methodological congruence in complex and collaborative mixed methods studies. *International Journal of Multiple Research Approaches*, 2, 2-14.
- TOLSON, D., MCINTOSH, J., LOFTUS, L. & CORMIE, P. 2005. Developing a managed clinical network in palliative care: a realistic evaluation. *Int J Nurs Stud*, 44, 183-195.
- TOUATI, N. & POMEY, M. 2009. Accreditation at a crossroads: Are we on the right track? *Health Policy*, 90.
- VALORI, R. M., ROGERS, C., JOHNSTON, D. K. & INGHAM, J. 2013. Developing a strategy for accreditation of clinical services. *Clinical Medicine*, 13, 538 - 542.
- VAN BELLE, G., FISHER, L. D., HEAGERTY, P. J. & LUMLEY, T. 2004. *Biostatistics: a methodology for the health sciences*.
- VAN DER KNAAP, L. M., LEEUW, F. L. & BOGAERTS, S. 2008. Combining Campbell Standard and the Realist Evaluation Approach: The Best of Two Worlds? *American Journal of Evaluation*, 29, 48-57.
- VAN HARTEN, W. H., CASPARIE, T. F. & FISSCHER, O. A. 2000. Methodological considerations on the assessment of the implementation of quality management systems. *Health Policy*, 54.
- VAN WEERT, C. 2000. Developments in professional quality assurance towards quality improvement: some examples of peer review in the Netherlands and the United Kingdom. *International Journal for Quality in Health Care*, 12, 239 - 242.
- VISWANATHAN, H. N. & SALMON, J. W. 2000. Accrediting Organizations and Quality Improvement. *The American Journal of Managed Care*, 6.
- WALLIMAN, N. 2005. *Your Research Project: A Step-by-Step Guide for the first time Researcher*, London, SAGE Publications.
- WALSHE, D. & DOWNE, S. 2004. Meta-synthesis method for qualitative research: a literature review. *Journal of Advanced Nursing*, 50, 204-211.
- WALSHE, K. 1999. Improvement through inspection? The development of the new Commission for Health Improvement in England and Wales. *Qual Health Care*, 8, 191-196.

- WALSHE, K. 2007. Understanding what works—and why—in quality improvement: the need for theory-driven evaluation. *International Journal for Quality in Health Care*, 19, 57–59.
- WALSHE, K. & SHORTELL, S. M. 2004. When Things Go Wrong: How Health Care Organizations Deal With Major Failures. *Health Affairs*, 23
- WALSHE, K., WALLACE, L., FREEMAN, T., LATHAM, L. & SPURGEON, P. 2001. The external review of quality improvement in health care organizations: a qualitative study. *International Journal for Quality in Health Care*, 13, 367-374.
- WANBERG, C. R. & BANAS, J. T. 2000. Predictors and Outcomes of Openness to Changes in a Reorganizing Workplace. *Journal of Applied Psychology*, 85, 132-142.
- WARD, S., DONOVAN, H. S. & SERLIN, R. C. 2003. An Alternative View on “An Alternative Paradigm”. *Res Nurs Health*, 26, 256–259.
- WAREHAM, N. J. 1993. External monitoring of quality of health care in the United States. *Quality in Health Care*, 3, 97 - 101.
- WEBER, P. S. & WEBER, J. E. 2001. Changes in employee perceptions during organizational change. *Leadership & Organisation Development Journal*, 22.
- WEINER, B. J. 2009. A theory of organizational readiness to change. *Implementation Science*, 4.
- WERNER, R. M. & ASCH, D. A. 2007. Clinical concerns about clinical performance. *Health Service Journal*, 5, 159-63.
- WHITLEY, R. & CRAWFORD, M. 2005. Qualitative research in psychiatry. *Can J Psychiatry*, 50.
- WHO 2003. Quality and Accreditation in Health Care Services: A Global Review. Geneva, Switzerland: World Health Organization.
- WILKINSON, S., JOFFE, H. & YARDLEY, L. 2004. Qualitative Data Collection: Interviews and Focus Groups. In: MARKS, D. & YARDLEY, L. (eds.) *Research methods for clinical and health psychology*. London: SAGE Publications.
- WILSON, T., BERWICK, D. M. & CLEARY, P. D. 2003. What Do Collaborative Improvement Projects Do? Experience from Seven Countries. *Joint Commission Journal on Quality and Safety*, 29.
- WOODHEAD, T. 2016. Five top tips for our checklist for building organisational improvement capability. *Five top tips for our checklist for building organisational improvement capability* [Online]. Available from: [http://www.health.org.uk/blog/five-top-tips-our-checklist-building-organisational-improvement-capability/?utm\\_source=charityemail&utm\\_medium=email&utm\\_campaign=april-2016&pubid=healthfoundation&description=april-2016&dm\\_i=4Y2,46OS6,F69VVBG,F8CZF,1](http://www.health.org.uk/blog/five-top-tips-our-checklist-building-organisational-improvement-capability/?utm_source=charityemail&utm_medium=email&utm_campaign=april-2016&pubid=healthfoundation&description=april-2016&dm_i=4Y2,46OS6,F69VVBG,F8CZF,1) [Accessed 8th May 2016 2016].
- WORRALL, A. 2011. The service context for clinical guidelines: Supporting guideline implementation by assuring and improving the quality of service in which clinicians work. *International Review of Psychiatry*, 23, 336 - 341.
- ZIKMUND, W. G., BABIN, B. J., CARR, J. C. & GRIFFIN, M. 2012. *Business Research Methods*, South-Western College Pub.

## Appendix A: Permission sought from authors

Sarah,

As far as I am concerned, you don't need my permission!

Do whatever makes sense and is helpful for your context.

If you can let us know when you have some results using these instruments or their modifications, we'd be interested.

Good luck!

Kurt

On Fri, Oct 26, 2012 at 10:35 AM, Jasim, Sarah <[sarah.jasim08@imperial.ac.uk](mailto:sarah.jasim08@imperial.ac.uk)> wrote:

Hi Kurt,

Many thanks for your help and co-operation for sending me the instruments.

Can I seek your permission of modifying some of the language to make it applicable to the setting of our study? For example: changing 'patient' to 'service user', and changing 'practice' to 'unit' etc.

Warm regards,

Sarah Jasim

---

From: Kurt Stange [[kcs@case.edu](mailto:kcs@case.edu)]

Sent: 25 October 2012 15:37

To: Jasim, Sarah

Cc: [bobiak@nccn.org](mailto:bobiak@nccn.org); Sara Torok

Subject: Re: Authorization for use of instrument

Hi Sara(h)s.

I've attached 2 versions of the instrument. The one described in Sarah Bobiak's paper (which is completed by an observer), and a parallel version for self administration by people working within the practice. We haven't done the psychometric analysis of the latter, but I think it has face validity from being the same items.

Good luck with your work, Sarah Jasmin!

Kurt Stange

On Tue, Oct 23, 2012 at 12:01 PM, Sara Torok

<[sac@case.edu](mailto:sac@case.edu)<<mailto:sac@case.edu>>> wrote:

Hello Sarah,

I'll be happy to help, but to start off with, let's run your request by Kurt Stange, himself. I've cc'd him here.

Kurt, please see the details of Sarah's request below. We'll call her "Sarah J." since there are so many of us Sara(h)s in this conversation. :-)

Sara.

Sara A. Torok, Senior Faculty Assistant

CWRU Department of Family Medicine & Community Health, Research Division

10900 Euclid Avenue, Cleveland, OH 44106-7136

Visitors/Deliveries: 11000 Cedar Ave., Ste 402

Internal Mail: BioEnterprise 402, LC 7136

ph: [216-368-2756](tel:216-368-2756)<tel:[216-368-2756](tel:216-368-2756)>; fax: [216-368-4348](tel:216-368-4348)<tel:[216-368-4348](tel:216-368-4348)>,

email: [sac@case.edu](mailto:sac@case.edu)<mailto:[sac@case.edu](mailto:sac@case.edu)>

## Appendix B: Focus group participant demographic checklist

CCQI Focus Group Participant Demographics		
<b>Date:</b>	<b>Time:</b> 12.30-13.30	<b>Place:</b> 4 <sup>th</sup> Floor Meeting Room (Large)
<b>Name:</b>		
<b>Age:</b>	<b>Duration at CCQI:</b>	<b>Projects worked on (since start):</b>
20-30 <input type="checkbox"/>	Less than 1 year <input type="checkbox"/>	
31-40 <input type="checkbox"/>	1-5 years <input type="checkbox"/>	
41-50 <input type="checkbox"/>	5-10 years <input type="checkbox"/>	
51-60 <input type="checkbox"/>	Over 10 years <input type="checkbox"/>	
Over 60 <input type="checkbox"/>		

## Appendix C: CCQI focus group topic guide

### PRE-FOCUS GROUP QUESTIONS:

1. Do they work?
2. What aspects of the process did you feel worked particularly well?
3. Have you come across any obstacles, or things that you feel did not work?

### INTRODUCTION:

*How many of you have attended focus groups before? Introduce myself as a facilitator, and the purpose and process of a focus group.*

*Thank you for making the time to talk to me today, and agreeing to share your views and experiences with each other. For some of you who may not be familiar with my research, I am looking into the process of peer-led quality improvement initiatives here at CCQI; in an attempt to understand the relationship between what goes on and the different outcomes that are generated. By discussing different themes and hearing your views and different experiences today, it will give me an 'insider's view' of the process, and really help me to build a foundation for my research.*

*For those of you who don't know each other – it would be helpful to start with some introductions; please continue to help yourself to refreshments throughout the session. If we could quickly go around the table and state name, your role at the CCQI and which network you work on, and how long you've worked here for – that would be great!*



TOPIC:

*Discuss pre-FG questions. Get participants to express their own views.*

*People have moved from different projects – ‘did you notice any difference from moving from one project to another?’ compare and contrast Accreditation staff views with Quality Network staff views throughout.*

*Try to illicit project worker <-> project worker ‘insider’ views.*

## CORE QUESTIONS:

1. **Why do you think services join?**
  - Motivation (*bottom-up vs. top-down*)
  - Role of the decision making party
  - Whose benefit?
  
2. **Why do some services change / improve more than others?**
  - Why do you say that?
  - Opportunities / obstacles
  - Awareness
  - Responsibility
  - Organisational set-ups helping / getting in the way?
  
3. **When (in the process) do these changes take place?**
  - Sign up, before, during, after?
  - Feedback delivery
  - How could this be improved / taken advantage of?

If time permits:

4. **What early signs have services shown, indicating their capacity to benefit?**
  - Why is this important?
  - What do you think of these aspects?
  
5. **What are the main differences between new / old members?**
  - Why do you think that?
  - Can anything be done about this?

## CLOSE:

*Any other questions?*

*Sum up and thank participants.*

## FILL IN DEMOGRAPHIC QUESTIONNAIRE:

*Name, age, etc.*

*Is there anything that you didn't say in the focus group that you'd like the researcher to know?*

## **Appendix D: Draft e-mail to member services**

Dear [Service Manager],

As an established member of the Quality Network for Forensic Mental Health, I would like to thank you for all your participation and commitment to improving quality.

To enable us to enhance the support we provide to services, you have been automatically selected to contribute as a voluntary participant in a research study funded by Imperial College London. The study aims to examine the influential contextual factors which impact peer-led quality improvement initiatives. This research will help to enhance the initiatives provided by the CCQI; ensuring they are delivering the necessary factors – pertinent to fostering improvement.

Your participation would involve a short 30-minute interview with yourself and some of your staff members (separately) over the course of July 2013 – July 2014. The interview will seek your opinions and experiences, following your decision to join the Quality Network for Forensic Mental Health. Participation will not affect your membership with the CCQI, and is not linked to the support you receive from them (i.e. peer review visit).

Please expect to be contacted with regards to scheduling an interview date over the next few months. However, if you do not wish to participate in this voluntary research study; please send an e-mail to [sarah.jasim08@imperial.ac.uk](mailto:sarah.jasim08@imperial.ac.uk) or call 07872693579 to opt-out.

Warm regards,

Sarah Jasim

## **Appendix E: Participant information sheet**



**Royal College of Psychiatrists'  
Centre for Quality Improvement  
21 Mansell Street  
4th Floor Standon House  
London, E1 8AA**

**Examining factors that influence the impact of peer-led quality improvement initiatives**

### **Participant Information Sheet (Version 2.0, 4 June 2013)**

You are being invited to take part in a research study. Before deciding whether or not to participate, it is important for you to understand the purpose of the research, and what it will involve. Please take time to read the following information carefully. Contact details are provided overleaf – please do not hesitate to ask if further information is required.

#### **What is the purpose of the study?**

Peer-led quality improvement initiatives (including peer review networks and accreditation programmes) aim to enhance standards of health care by engaging and supporting clinicians and managers to improve the quality of care they provide. Previous research of the quality of care provided by services that participate in these initiatives have demonstrated that compliance with standards increases over time. However, the true impact of participation is largely unknown. This indicates a knowledge gap, which requires more evidence on the factors that impact peer-led quality improvement initiatives.

### **Who is funding it?**

The study has been funded by the Faculty of Medicine at Imperial College London.

### **How does the study work?**

The study proposes to use evidence from semi-structured interviews with staff members to examine the process of peer-led quality improvement initiatives, methods of feedback, and factors which influence the impact of these programmes. The outcomes of the study will be used to inform future improvements made to the initiatives, which subsequently will help services maintain and achieve improved quality of care.

### **Why have I been chosen?**

Staff will be recruited from services whom are existing members of CCQI networks. Staff members are invited to participate in semi-structured interviews, to express their views on joining their respective network, so they can share their experiences in participation so far.

### **What does taking part involve?**

A researcher will spend up to one day at your service to interview approximately four members of staff to collect data for the study. Individual participation will involve undertaking an interview with one researcher. If agreed to each interview will be audio recorded, and will last approximately 30 minutes. If you cannot undertake the interview on the day of the researcher's visit, there will be an opportunity to reschedule the interview for another visit, or if no arrangements can be made – you can answer the researcher's questions over the telephone on a time and date that is convenient to you.

### **Do I have to take part?**

No, participation in this study is entirely voluntary. If you decide against participation, this will bear no effect on your services involvement in the network. If you initially agree to participate but change your mind after completing it, just inform the researcher and it will be ensured that the information you provided will not be included. Again, this will not affect your services involvement in the network.

### **Are there any risks in taking part?**

We do not think that your involvement in the study will pose any risk to you. If you find it difficult or upsetting to complete the interview, you can stop at any time, skip questions you do not want to answer or withdraw from the study at any time. If you feel you would like help or support due to any distress felt through taking part in the study, you can contact my supervisor at the CCQI – Dr Alan Quirk: [aquirk@cru.rcpsych.ac.uk](mailto:aquirk@cru.rcpsych.ac.uk)

### **Is the study confidential?**

Yes. Any information that you give us will be treated in strict confidence and it will not be traceable to you individually. Furthermore, details will not be shared with employers, and any identifiable information, including the tape recording, will be destroyed at the end of the study. However, any concerns which regarding harm to patients or staff, which arise from interviews, will be escalated to my supervisor. Published qualitative research often involves direct quotations to strengthen discussion or analysis findings for the audience. For this reason, direct quotations (all identifiable information removed) may be used in publications from this study.

### **Can I see the results?**

The findings from this study will be published in the researchers PhD thesis (October, 2015). Following approval from the Imperial College Examination Board – this will be made available to all Imperial College students and

researchers via the Imperial College Digital Repository: Spiral. Papers based on study findings will be prepared for submission in peer-reviewed journals. Oral presentations of the study findings will be presented to staff at the CCQI and at national conferences relating to quality improvement. A summary of study findings will also be sent to any participant who requests this, and made publicly available via the CCQI website.

**Who do I contact for more information?**

If you would like further information about the study, please contact:

Sarah Jasim, PhD Researcher – Email: [sj808@imperial.ac.uk](mailto:sj808@imperial.ac.uk) Tel: 07872693579

## Appendix F: Participant consent form



**Royal College of Psychiatrists'  
Centre for Quality Improvement  
21 Mansell Street  
4th Floor Standon House  
London, E1 8AA**

### **Examining factors that influence the impact of peer-led quality improvement initiatives**

#### **Participant Consent Form (Version 2.0, 4 June 2013)**

- 1 I confirm that I have read and understood the Participant Information Sheet (Version 2.0- 04.06.2013) for the above study and have had the opportunity to ask questions.
  
2. I agree for the researcher from the Royal College of Psychiatrists' Centre for Quality Improvement to collect the information required for the study, as described in the Participant Information Sheet (Version 2.0- 04.06.2013).
  
3. I agree to be interviewed by the researcher in order to provide the information specified in the Participation Information Sheet (Version 2.0- 04.06.2013).
  
4. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving a reason.
  
5. I understand that all information provided will be kept confidential.
  
6. I agree for this interview to be audio-recorded



7. I agree to take part in the above stated study.

\_\_\_\_\_

Name of Participant

Date

Signature

\_\_\_\_\_

Name of Researcher

Date

Signature

## Appendix G: Interview participant demographic checklist

Personal Information	Answer
Your name:	
Name and postcode of your service:	
Name of your network:	
Length of time with network (if known):	
Please state your role:	
Length of time working for service:	
How many times have you participated in a Peer Review of your own service?	
How many times have you been a reviewer on a Peer Review Team?	

## Appendix H: Interview topic guide

### 1. PERCEPTIONS OF MEANING / PURPOSE OF THE NETWORK

Question: What does QI (and the low secure network in particular) mean to you?

Prompts:

*-Members of other QI initiatives*

Question: Could you tell me what you know about the network?

Prompts:

*-Define / identify meaning*

Question: What is your understanding of the purpose of the network?

Prompts:

*-What do you hope to gain from the network?*

### 2. MECHANISMS OF ACTION / INVOLVEMENT IN NETWORK

Question: What sort of involvement have you had in the network?

Prompts:

*-How did you first become aware of the network / first impressions?*

*-Do you know why / how it came about that your service joined the network?*

*-Who is involved? (Actors: colleagues, senior staff)*

*-What do you think about this unit being involved? Were you involved in this decision?*

*-What preparation is involved for peer review visits?*

*-Tell me what your experience of peer review visits is? (Going on visits, receiving this)*

*-What was your involvement in the last peer review visit?*

*-What did that feel like? (Threatened, supported, learning, time wasted)*

*-Do you think participation in this network does / could benefit this service?*

*-How does the process help your service? You? Affect your job role?*

*-What do you think might change as a result of the network? (Staff, unit, personally – staff level)*

*-Compare experiences with other QI initiatives (mentioned at the start)*

### 3. CONTEXTUAL INFLUENCES

Question: How has the CCQI network helped to change practice – (for good or ill)?

Question: Can you give me an example of where your service's involvement in the low secure network has led to changes in practice on the ward?

Question: How did the network help to bring about these changes?

Questions: Are there any benefits other than quality improvement? (prestige, staff morale etc.)

Prompts:

-What are the most helpful parts of the process?

-Are there any people that have helped / hindered this process?

- Any other factors that have helped / hindered this process (other initiatives, resources)

### LEADERSHIP

Prompts:

-Role of lead contact / ward manager (*Who is the lead contact from your service? What do they do?*)

- (If this is not the leader) – *what was your experience of what was done / carried out? What would you have done differently?*

-Tactics that have been used, examples of how processes have been led (how to do it, how not to do it)

### WARD CULTURE

Prompts:

-*What does it feel like to work on this ward?*

-Staff engagement

-Group / individual motivation

-QI reminders

-QI awareness

## Appendix I: Readiness for change tool ([Bobiak et al., 2009](#))

1.	Degree of conflict among staff and clinicians.		very high	high	moderate	Low	very low
2.	Degree of conflict among staff.		very high	high	moderate	Low	very low
3.	Degree of staff friendliness toward patients.		very high	high	moderate	Low	very low
<b>Please rate how often the following are observed.</b>							
4.	There is a lot of tension between people in this unit.	almost always	very often	often	occasionally	Seldom	almost never
5.	People usually feel supported by the other people who work in this unit.	almost always	very often	often	occasionally	Seldom	almost never
<b>Please rate how strongly you agree or disagree with the following statements.</b>							
6.	Staff and clinicians put a great deal of effort into meeting the patients' expectations.		strongly agree	agree	neutral	Disagree	strongly disagree
7.	This unit has a shared vision among members.		strongly agree	agree	neutral	Disagree	strongly disagree
8.	People in this unit are willing to adjust usual routines in response to what is happening around them.		strongly agree	agree	neutral	Disagree	strongly disagree
9.	This unit is typically able to adapt new standards or procedures, even those forced upon us.		strongly agree	agree	neutral	Disagree	strongly disagree
10.	Changes people propose are usually shot down by other people in this unit.		strongly agree	agree	neutral	Disagree	strongly disagree
11.	People in this unit have similar opinions about important issues.		strongly agree	agree	neutral	Disagree	strongly disagree
12.	It's hard to make any changes in this unit because so busy seeing patients.		strongly agree	agree	neutral	Disagree	strongly disagree
13.	Non-clinical staff often receive incompatible requests from two or more people.		strongly agree	agree	neutral	Disagree	strongly disagree
14.	People in this unit take time to think about what is important to the unit.		strongly agree	agree	neutral	Disagree	strongly disagree
15.	This unit is generally willing to try new things to improve patient care.		strongly agree	agree	neutral	Disagree	strongly disagree
16.	This unit sees community resources as helpful.		strongly agree	agree	neutral	Disagree	strongly disagree
17.	Leadership in this unit articulates a vision for the unit.		strongly agree	agree	neutral	Disagree	strongly disagree
18.	This unit generally sees change as possible.		strongly agree	agree	neutral	Disagree	strongly disagree
19.	This unit's leadership promotes an environment that creates an enjoyable place to work.		strongly agree	agree	neutral	Disagree	strongly disagree
20.	This unit utilizes community resources to meet the health care needs of patients.		strongly agree	agree	neutral	Disagree	strongly disagree
21.	This unit tends to be very flexible in dealing with change.		strongly agree	agree	neutral	Disagree	strongly disagree
22.	This unit is able to stay on schedule.		strongly agree	agree	neutral	Disagree	strongly disagree
23.	People in this unit see the health care system as helpful in meeting the needs of the unit.		strongly agree	agree	neutral	Disagree	strongly disagree
24.	People in this unit are connected with people in other units.		strongly agree	agree	neutral	Disagree	strongly disagree
25.	Things change so fast in this unit it's hard to keep up with what's going on.		strongly agree	agree	neutral	Disagree	strongly disagree

## Appendix J: Permission to republish tool ([Bobiak et al., 2009](#))

### WOLTERS KLUWER HEALTH, INC. LICENSE TERMS AND CONDITIONS

This Agreement between Sarah Jasim ("You") and Wolters Kluwer Health, Inc. ("Wolters Kluwer Health, Inc.") consists of your license details and the terms and conditions provided by Wolters Kluwer Health, Inc. and Copyright Clearance Center.

License Number	3854990172072
License date	Apr 23, 2016
Licensed Content Publisher	Wolters Kluwer Health, Inc.
Licensed Content Publication	Quality Management in Health Care
Licensed Content Title	Measuring Practice Capacity for Change: A Tool for Guiding Quality Improvement in Primary Care Settings.
Licensed Content Author	Bobiak, Sarah; Zyzanski, Stephen; Ruhe, Mary; MPH, BS; Carter, Caroline; BA, LSW; Ragan, Brian; Flocke, Susan; Litaker, David; MD, PhD; Stange, Kurt; MD, PhD
Licensed Content Date	Jan 1, 2009
Licensed Content Volume Number	18
Licensed Content Issue Number	4
Type of Use	Dissertation/Thesis
Requestor type	Individual
Portion	Figures/table/illustration
Number of figures/tables/illustrations	2
Figures/tables/illustrations used	Figure 1 and Table 2
Author of this Wolters Kluwer article	No