High performance in colorectal surgery

A thesis presented for the degree of Doctor of Philosophy (PhD)

Benjamin Edward Byrne MA (Cantab), MB BChir

Centre for Patient Safety and Service Quality
Department of Surgery and Cancer
Imperial College London

May 2016
Abstract

A large body of international outcomes research has documented significant variation in the results of health care, beyond differences attributable to patient age, comorbidity or chance. Naturally, quality of care, however measured, may vary by provider within a health care system. Yet detailed understanding of the relationship between quality of care and health care outcomes continues to elude researchers. Much research has focused on the patient level, determining which patient-focused clinical processes deliver the best outcomes. By contrast, there is a relative lack of research examining intermediate and higher levels, to understand team performance and how teams work to provide high quality care, though research in this area is growing. This thesis aims to develop a greater understanding of how the best colorectal surgical units may be identified, and how they achieve their results.

Chapter 1 provides background to the present approach to the assessment of performance in health care. Chapter 2 summarises salient surgical outcomes research, and chapter 3 presents a literature review of evidence associating specific organisational structures and processes with clinical outcomes. Chapter 4 presents a patient questionnaire study, undertaken to assess the involvement of patients with gastrointestinal cancer in choosing a provider, and what provider-level information patients consider important. Chapters 5 to 8 describe the methods and results of a series of studies using routine administrative data to explore changes within colorectal surgery over time, as well as the relationship between different outcome measures at the unit level. Chapters 9 to 12 present research designed
to understand how units achieve their results. This work included developing a semi-structured interview to better understand the key organisational factors determining length of stay after elective colonic surgery. Chapter 13 summarises the main findings and limitations of this thesis, and discusses its implications for practice and future research.
# Table of contents

Abstract 3

Table of contents 5

List of tables 15

List of figures 19

Declarations 21

Acknowledgements 23

Thesis outputs 25

Abbreviations 29

## I  Introduction 31

1  Thesis summary and background 35

1.1  Chapter overview 35

1.2  Expanded summary of thesis 35

1.3  Historical context of performance assessment 37

   1.3.1  The ‘professional’ model - self assessment 38

   1.3.2  The context for change 39

   1.3.3  The ‘transparency’ model - increasing external assessment 41
# Table of contents

1.4 Performance *versus* quality ................................. 42
1.5 Why examine performance? ................................. 43
1.6 Defining high quality care ................................. 44
1.7 What to assess ........................................... 46
1.8 Frameworks for assessing performance ..................... 46
1.9 Chapter summary ......................................... 49

2 Outcomes in surgery ........................................ 51
2.1 Chapter overview ......................................... 51
2.2 Advantages and disadvantages of outcome measures ........ 51
   2.2.1 Advantages ........................................... 52
   2.2.2 Disadvantages ........................................ 55
2.3 Types and examples of surgical outcomes .................. 60
2.4 Unit-level surgical outcomes .............................. 61
   2.4.1 Unexplained variation ................................ 62
   2.4.2 Relationship between outcome measures .............. 63
2.5 Discussion ................................................ 65
2.6 Chapter summary ......................................... 67

3 Understanding organisational outcomes in surgery: a literature review ........................................ 69
3.1 Chapter overview ......................................... 69
3.2 Structure and process measures of performance ............ 70
3.3 Aims ....................................................... 71
3.4 Methods .................................................... 72
   3.4.1 Identification of key factors .......................... 73
3.4.2 Search strategy ........................................... 74
3.5 Selected results ............................................ 75
  3.5.1 Structural factors ...................................... 75
  3.5.2 Process factors ...................................... 79
3.6 Discussion .................................................. 87
3.7 Chapter summary ......................................... 89

Introduction summary and thesis questions 91

II Identifying patient preferences 95

4 Study 1: Decision-making and information needs about providers of surgical care 99
  4.1 Chapter overview ....................................... 99
  4.2 Introduction ............................................ 99
  4.3 Methods .................................................. 102
    4.3.1 Study design ....................................... 102
    4.3.2 Questionnaire design ................................ 102
    4.3.3 Data analysis ....................................... 104
    4.3.4 Ethical approval .................................... 106
  4.4 Results .................................................. 106
    4.4.1 Response rates and participant characteristics .... 106
    4.4.2 Provider choice .................................... 108
    4.4.3 Factors associated with experiences and preferences .... 112
    4.4.4 Information needs .................................. 113
    4.4.5 Relationship between decision-making and information needs 115
III Investigating colorectal outcomes 123

5 Using routine administrative data 127
  5.1 Chapter overview 127
  5.2 About the HES database 127
    5.2.1 Origins 127
    5.2.2 Contents 128
    5.2.3 Data strengths and weaknesses 130
  5.3 Using HES data 135
    5.3.1 HES data extraction and processing 135
    5.3.2 HES data analysis 138
  5.4 Chapter summary 142

6 Study 2: Mortality after colorectal cancer surgery between 1998 and 2012 143
  6.1 Chapter overview 143
  6.2 Introduction 144
  6.3 Methods 145
    6.3.1 Inclusion and exclusion criteria 146
    6.3.2 Definition of variables and outcomes 146
    6.3.3 Statistical analyses 148
  6.4 Results 150
    6.4.1 Laparoscopy 151
6.4.2 Screening ........................................... 155
6.4.3 Complication management ......................... 155
6.5 Discussion ............................................. 159
6.6 Chapter summary ...................................... 165

7 Study 3: Population-based cohort study comparing 30- and 90-day institutional mortality rates after colorectal surgery 167
7.1 Chapter overview ....................................... 167
7.2 Introduction ........................................... 167
7.3 Methods ............................................... 169
   7.3.1 Inclusion and exclusion criteria .................... 169
   7.3.2 Definition of variables and outcomes ............... 169
   7.3.3 Statistical analysis ................................ 171
7.4 Results ............................................... 171
   7.4.1 Cohort characteristics ............................ 171
   7.4.2 Mortality rates .................................... 171
   7.4.3 Factors associated with mortality .................. 173
   7.4.4 Relationship between mortality time periods ....... 173
7.5 Discussion ............................................. 181
7.6 Chapter summary ...................................... 185

8 Study 4: Outlier identification in colorectal surgery should separate elective and non-elective service components 187
8.1 Chapter overview ....................................... 187
8.2 Introduction .......................................... 187
8.3 Methods ............................................... 189
IV Investigating colorectal structures and processes 203

9 Study 5: Comparing short and long length of stay units using existing data 207

9.1 Chapter overview 207
9.2 Introduction 207
9.3 Methods 208
9.3.1 Unit selection 208
9.3.2 Data for comparison 210
9.3.3 Ethical considerations 211
9.4 Results 213
9.5 Discussion 214
9.6 Chapter summary 216

10 Study 6: Understanding length of stay in elective colorectal surgery - methods and qualitative analysis 219
10.1 Chapter overview ........................................... 219
10.2 Introduction .................................................. 219
10.3 Methods ....................................................... 222
   10.3.1 Overall design ........................................ 222
   10.3.2 Choice of methodology ................................. 223
   10.3.3 Unit selection ........................................... 224
   10.3.4 Interview development ................................. 225
   10.3.5 Interviewee selection ................................. 230
   10.3.6 Interview conduct ..................................... 231
   10.3.7 Data analysis ........................................... 231
   10.3.8 Ethical considerations ................................. 232
10.4 Results ....................................................... 233
   10.4.1 Participation ........................................... 233
   10.4.2 Descriptive unit characteristics ....................... 234
   10.4.3 Qualitative results .................................... 234
10.5 Discussion ................................................... 241
10.6 Chapter summary ............................................ 246

11 Study 6: Understanding length of stay in elective colorectal surgery
   - exploratory quantitative analysis ......................... 247
11.1 Chapter overview ........................................... 247
11.2 Introduction .................................................. 247
11.3 Methods ....................................................... 249
   11.3.1 Statistical analysis .................................... 249
11.4 Results ....................................................... 250
11.4.1 Distribution of scores 251
11.4.2 Overall scores 251
11.4.3 Scores by professional group 251
11.4.4 Scores by performance status 252
11.4.5 Scores by profession and performance 254
11.5 Discussion 257
11.6 Chapter summary 260

12 Study 6: Post hoc analysis - comparison of participant and non-participant units using existing data 263
12.1 Chapter overview 263
12.2 Introduction 263
12.3 Methods 264
12.4 Results 264
12.5 Discussion 265
12.6 Chapter summary 266

V Discussion and conclusions 267

13 Discussion 271
13.1 Chapter overview 271
13.2 Key findings 271
13.3 Limitations 275
13.4 Understanding high performance 277
13.5 Implications and future work 279
13.6 Concluding remarks 285
Table of contents

F.1 Appendix overview ........................................... 371
F.2 Expert interviews ........................................... 371
F.3 Peer interviews ............................................. 373
F.4 Appendix summary .......................................... 375

G Study 6: interview protocol 377
G.1 Appendix overview .......................................... 377

H Study 6: quantitative analysis supplement 393
H.1 Appendix overview .......................................... 393
H.2 Distribution of overall scores ............................. 393

I Third party permissions 397
I.1 Appendix overview .......................................... 397
List of tables

1.1 Structures, processes and outcomes at various organisational levels within a hospital. .......... 49

2.1 Advantages of outcomes as performance measures. .............. 53
2.2 Disadvantages of outcomes as performance measures. ............ 56
2.3 Outcomes relevant to an operation for colonic cancer. .......... 60

3.1 Structures and processes examined in literature review. ............ 73

4.1 Single-item Control Preferences Scale ......................... 103
4.2 Respondent characteristics. ......................... 107
4.3 Provision of choice of provider at different decision points. ........ 108
4.4 Decision-making experiences and preferences .................... 109
4.5 Decision-making experience and preference matrices. ............ 110
4.6 Decision-making role matching ......................... 111
4.7 People with whom participants would like to discuss choice. ......... 111
4.8 Significance of variables associated with experiences or preferences. 113
4.9 Importance ratings of information items ......................... 114

5.1 Basic HES data processing steps and illustrative numbers. .......... 137

6.1 OPCS codes for colorectal procedures selected for inclusion. ........ 146
6.2 OPCS codes used to identify surgical failure to rescue procedures. 147
6.3 Resections and deaths by year and urgency. ......................... 151
6.4 Model results for laparoscopy rates over time. .......................... 152
6.5 Model results for mortality rates over time, by urgency of admission,
surgical approach and sex of patient. ................................. 153
6.6 Model results for mortality rates over time, by urgency of admission,
age group and sex of patient. ........................................... 157
6.7 Model results for surgical failure to rescue rates over time. ........ 157
6.8 Model results for mortality rates over time, by urgency of admission,
surgical failure to rescue and sex of patient. ......................... 159
7.1 Colorectal operations selected for inclusion. ............................ 170
7.2 Patient characteristics for cohort and different time-based mortality
subgroups. ........................................................................... 172
7.3 Multiple logistic regression results for each postoperative time period. 174
7.4 Numbers of outlying units identified by postoperative time period. . 180
7.5 Correlation of observed-to-expected mortality across different post-
operative time periods. ....................................................... 180
8.1 Colorectal operations selected for inclusion. .............................. 190
8.2 Patient characteristics for cohort and urgency-based subgroups. .. 192
8.3 Multiple logistic regression results for elective and non-elective co-
horts. ............................................................................... 193
9.1 Data compared between short and long stay units. .................... 212
9.2 Length of stay results and participation of 20 units approached. .. 213
9.3 Difference between low and high LLoS units. ............................ 214
10.1 Patient pathway factors important in determining the performance
of a colorectal unit. ......................................................... 226
10.2 Over-arching factors important in determining the performance of
a colorectal unit. ......................................................... 227
10.3 Areas of care organisation examined in the interview. ................. 230
10.4 Unit participation against length of stay results. ......................... 233
10.5 Unit characteristics. .................................................. 235

11.1 Descriptive statistics for interview scores. ............................... 252
11.2 Descriptive statistics for interview scores, by profession. ............... 253
11.3 Descriptive statistics for interview scores, by performance. ............ 255
11.4 Average interview total by profession and performance. .............. 256

12.1 Differences between participating and non-participating units. ....... 265

A.1 Example literature search. ........................................... 338
A.2 Papers retrieved from example search. ................................ 339

D.1 Resections and deaths by year, urgency, surgical approach and sex
of patient. ................................................................. 362
D.2 Resections and deaths among screening age patients by year, ur-
gency and sex of patient. .............................................. 363
D.3 Resections and deaths among non-screening age patients by year,
urgency and sex of patient. ............................................ 364
D.4 Resections and deaths by year, urgency, surgical failure to rescue
and sex of patient. ....................................................... 365
List of figures

1.1 Expanded structures, processes and outcomes framework. . . . . . . . 48

2.1 Well-being, illness and surgical intervention. . . . . . . . . . . . . 53

6.1 Rates of laparoscopic surgery against year, by urgency of admission
   and sex of patient. . . . . . . . . . . . . . . . . . . . . . . . . . . . 152

6.2 90-day mortality rates against year, by urgency, surgical approach
   and sex of patient. . . . . . . . . . . . . . . . . . . . . . . . . . . . 154

6.3 90-day mortality rates against year, by urgency, age group and sex
   of patient. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 156

6.4 Rates of surgical failure to rescue procedures against year, by ur-
   gency of admission and sex of patient. . . . . . . . . . . . . . . . 158

6.5 90-day mortality rates against year, by urgency, surgical failure to
   rescue and sex of patient. . . . . . . . . . . . . . . . . . . . . . . . 160

7.1 Postoperative 30-day mortality funnel plot. . . . . . . . . . . . . . 175

7.2 Postoperative 90-day mortality funnel plot. . . . . . . . . . . . . . 176

7.3 Postoperative 180-day mortality funnel plot. . . . . . . . . . . . . . 177

7.4 Postoperative 365-day mortality funnel plot. . . . . . . . . . . . . . 178

8.1 Postoperative elective mortality funnel plot, elective outliers marked.194

8.2 Postoperative non-elective mortality funnel plot, elective outliers
   marked. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 195
8.3 Postoperative non-elective mortality funnel plot, non-elective outliers marked. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 196
8.4 Postoperative elective mortality funnel plot, non-elective outliers marked. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 197
11.1 Scatter plot of doctors’ and nurses’ interview scores. . . . . . . . . . 254
H.1 Frequency histograms of scores for questions 2 to 7. . . . . . . . . . 394
H.2 Frequency histograms of scores for questions 8 to 12. . . . . . . . . 395
H.3 Frequency histogram of total interview scores. . . . . . . . . . . . . 396
Declarations

I declare that the work presented in this thesis is my own, and that all contributions from other sources have been appropriately acknowledged.

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.
Acknowledgements

I would like to thank Professor Charles Vincent for his fantastic supervision and sage advice, giving me freedom to explore my ideas while gently steering me towards producing this thesis. I would also like to thank Mr Omar Faiz for his guidance and support, without which this thesis would not have been realised.

I must also thank my friends and colleagues in the Centre for Patient Safety and Service Quality, and at the Surgical Epidemiology, Trials and Outcomes Centre at St Mark’s. A sense of humour and friendly faces have been invaluable in tackling the challenges involved, and enjoying my three years at Imperial.

The origins of this journey may be traced back to a present my father bought me just before I got married - Atul Gawande’s ‘The Checklist Manifesto’. Unusually for a book I received from my father, I actually read this while on honeymoon in 2010. It sparked an interest in quality and safety research that continues to shape my career. So I am grateful to both my parents for their support and influence on me.

Lastly, I am immensely lucky to have such a wonderful, supportive and understanding wife. She was happy for me to pause my surgical career for three years, allowing me to indulge my curiosity and develop my interest in clinical research. This has now progressed to a full-scale relocation from Winchester to Bristol, so that I can pursue a combined clinical and academic career. Thank you!
Thesis outputs

Publications

Peer-reviewed papers


Presentations

Oral

International


National

• Byrne B E, Mamidanna R, Vincent C A, Faiz O D. Variation in outcomes between elective and non-elective colorectal surgery: a population-based cohort

**Regional**


**Poster**

**International**


**National**


Regional


**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACS</td>
<td>American College of Surgeons</td>
</tr>
<tr>
<td>CCI</td>
<td>Charlson Comorbidity Index</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>ERP</td>
<td>Enhanced Recovery Program</td>
</tr>
<tr>
<td>FCE</td>
<td>Finished Consultant Episode</td>
</tr>
<tr>
<td>GI</td>
<td>Gastrointestinal</td>
</tr>
<tr>
<td>HES</td>
<td>Hospital Episode Statistics</td>
</tr>
<tr>
<td>HMU</td>
<td>High Mortality Unit</td>
</tr>
<tr>
<td>HSCIC</td>
<td>Health and Social Care Information Centre</td>
</tr>
<tr>
<td>IBD</td>
<td>Inflammatory Bowel Disease</td>
</tr>
<tr>
<td>ICD</td>
<td>International Classification of Diseases and Related Health Problems</td>
</tr>
<tr>
<td>LLoS</td>
<td>Long length of stay</td>
</tr>
<tr>
<td>LMU</td>
<td>Low Mortality Unit</td>
</tr>
<tr>
<td>M&amp;M</td>
<td>Morbidity and Mortality</td>
</tr>
<tr>
<td>NHS</td>
<td>National Health Service</td>
</tr>
<tr>
<td>NSQIP</td>
<td>National Surgical Quality Improvement Program</td>
</tr>
<tr>
<td>O/E</td>
<td>Observed-to-Expected</td>
</tr>
<tr>
<td>ONS</td>
<td>Office of National Statistics</td>
</tr>
<tr>
<td>OPCS</td>
<td>OPCS Classifications of Interventions and Procedures</td>
</tr>
<tr>
<td>OR</td>
<td>Odds Ratio</td>
</tr>
<tr>
<td>RTMv8</td>
<td>Real Time Monitoring version 8</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>US</td>
<td>United States</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>VA</td>
<td>Veterans Affairs</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organisation</td>
</tr>
</tbody>
</table>
Part I

Introduction
Section overview

The first chapter of this introductory section presents a summary of the thesis and explains why the work presented was needed. It places the thesis in historical context, and considers the theoretical underpinnings of the approach taken to better understand colorectal surgical performance. Chapters two and three in this section move on to examine selected relevant evidence that helped direct the original work presented in later sections. At the end of this section, I summarise the material covered in the first three chapters, and present three central questions which the subsequent work attempts to answer.
Chapter 1

Thesis summary and background

1.1 Chapter overview

This introductory chapter presents an expanded summary of this thesis, describes the important background to the work that has been undertaken, and briefly considers the individual studies reported. It places this thesis in historical context, describing how the approach to examining performance in health care has changed over time. The definition of quality of care is considered, and the key theoretical framework upon which this thesis rests is presented and discussed.

1.2 Expanded summary of thesis

It has been well established by other researchers that the outcomes of surgical care vary significantly between providers. This important observation may provide the opportunity to better understand the key determinants of the results of care, by studying units with different outcomes. It may also be used as a lever to try and drive improvements in care by promoting competition among providers. Provider-level performance is important, and this thesis aims to contribute towards a better understanding in this field.

The first three chapters of this thesis provide the background and context for
the original studies presented. Chapter one reviews how performance has been assessed within the health care setting over the last hundred years, highlighting how performance assessment has changed, and the reasons for this. It also defines the distinction between the terms ‘performance’ and ‘quality of care’, as used in this thesis, and explains the key framework upon which this thesis is based. Chapters two and three review relevant literature, specifically focused on unit-level performance. Chapter two describes selected outcomes research, relevant to colorectal surgical performance, and chapter three presents a review of the literature examining evidence linking specific structures and processes to department-level outcomes.

After these introductory chapters, the six original studies included in this thesis will be described. The first study is a patient questionnaire designed to determine whether patients are, or would like to be, involved in choosing their provider, and what provider-level information they consider important. This study helps understand whether the political emphasis on patient choice, promoted in part as a way of driving improvement through competition, is operating effectively at the provider level. It also helps assess whether the priorities of patients, when considering provider performance, are similar to those of the clinical and research communities.

Studies two, three and four utilise the Hospital Episode Statistics (HES) database, a national administrative dataset, to explore and better understand postoperative mortality after colorectal surgery in England. Study two examines the fall in mortality after colorectal surgery for cancer between 1998 and 2012. It attempts to better understand the role of some of the most salient developments, such as laparoscopic surgery and bowel cancer screening, in contributing to the significant
improvements seen. Study three explores the impact of varying the definition of postoperative mortality on the stratification of unit-level performance, by changing the time period during which deaths are included in this measure. The fourth study examines the relationship between the outcomes achieved for patients undergoing colorectal surgery within the same institution according to the urgency of their admission. Together, these studies help refine the present understanding of the complexity of unit-level performance in colorectal surgery.

The fifth and sixth studies present original work to try and understand how colorectal units achieve high or low outlying risk-adjusted length of stay after surgery. Study five explores whether existing, routinely collected NHS data can be used to help understand the different levels of performance achieved among the selected units. The sixth study uses a novel, semi-structured telephone interview, to develop a deeper understanding of how the organisation of care within the study units determined their outcomes. The interviews yielded both qualitative and quantitative data for analysis, which are presented in successive chapters.

The final chapter presents a discussion of the findings of these studies, and their implications for future research and clinical practice.

The next section of this first introductory chapter will explain the history of performance assessment, before considering why and how we should examine health care performance.

1.3 Historical context of performance assessment

While medical practice has always been subject to some form of scrutiny, there have been significant changes in the way performance is monitored over the last
hundred years. This section discusses the historical approach to performance assessment, the context that drove significant changes, and the contemporary model that applies to practice in the early 21st century.

1.3.1 The ‘professional’ model - self assessment

Until the later decades of the 20th century, the assessment of performance within healthcare was characterised by the features that define a profession. Doctors had privileges of autonomy and societal standing, with associated responsibilities for self-regulation and assessment. Their performance was not subject to public scrutiny, but rather that of their peers. The ‘Morbidity and Mortality’ (M&M) meeting is a good example of this type of peer-based regulation, and has a strong tradition in the surgical specialties. M&M meetings involve the presentation of individual patients who have experienced a complication or died whilst under the care of a specific consultant or team within a set period of time. The presentation is often given by junior members of the team to the rest of the local surgical department. Relevant published literature may be presented to highlight learning points. The attending doctors discuss the cases presented to establish if the results of the care provided could have been improved. While attitudes to M&M meetings are generally positive\(^1\), they may be subject to bias, including under-reporting of adverse events such as deaths or unplanned return to theatre\(^2\). Such meetings may allow appropriate and thorough assessment of the performance of a surgeon or team, but they lack rigor and may be significantly influenced by local dynamics.
1.3.2 The context for change

The degree of autonomy afforded to the medical profession was arguably supported by the continued progress made by medical science during the 20th century. The first antibiotic, penicillin, was successfully isolated in 1940, and by the end of the World War II penicillin was being produced on an industrial scale. Subsequently, a number of other antibiotics were developed that together provided a range of effective treatments for bacterial infections. The first successful human dialysis machine was used in 1945. In 1953 Watson and Crick described the structure of DNA. The first mid- and long-term successes with solid organ transplantation were achieved in the 1950s and 1960s. Against this backdrop of demonstrable progress, the privilege of autonomy was not questioned as medicine appeared to be making significant advances and delivering improved care over time.

Towards the end of the 20th century, one may find the beginnings of a shift towards increasing external and public scrutiny of the performance of health care providers. In the United States (US), the Health Care Financing Administration began publishing hospital-level mortality data in 1986. This must be placed in the context of a privatised health care system, driven by open competition. In such an environment, information on provider performance was sought to inform the purchasing of health care from hospitals by insurance companies and state-funded schemes such as Medicare and Medicaid.

The National Health Service (NHS) in the United Kingdom (UK) has seen a gradual introduction of increased competition between providers, and greater emphasis on patient choice. After the 1989 white paper, Working for Patients, the Conservative government introduced an internal market through the separation
of health care authorities’ purchasing and provider roles. It was hoped that this would help limit spiralling costs and bring down waiting lists. In 2000, The NHS Plan outlined the Labour government’s plans for the NHS\textsuperscript{10}. This included increasing involvement of private sector providers and a focus on empowering patients. Among other things, this required more information for patients and strengthening patient choice. An electronic clinic appointment booking system called ‘Choose and Book’, introduced in 2004, exemplified the drive to increase patient choice. In the 2008 publication \textit{High Quality Care For All}, Lord Darzi described the major challenge facing the NHS at that time to be raising the standard of care across all aspects of its services\textsuperscript{11}. One component of tackling variation in the quality of care would be providing patients with greater information and choice\textsuperscript{11}. In 2009, \textit{The NHS Constitution} enshrined patients’ rights to informed choice when selecting a GP practice, a particular doctor at a GP practice, and other broader choices about NHS care\textsuperscript{12}. The choices made by patients may in many instances reflect access and convenience, but it has been repeatedly recognised that facilitating truly informed choice requires information about providers.

Besides the shift in focus towards patient choice and the accompanying need for public information on provider performance, a number of high profile failures of the ‘professional’ model of self-assessment and regulation provided further impetus to the drive for change.

In the late 1990s a scandal surrounding the quality of care received by paediatric cardiac surgical patients at the Bristol Royal Infirmary received great media attention and resulted in an extensive public inquiry\textsuperscript{13}. The inquiry examined practice at the hospital between 1984 and 1995, and it was estimated that 30 to 35 more children under 1 year of age died between 1991 and 1995 than would have
been expected at a similar hospital in England at that time. The report documented failings of individuals and organisations, including the lack of a system to assess the quality of care delivered.

In 1998, a general practitioner, Harold Shipman, was arrested and later convicted of murdering 15 of his patients, though it is suspected he was responsible for the death of many more. The case attracted significant press coverage and an inquiry was set up after his conviction in 2000. The inquiry found that the system of death certification was fragmented and did not function as was expected or required\textsuperscript{14}.

More recently, the Mid Staffordshire NHS Foundation Trust has been subject to two inquiries to investigate the quality of care it provided between 2005 and 2009\textsuperscript{15,16}. The most recent report highlighted a culture that tolerated poor standards and risk, where communication was poor and organisational targets were pursued over patient care.

Such repeated failures of health care have provided the context for a move away from self-regulation towards a system for evaluating the performance of health care that is more open and transparent.

1.3.3 The ‘transparency’ model - increasing external assessment

Over the last decade, an increasing number of formalised systems have been developed to assess the performance of health care professionals and institutions. These have placed increased emphasis on the scrutiny of performance by individuals outside the institution under examination. Local processes include mandatory
Thesis summary and background

appraisal and audit, as well as the traditional M&M meetings. National systems include peer review visits, voluntary disease registries and publication of aggregated outcome data by the NHS or independent organisations such as Dr Foster Intelligence. In addition, the first cycle of periodic revalidation of practitioners who have completed their training began in April 2013 and will be completed by the end of March 2016\(^1\). In December 2012, the NHS Commissioning Board announced plans to publish survival rates and other quality measures for all consultants practising in each of 10 primarily surgical specialties from Summer 2013\(^2\). These many and varied processes indicate an opening up of the practice of health care to assessment by a greatly increased number of people, with increasing amounts of data entering the public domain.

Within the UK, cardiac surgeons have been early adopters of the open publication of their results. This may well have been significantly influenced by the Bristol Royal Infirmary Inquiry, as one of its key recommendations was to develop rigorous systems for internal and external monitoring of outcomes. The Society for Cardiothoracic Surgery in Great Britain and Ireland have been publishing hospital- and consultant-level mortality rates after cardiac surgery since 2005. Since then, improvements in mortality results have been demonstrated, with no evidence of adverse consequences due to the publication of outcome data\(^3\). Only time will determine whether the same will be found for other specialties.

1.4 Performance versus quality

It is important to define how these terms will be used in this thesis. The term ‘performance’ will be used to denote the delivery of, and / or the results achieved
by health care. There are a number of potential factors that may exert important influences on performance, such as an individual patient’s age or frailty. The term ‘quality’ will be used to indicate the standard to which care is delivered. Quality is therefore a factor that may influence outcomes, and it is an aspect of care that is under the control of the health care provider. Once other important factors have been taken into consideration, such as a patient’s age, performance may be considered a reflection of the quality of care provided. A more detailed consideration of the definition of ‘quality’ is provided in section 1.6.

1.5 Why examine performance?

The purpose of health care is to treat patients. At its simplest, examination of the performance of health care is undertaken to ensure patients receive a good standard of care. However, deciding what represents ‘good’ care is not straightforward. Quality of care has been variously defined, and some of the most widely used definitions are discussed in section 1.6.

It is known that there is wide variation in the results of health care. Many have attributed a proportion of this variation to differences in quality of care. For example, Birkmeyer and Dimick argued that ‘variation can be attributed to three contributing factors: chance, case mix, and quality of care.’ Statistical techniques can be used to make allowance for measured differences in case mix, such as the age and comorbidity of patients. The role of chance can also be quantified statistically. After applying such techniques, residual variation may then be attributed to differences in quality of care. However, some authors have challenged this. In a systematic review of the relationship between risk-adjusted mortal-
ity rates and quality of care, Pitches and colleagues proposed three reasons why mortality may vary even after allowing for case mix and chance: differences in risk not already captured (unmeasured or confounding variables); differences in how variables or outcomes are defined, or how definitions are applied; and genuine differences in quality of care\textsuperscript{33}. While the majority of clinicians and researchers may agree with the former stance, that there is a link between variation in risk-adjusted outcomes and quality of care, it is important to remember that this view is not held by all. If the validity of such a link is accepted, the observed variation in outcomes presents an opportunity to better understand healthcare performance by examining practice in different units. A better understanding of this variation may help reduce inequalities in healthcare provision and inform global quality improvement.

1.6 Defining high quality care

Given that the purpose of assessing the performance of healthcare is to ensure patients receive ‘good’ care, it is important to consider what is defined as ‘good care’. The World Health Organization (WHO) described six domains of healthcare provision that pertain to quality\textsuperscript{35}. High quality care should be: effective; efficient; accessible; acceptable and patient-centred; equitable; and safe. The Institute of Medicine described quality of care as

\[ \text{...the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge.}\textsuperscript{36} \]

Others have provided alternative but similarly broad definitions of quality of
General definitions of high quality care such as these provide a starting point for the examination of quality specific to certain conditions or procedures. While all of the domains described by the WHO may be relevant to a particular speciality, specific considerations may vary. For example, the safety considerations relevant to a drug that is prescribed to reduce the risk of a heart attack for a particular patient over a 10-year period will be very different to those for a procedure to remove a large cancer of the bladder. As the focus of study becomes increasingly specific, the definition of quality may become inextricably linked with how it is measured. For colorectal surgery, McGory et al utilised the Delphi technique to develop a list of 92 quality indicators across 6 domains. Disease- or condition-specific quality criteria such as these may dictate the methods used to assess performance within a speciality.

To date, the majority of medical research has focused on care provided at the level of individual patients, for example, identifying the most effective treatment for a specific condition. Over recent years, increasing attention has been given to other aspects of the quality of care, such as safety and patient experiences of treatment. There has also been a broadening of the scope of health care research, with growing recognition of the importance of meso- and macroscopic influences on clinical outcomes. The work in this thesis is focused on understanding performance at the intermediate, unit level, and a broad range of influences on performance are considered. If researchers and providers of health care truly seek to improve quality of care, efforts must be directed towards understanding all the facets of performance implicated in the broad definition of quality.
1.7 What to assess

As the definition of quality is broad, it is unlikely that the overall quality of care provided by an institution can be described by a single indicator of performance\textsuperscript{20,40,41}. Rather, provider performance should be assessed across a wide range of features. For colorectal surgery, these may include outcomes such as short- and long-term mortality and morbidity, in-hospital length of stay, patient reported outcome measures, wait times and care costs.

Different parties may have different interpretations of quality and what aspects are important to them. A patient may be most interested in minimising wait time for treatment and avoiding disabling complications. A doctor may combine these concerns with an emphasis on indices of high performance such as short length of stay or cancer-specific results. A hospital manager may be under pressure to ensure resource efficiency and meeting of local or national targets. Specific domains within the broader definition of quality may therefore be given greater or lesser priority by different people.

1.8 Frameworks for assessing performance

Ernest Codman has been widely credited as one of the first people to scrutinise the performance of health care according to ‘modern’ principles. During the 1910s, he developed his ‘End Results System’, suggesting that hospitals and doctors should follow every patient treated, for as long as required, to determine if the care provided had been successful or not\textsuperscript{42}. Integral to this approach was an assessment of unsuccessful results to facilitate learning. His system challenged the existing
traditions of paternalism and nepotism within medicine, and was not popular among his contemporaries. However, over the following decades, this approach to understanding the performance of health care gained increasing acceptance.

In 1966, Avedis Donabedian described a framework for studying the performance of health care organisations which followed Codman’s paradigm. According to Donabedian, the key elements of a health care organisation’s performance were its structures, processes and outcomes. Structures may be considered as the environment in which health care is delivered, in the broadest sense. This includes the physical environment in which health care workers and patients interact, the financial context, and the experience and qualifications of the staff. Processes are the individual interactions between health care workers and patients. Examples include a doctor examining a patient in clinic, or the team briefing in an operating theatre before an operation starts. Donabedian referred to outcomes as ‘the end results’, drawing on Codman’s principles. Outcomes range from death rates to length of hospital stay and patient satisfaction ratings. Structures effectively form the preconditions for health care, shaping and informing the processes that represent the actual delivery of medical care to patients. Through these processes, patients respond to the treatment provided and measurable outcomes result. Donabedian’s framework has been widely applied to the assessment of health care since its initial description.

Lilford et al further expanded on the three elements of organisational performance described by Donabedian (see figure 1.1). They grouped different elements according to the level at which they were relevant within an institution. Some processes, such as local appraisal of a doctor’s performance, are of particular relevance at the organisational level, whereas clinical processes may be
considered as patient-specific, for example the administration of antibiotics before an operation. Likewise, certain outcomes, such as overall waiting times, are of importance macroscopically, whereas individual complications are of greatest impact on particular patients. The primary focus of this thesis is to examine speciality-specific performance at the departmental level. This represents an intermediate level, when considered within the context of hospital-based care. The different levels and components of care may be summarised as shown in table 1.1. The frameworks discussed here provide the central conceptual structure upon which this thesis is based.

Figure 1.1: Expanded structures, processes and outcomes framework. Reproduced from Lilford et al with permission. 

---

48 Thesis summary and background
Table 1.1: Structures, processes and outcomes at various organisational levels within a hospital.

## 1.9 Chapter summary

This chapter has summarised the outline of this thesis, and placed the original work that will be presented in later chapters in historical context. The central assumption and theoretical framework upon which this thesis is based have been described. To recap, the research presented is largely predicated on the assumption that some proportion of the unexplained variation in surgical outcomes may be attributed to differences in quality of care between providers. Once patient factors and chance have been taken in to consideration, outcomes may therefore be considered as measures of quality of care. The approach taken to try and better understand the observed variation in performance is based upon Donabedian’s structures, processes and outcomes framework. Studies two to four focus on examining outcomes, while studies five and six focus on key structure and process determinants of outcomes. In this way, the studies presented combine to help develop a deeper understanding of unit-level performance in surgery.
Chapter 2

Outcomes in surgery

2.1 Chapter overview

In chapter 1, Donabedian’s three key facets of health care performance (structures, processes and outcomes) were explained. The outcomes of care appear to represent the most useful indicators of performance for patients and clinicians. However, their direct link to specific areas of clinical practice can be difficult to determine. This can make it challenging to know how to address poor outcomes, or how to identify good practice where excellent results have been achieved. This chapter discusses both the advantages and disadvantages of outcomes as measures of performance, and then summarises selected, relevant department-level outcomes research within the surgical literature.

2.2 Advantages and disadvantages of outcome measures

Of the three domains of health care performance described by Donabedian, the outcomes of health care are usually the criteria against which all other aspects are measured. There are good reasons for this, and there are specific advantages to outcomes as performance metrics, when compared with structure and process
measures. However, there are also problems with an excessive focus on outcomes, when examining structure or process aspects of performance may be more helpful in directing clinical care, in practical terms. Some of the advantages and disadvantages of outcome measures are discussed below.

### 2.2.1 Advantages

As measures of performance, outcomes have a number of advantages, summarised in table 2.1. In accordance with Codman’s ‘End results’, outcomes represent the end of the journey for patients who need health care. Many interventions offered by modern medicine have clear immediate effects to further worsen an unwell patient’s symptoms or quality of life, before it can be hoped that an improvement is seen (illustrated in figure 2.1). This is particularly relevant to major surgery. The physical nature of operating on a person, cutting through layers of skin, muscle and other organs, before suturing them back in place, carries obvious short-term consequences and risks. Patients are willing to undergo such gross interventions, and clinicians are willing to perform them, for the results that are achieved once a recovery has occurred, days, weeks or perhaps even months later. Considered in this light, it is not surprising that outcomes have been a central focus of scientific research in medicine. Outcomes, such as postoperative mortality or length of stay, have obvious importance and relevance, or face validity, as measures of performance. By association, they may also therefore be useful indicators of quality of care.

Most outcomes are amenable to quantitative assessment, and this also contributes to their appeal. Health care has become increasingly complicated over
<table>
<thead>
<tr>
<th>Advantage</th>
<th>Explanation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face validity</td>
<td>Have obvious relevance to those involved in care</td>
<td>Postoperative death within 30 days</td>
</tr>
<tr>
<td>Quantitative</td>
<td>Measured numerically</td>
<td>Average length of stay after surgery</td>
</tr>
<tr>
<td>Easily measured</td>
<td>Determination of outcome does not require specific training or expertise</td>
<td>Postoperative death</td>
</tr>
<tr>
<td>Objectivity</td>
<td>Observation and recording of outcome unlikely to be influenced by observer bias</td>
<td>Postoperative death</td>
</tr>
</tbody>
</table>

Table 2.1: Advantages of outcomes as performance measures, with explanation and examples.

Figure 2.1: Well-being through illness and surgical treatment. Adapted from Vincent with permission\textsuperscript{44}. 

recent decades, involving ever-growing teams of clinicians and support staff, with ever greater reliance on sophisticated technologies. Representing its performance with a single numerical measure, such as in-hospital mortality rate, or a limited selection of metrics, appears to make such complexity easy to understand for all interested parties, from clinicians and patients to managers and politicians.

Outcomes are generally easier to measure than structures or processes, and some can be measured retrospectively. It is an obvious statement that any observer can determine if a patient lived and was discharged, or died in hospital after treatment. The definition of ‘death’ is not contentious. Mortality data can therefore be collected retrospectively without excessive concern about its measurement. Other outcomes, such as length of hospital stay or readmission to hospital, are also easy to measure.

Outcomes appear to be objective indicators of performance. Not only is it typically easy to determine if a patient died after surgery, this observation is so plain that observer bias is unlikely to affect its measurement. However, this appearance belies the practical truth of the measurement and reporting of outcomes. Problems with the objectivity of outcomes, such as postoperative mortality rates, are discussed in section 2.2.2.

Outcomes have formed the basis of the assessment of most, if not all, major developments in health care. The methodology of randomised controlled trials may be considered a formal system for isolating and varying a single or limited number of processes, and observing the effect of that process on a pre-defined set of outcomes. Much organisational-level research has focused on a limited number of outcomes as primary indicators of quality. Many studies have then examined factors associated with outcomes as a method of understanding the determinants
of high quality care. When the face validity, quantitative nature, easy measurement and apparent objectivity of many common outcome measures are considered together, it is not surprising that they have had such central importance in the medical literature.

### 2.2.2 Disadvantages

The importance of outcomes should not be overlooked, but there are a number of reasons why they might not always be the optimal measure of performance. Some of the disadvantages of outcomes are summarised in table 2.2.

Outcomes represent the end result of the process of health care, and do not directly lend understanding to how particular results can be achieved. The activities that comprise health care, Donabedian’s processes, are often temporally separated from the determination of their outcome. Interpreting and understanding outcomes can therefore be challenging. For example, some may consider in-hospital mortality to be a marker of the quality of surgery provided in theatre. If a particular patient dies 12 days after an operation, it must be considered that the patient’s death may be related to care received at the time of their operation. However, any number of events that could have directly caused or contributed to the patient’s death may have occurred during the intervening period. The outcome and process under consideration are not always closely coupled, and a large number of other variables may compete with or supercede the importance of the operation in determining a fatal outcome. While there may be a reasonably strong underlying causality, the presence of intervening variables may obscure that relationship. Conversely, this loose coupling may be used by some to ‘explain away’ an apparent
Table 2.2: Disadvantages of outcomes as performance measures, with explanation and example.

<table>
<thead>
<tr>
<th>Disadvantage</th>
<th>Explanation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of explanatory power</td>
<td>Difficult to identify key clinical activities associated with a particular outcome</td>
<td>Postoperative death due to a postsurgical infection</td>
</tr>
<tr>
<td>Apparent objectivity misleading</td>
<td>Not immediately clear that a particular outcome is associated with a particular clinical activity</td>
<td>Problems with specificity of postoperative care</td>
</tr>
<tr>
<td>Over-simplify performance of outcomes</td>
<td>Performance domains include efficacy, efficiency and patient-centred care, and cannot be represented with a single outcome</td>
<td>Inpatient outcome reporting of 90-day mortality</td>
</tr>
<tr>
<td>Difficulty to improve</td>
<td>Performance becomes increasingly difficult to improve with a single outcome</td>
<td>Low postoperative mortality following colorectal cancer surgery</td>
</tr>
<tr>
<td>Sample size considerations</td>
<td>Uncommon or rare events require large sample sizes to reliably measure performance</td>
<td>A 2-arm randomised trial would require 1810 patients to have 80% power to show a difference in absolute mortality if p &gt; 0.05</td>
</tr>
</tbody>
</table>

*Table 2.2: Disadvantages of outcomes as performance measures, with explanation and example.*
relationship. A focus on processes is required to help guide practitioners in their daily practice.

The apparent objectivity of outcomes can be misleading. Even mortality rates can be subject to variation in definition by time period or by cause of death. Common measures of post-operative mortality used in outcomes research include in-hospital, 30- and 90-day mortality. The inclusion of deaths within a specific time-period is arguably arbitrary, determined only by convention rather than based on specific scientific considerations. Furthermore, the decision to use a particular metric for a specific piece of work may be influenced by conscious or unconscious bias, introducing a degree of subjectivity that may not at first be appreciated. In addition, reporting bias may result in under-estimation of death rates. Other outcomes may be much more difficult to define. For example, the diagnosis of certain post-operative complications, such as a chest infection, can be clear and easy; however, on some occasions this diagnosis may require careful clinical judgement. The stress that major surgery places on the body often results in a high temperature during the following 48 hours. A four-hour general anaesthetic may be expected to result in a degree of collapse of the small airways in the lungs, with associated retention of mucus, which may be heard as crackling sounds using a stethoscope. It can therefore be difficult to decide if a patient seen 48 hours after major surgery has a chest infection, as there are other good reasons for such a patient to have a high temperature and signs of a chest infection when examined. One approach to resolving such difficulties may be to develop comprehensive diagnostic criteria, using expert consensus methods. For example, the International Study Group of Rectal cancer published a clear definition and classification of anastomotic leakage after anterior resection, a specific complication encountered
Outcomes in surgery

Another strategy is to classify complications by the treatment required. This obviates the need to have an exhaustive list of definitions for all eventualities. This approach is taken by the Clavien-Dindo system for grading complications\textsuperscript{46,47}. However, the success of either approach in bringing uniformity and a degree of objectivity to specific outcomes relies on their uptake by practitioners and researchers.

The accessibility and apparent simplicity of many outcome measures may also belie the true complexity of health care performance. As described in chapter 1, most definitions of quality of care are broad. For example, the WHO definition covers 6 domains, including effectiveness, efficiency and safety. Evaluating care using a single outcome measure will not provide a comprehensive appraisal of overall performance.

As health care makes advances and outcomes improve, further improvements may become increasingly difficult to achieve. If new treatments are to be evaluated against existing therapies, it may be very difficult to demonstrate any further improvement to justify their adoption. Less frequently studied outcomes, or alternative measures, may be required to evaluate the benefits of new approaches. For example, key-hole surgery has been shown to improve short-term outcomes like pain and length of hospital stay after bowel surgery\textsuperscript{48,49}. However, there has been no high-level evidence of mortality benefit. Despite this, most surgeons would consider such an operative approach to deliver improved care. In comparison, robotic techniques may make certain operations easier to perform, and this too could represent an important improvement. However, there is currently no high quality randomised evidence of improved outcomes after robotic surgery\textsuperscript{50}. The wide-spread adoption of robotic techniques cannot be justified if mortality or on-
cological outcomes are the key measures of performance upon which uptake of this technology is based.

There can also be statistical challenges with using certain outcomes to assess performance in health care. If an event is relatively uncommon in a population, such as postoperative death or a particular complication, measuring a difference in event rate within a study population can require large sample sizes. For example, if the population average mortality rate after removal of part of the gullet or oesophagus is 6.7%, a sample size of 148 procedures would be required to have 95% confidence ($\alpha = 0.05$) and 80% power ($\beta = 0.2$) to detect a surgeon with a mortality rate that was twice the average. With a median of 11 operations per year for individual surgeons, a new consultant would take over 13 years to accrue the necessary number of cases. The gold standard assessment for new treatments is a randomised controlled trial, where two groups are compared, such as the existing standard treatment for a condition and a new technique or therapy. If accepted standard technique A has a postoperative mortality of 5%, and technique B is expected to have a mortality of 2.5%, 1810 patients would need to be included in the trial to have the same confidence and power to detect this large and important difference in treatment outcomes. Few new techniques or treatments could be expected to result in a 50% relative risk reduction for a specific outcome, particularly when the existing treatment has been developed and improved over many years. Therefore, other measures may be required to establish the benefits of new approaches when good results are already being achieved using established techniques.

It can therefore be seen that while outcomes have been, and will continue to be, a central aspect of the assessment of treatments and performance in health care,
their use is not without challenges and potential pitfalls. A broad consideration of all aspects of performance may provide a richer understanding of how further improvements may be achieved in the future.

2.3 Types and examples of surgical outcomes

There is a huge range of outcomes that may be relevant within different areas of health care, and different outcomes may be important to the different stakeholders involved. In this section, the different types of outcome relevant to colorectal surgical practice are discussed. Outcomes may be classified as general, specific to a particular condition or disease, or specific to a given treatment. Examples for a patient undergoing surgery for colonic cancer are provided in table 2.3.

<table>
<thead>
<tr>
<th>General</th>
<th>Disease-specific</th>
<th>Treatment-specific</th>
</tr>
</thead>
<tbody>
<tr>
<td>• One-year survival</td>
<td>• QLQ-CR38 questionnaire</td>
<td>• Complications, such as anastomotic leak</td>
</tr>
<tr>
<td>• SF36 questionnaire</td>
<td>• Bowel motions passed per day</td>
<td>• Blood transfusion</td>
</tr>
<tr>
<td>• EQ-5D questionnaire</td>
<td>• Weight lost</td>
<td>• One-year disease-free survival</td>
</tr>
<tr>
<td>• Length of hospital stay</td>
<td>• Appetite</td>
<td>• Pathological resection margin status</td>
</tr>
<tr>
<td>• Readmission to hospital</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.3: Outcomes relevant to an operation for colonic cancer.
SF36 - Short Form 36; EQ-5D - EuroQuol 5 Dimension; QLQ-CR38 - Quality of Life Questionnaire for Colorectal Cancer 38.

General outcomes may be relevant not just to colorectal surgery, but to almost any aspect of health care. The most universal outcome is whether or not a patient is alive at a given point in time. Other general outcomes include quality of life measures, such as the Short Form 36 (SF36) questionnaire or EuroQuol five dimensions (EQ-5D) instrument. Some general outcomes may not be universally
relevant, but still apply to a significant proportion of health care activity. For any care requiring admission to hospital, broadly relevant outcomes include length of stay and readmission.

Within colorectal surgery, specific outcomes may relate directly to the underlying condition or cause of ill health, or to the treatment provided. A variety of condition-specific quality of life measures exist, such as the QLQ-CR38 colorectal cancer-specific quality of life questionnaire.

Specific outcomes in the surgical treatment of cancer include disease-free survival (survival with no evidence of cancer recurrence), as well as local, regional and distant recurrence, according to the site of tumour regrowth if this occurs. Other outcomes may be directly related to the treatment provided, such as anastomotic stricture or leak. An anastomosis is a join between two hollow organs, such as two pieces of bowel or blood vessels. If such a join narrows, it is called a stricture or stenosis, and this can cause problems with flow through the organ; alternatively, if the join is not effective, the contents of the organ can leak and cause problems with infection or bleeding.

A large and diverse range of outcomes may therefore be relevant to a patient undergoing colorectal surgery, or when appraising the performance of a particular unit.

### 2.4 Unit-level surgical outcomes

In this section, two of the key areas explored by existing surgical outcomes research that is focused on unit-level performance are discussed: unexplained variation in outcomes; and the relationship between different outcome measures. Variation in
Outcomes in surgery

outcomes has rightly received a great deal of attention, as this variation is commonly interpreted to show that the quality of care varies significantly between units. Despite much work, this observed variation has yet to be adequately explained. Researchers have also undertaken studies exploring the relationship between different outcomes, varied either by definition or area of practice. These avenues of research are important in contextualising the studies presented in this thesis.

2.4.1 Unexplained variation

Much existing research examining outcomes at the departmental or organisational level has described greater variation in outcomes than can be explained easily. Measurement of any variable using a sample from a population represents an estimate of the true population value. The accuracy of that estimate is influenced by the sample size. Therefore, some degree of variation in outcomes may be explained by strictly statistical considerations related to sampling. It is also known that certain groups within a population may be more or less likely to experience certain outcomes. For example, an elderly patient, or someone who has previously had multiple heart attacks, will have a greater chance of dying after major abdominal surgery than a young and fit patient. Patient factors such as age and comorbidity can be measured, and statistical techniques can be used to adjust outcomes to allow for differences in patient characteristics at different hospitals. This process is called risk- or casemix-adjustment, and it has been shown to influence relative performance between organisations\textsuperscript{56}. However, even after risk-adjustment and the role of chance are incorporated into the assessment process, many studies have
shown more variation in organisational outcomes than expected. For example, Morris et al showed significant variation in 30-day risk-adjusted mortality after colorectal surgery in England\textsuperscript{24}. Other studies of colorectal surgery have shown wide variation in reoperation rates\textsuperscript{22,28}, length of stay\textsuperscript{25} and failure-to-rescue\textsuperscript{57}, as well as other measures such as readmission and abdominoperineal excision rates\textsuperscript{20}. Some research has been conducted to try and better understand this unexplained variation, typically by examining the association between structural factors such as operative volume, and outcomes. Studies examining the relationship between organisational structures, processes and outcomes are discussed in chapter 3.

### 2.4.2 Relationship between outcome measures

Some researchers have demonstrated that varying the definition of a particular metric may result in a different assessment of performance at the institutional level. Using observed-to-expected mortality (a risk-adjusted outcome measure), two studies reported that lengthening the period of follow-up can have a notable impact on the identification of units with outlying performance, although statistical correlation between mortality rates for different periods was high\textsuperscript{58,59}.

Others have examined the relationship between different metrics within a speciality. Particular attention has been paid to the outcomes of patients who have experienced a complication. The term ‘failure-to-rescue’ was first coined by Silber in 1992 to describe a fatal outcome after a complication of health care\textsuperscript{60}. In a later paper, Silber and colleagues examined outcomes after coronary artery bypass grafting\textsuperscript{61}. There was no association between postoperative mortality and complication rates, or between failure-to-rescue and complication rates. This means
that units with the highest complication rates did not have the highest overall postoperative mortality rates, nor did they have the highest death rates following detection of a postoperative complication. Using data from the American College of Surgeons (ACS) National Surgical Quality Improvement Program (NSQIP), Ghaferi and coworkers studied units stratified by overall risk-adjusted mortality rates across different procedures, and within specialties. Across mortality quintiles, there was no difference in overall complication rates or the occurrence of major complications. However, there was an association between overall mortality quintile, and mortality after major complications. Almoudaris et al examined the relationship between a broad range of outcomes after colorectal surgery: 30-day postoperative mortality, return to theatre rate, mortality rate after a return to theatre (termed surgical failure-to-rescue), 28-day readmission rate, length of stay, and abdominoperineal excision rate. The cited study found a significant weak association between 30-day mortality and the rate of return to theatre (Pearson’s r = 0.191, p = 0.020), with a stronger, albeit still only moderate, association between 30-day mortality and the surgical failure-to-rescue rate (Pearson’s r = 0.445, p <0.001). However, there was no association between 30-day mortality rates and other measures, such as length of stay or readmissions.

Other studies have examined the relationship between measures of performance in different areas of practice within the same speciality. In two separate studies, Ingraham et al reported that, within the specialties of colorectal and general surgery, elective and emergency outcomes were not strongly correlated. While the same surgeons may be providing these services in the same hospital, relative performance measured by aggregated morbidity and mortality rates was not comparable.

More broadly, others have examined the relationship between outcomes across
different specialties within a hospital. Two studies, one comparing trauma surgery with elective and emergency general surgery, and another comparing in-patient and out-patient surgical procedures, did not show any significant relationships between these different fields of surgical practice within individual institutions\textsuperscript{64,65}. However, an earlier study by Dimick et al found a tendency for mortality rates after high-risk surgical procedures to correlate at the hospital level\textsuperscript{66}. Outside the surgical field, DesHarnais and coworkers found no association between risk-adjusted hospital-wide mortality, complication or readmission rates\textsuperscript{8}.

2.5 Discussion

Existing outcomes research has made important contributions to the present understanding of unit-level performance in surgery. For example, we now have a much better grasp of the significance of complications during the postoperative period. Existing outcomes research has shown that the units with the lowest postoperative mortality rates do not have the lowest complication rates. It may previously have been considered by many surgeons that the way to improve outcomes was to reduce the incidence of complications. Anecdotally, many clinicians will have worked in units where an individual surgeon proudly states that, ‘I don’t have complications.’ While the pursuit of ‘zero complications’ may be laudable, up to a point, the evidence suggests that a low rate of complications is not associated with the lowest postoperative mortality rates. Therefore, the highest performing units do not achieve superior results by lowering their rate of complications. The study by Ghaferi et al showed that units with lower overall mortality had average complication rates, but managed to achieve lower mortality after a complication
occurred. Therefore, high performing units seem to achieve demonstrably better results after a complication has occurred. This suggests that the management of complications may be as, or more, important than their prevention. This represents a significant shift, and requires a change in attitude among practitioners. Having complications is not a sign of failure or poor quality surgery. All units have complications. Appropriate and expert management of complications may be one of the distinguishing characteristics of the best surgical teams.

In addition, the studies discussed in this chapter go some way to deconstructing any notion of an uniformly high performing unit or hospital. Speciality teams may perform well on one metric, but revision of the definition of the metric, or examination of other areas of practice within that team’s activity, may reveal average or below-average results. Individual units are likely to have strengths and weaknesses, and it appears that there are no (or perhaps very few) globally high performing surgical institutions or teams across metrics. Within a hospital, different specialties may perform at significantly different levels, despite sharing common facilities and supporting resources. Outside the peer-reviewed literature, examinations of failing organisations support this picture of performance. Reports from the Mid Staffordshire NHS Foundation Trust Public Inquiry and the recent ‘Review into the quality of care and treatment provided by 14 hospital trusts in England’ describe pockets of good quality care even when the culture and systems within these organisations were clearly problematic. Overall, these findings perhaps reflect the inherent complexity of modern health care.

While it may seem disappointing to find that an idealised, truly high performing institution, with globally high standards of care and good outcomes, does not seem to exist, this finding may be considered in a positive light. Quality of care
may be considered as either a reflective or formative construct\textsuperscript{40}. If considered to be a reflective construct, performance measures are examined for correlation, in the quest of indirectly assessing an unmeasurable latent construct that represents overall quality of care. However, if quality of care is conceptualised as a formative construct, derived from the combination of other measures, one would not necessarily expect the components of an overall measure of quality to correlate. It may actually be preferable to use uncorrelated measures to avoid redundancy. While this may result in a more complicated picture, it may also yield a more nuanced, and importantly a more accurate, appreciation of department-level performance in health care. Individual surgical units may be good at some things and on some measures, and they may be average or poor in other areas. An overall assessment of the quality of a unit may incorporate information on a number of areas of performance to provide a high-level focused summary, which, combined with the component measures, may result in an accurate picture of how a particular team is functioning.

2.6 Chapter summary

This chapter has discussed some of the advantages and disadvantages of outcomes as measures of performance, and reviewed some important aspects of unit-level performance covered by existing surgical outcomes research. While the focus of the chapter has been broad, it is of direct relevance to the work presented in this thesis, which builds upon these areas.
Chapter 3

Understanding organisational outcomes in surgery: a literature review

3.1 Chapter overview

A key aim of this thesis is to develop a better understanding of how the best colorectal units achieve their results, so that recommendations may be made for future quality improvement. It is therefore important to examine the existing evidence in this area. This chapter presents the methods and results of a review of the literature, covering research that links specific structures or processes with outcomes. The review covers a broad range of studies, relevant to surgical performance, and completes the background picture, against which the thesis studies are presented. This review was also of particular importance in guiding the development and refinement of the interview protocol for study 6.
3.2 Structure and process measures of performance

As described in chapter 1, structures and processes represent the ingredients that determine health care outcomes. This section will discuss examples of structures and processes, and their role in understanding health care performance.

Structures may be considered as the ‘pre-conditions’ of work, and they include physical and other factors that create the setting in which specific processes are delivered. The volume of procedures performed by a team or individual clinician has probably been the single most studied structural factor within the surgical literature. Other examples, such as the number and qualifications of staff, have also been examined. Some structures may be considered to have intrinsic value, such as good quality and supply of technical equipment. Others may require assessment by other criteria, such as outcomes, to determine their role in determining the level of performance achieved. For example, operative case volume has typically been examined against postoperative mortality rates. While structural factors are undoubtedly important in framing the processes of health care, it can be difficult to act upon the knowledge that a specific structural factor is associated with improved outcomes. Increasing the operative volume in surgical units may require low-volume units be closed, and greater support for high-volume units that may consequently receive increasing numbers of referrals for surgery. This may be difficult, expensive and time-consuming to achieve.

Processes are the activities performed in the delivery of care. The surgical approach to removing a section of the large bowel, be it laparoscopic or open, is a key clinical process in the care of a patient undergoing surgery to remove a bowel
cancer. As with structural factors, certain processes may have intrinsic value, such as involvement of patients in decisions which affect their care. However, many processes require evaluation by assessing for association with outcomes. For example, the administration of antibiotics when patients are having a bowel operation may be evaluated by assessing the rate of wound infections within the first 2 weeks after surgery. In general, implementing knowledge that a specific process is associated with better performance is easier than changing structural factors. However, data on process factors can be difficult and expensive to collect.

The examination of structures and processes is therefore key to quality improvement research. Understanding what health care workers actually do to achieve specific outcomes for patients will inform recommendations for improvement, and such knowledge may guide the development of specific interventions to improve care.

3.3 Aims

The narrative literature review described in this chapter was performed for two reasons. Firstly, it was intended to develop a broad understanding of the existing literature to guide the studies presented in this thesis. Secondly, it was conducted to inform the development of the interview protocol for study 6.

The focus of the review was to identify evidence examining for a relationship between specific structures or processes, and outcomes, relevant to department-level performance in surgery. While some structure and process factors may be considered important contributors to performance on face value, through expert opinion or survey data, only studies exploring for an association with outcomes
were included in the review.

It was aimed to obtain an overview of the literature. A systematic review and meta-analysis was not appropriate or feasible due to the breadth of understanding sought, and the disparate nature of the evidence collected. Studies were reviewed without intention to reach definitive conclusions about the relationship between all factors examined and any associated outcomes. Where evidence of a relationship was mixed, this is reported to accurately reflect the state of the existing research.

3.4 Methods

This review examined a range of structures and processes that were considered relevant to the outcomes of a surgical department. A pragmatic, narrative approach was adopted for two key reasons. Firstly, the objective was to obtain a broad overview of the relevant literature, rather than to exhaustively examine all the literature within a narrow, focused area of surgical care. Secondly, a flexible approach was required, as the quantity and quality of evidence in each area examined was highly variable. Only one similar review, examining organisational factors in the intensive care unit, was identified\textsuperscript{68}.

The following sections describe the process followed to identify important structure and process factors, each of which was then the focus of a separate literature search. The strategy for identifying relevant literature for individual searches is then described.
3.4.1 Identification of key factors

Important structures and processes were identified in the context of various frameworks that were used to provide a broad overview of surgical performance. This thesis is founded upon Donabedian’s structures, processes and outcomes model for understanding health care performance\(^{43}\). Surgical care is a largely linear process consisting of a series of discrete phases. Factors influencing performance can therefore be classified as pre-operative, operative, and post-operative. There may also be over-arching themes, relevant to all stages of the patient journey. In addition, a description of the domains of health care delivery was sought. While Vincent’s ‘Factors that influence clinical practice’ was primarily intended for the analysis of risk and safety in medicine, it also has immediate relevance to performance in health care more broadly\(^{69}\). These frameworks, combined with the author’s clinical experience, yielded a list of factors that were considered relevant to unit-level surgical performance. These factors formed the basis of the series of literature searches performed in the present review, and are listed in table 3.1.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Process</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Operative volume</td>
<td>• Preoperative assessment and optimisation</td>
</tr>
<tr>
<td>• Technology</td>
<td>• Operative management</td>
</tr>
<tr>
<td>• Staffing level</td>
<td>• Postoperative management</td>
</tr>
<tr>
<td>• Building layout and design</td>
<td>• Leadership</td>
</tr>
<tr>
<td>• Out-of-hours care</td>
<td>• Quality assessment and improvement</td>
</tr>
<tr>
<td>• Total work hours</td>
<td>• Standardisation and protocols</td>
</tr>
<tr>
<td></td>
<td>• Communication and collaboration</td>
</tr>
<tr>
<td></td>
<td>• Management of adverse events and error</td>
</tr>
</tbody>
</table>

Table 3.1: Structures and processes examined in literature review.
3.4.2 Search strategy

A specific strategy for identifying relevant research for retrieval and review was developed to meet the aims of this review. The two key components of this strategy were a modular approach and an iterative process of refinement. Separate series of searches were conducted to identify evidence relevant to each of the factors listed in table 3.1. An example is presented in appendix A.

Individual searches adopted a modular, concept-based approach. Concepts were built using MeSH and free-text search terms applied to title and/or abstract fields, combined using the ‘OR’ Boolean operator. For example, the concept of outcomes was constructed using the following search terms: mortality, death, morbidity, complication, length of stay, patient satisfaction, or the term ‘outcomes’. Concepts were then grouped using the ‘AND’ operator to identify research articles relevant to the aims of the review.

After initial completion of a concept-based search, the number of hits and top listed article titles and abstracts were screened to determine if further refinement was required. If the literature returned was not relevant, or an unfeasibly large number of articles was returned (>500 as a rule-of-thumb), additional concepts were added to narrow the search and/or reduce the number of articles for screening. For example, the additional concept of surgery could be added, with further refinement to colorectal surgery if appropriate. In this way, searches were performed iteratively, with initial broad searches being refined and made increasingly specific depending upon the relevance, quality and quantity of literature identified.

The literature was searched using the OvidSP search engine to access the MEDLINE database, including in-process and non-indexed articles. Searches were per-
formed during April and May 2013, and selected updates were made in early 2016. Wildcard operators were used to allow for alternative spellings or word endings. Articles were only eligible for inclusion if they related to studies in humans and were published in English. All article types, including reviews, were considered. Reference-checking and citation-tracking was performed to identify further relevant research.

3.5 Selected results

Fourteen individual searches were performed across all the structure and process factors listed in table 3.1. The level, volume and relevance of the existing evidence in each area was highly variable. For example, a number of systematic reviews and meta-analyses were identified examining the relationship between operative volume and surgical outcomes in a range of different surgical specialties. In other areas, there was very limited existing research. No studies were identified assessing the impact of a general pre-operative assessment clinic. Therefore, the separate functions of such a clinic were considered, and relevant aspects were examined individually, such as pre-operative risk stratification. Due to the large number of searches performed, a selection of the literature identified and reviewed is presented in this chapter. A summary of the literature in the other areas reviewed is included in appendix B.

3.5.1 Structural factors

In this section, the evidence examining the relationship between operative volume and staffing levels, and outcomes will be reviewed.
Operative volume

There is a significant amount of evidence supporting a link between operative volume and outcomes. A number of meta-analyses have shown an association between increased surgeon or hospital caseload and reduced mortality for non-colorectal surgical specialties\textsuperscript{70–75}. Specific to colorectal surgery, a Cochrane Systematic Review reported improved five year survival after surgery at high-volume hospitals, by high volume surgeons, and by colorectal specialists\textsuperscript{76}. There was a significant reduction in operative mortality among high-volume surgeons and colorectal specialists, though there was no relationship with hospital caseload. High-volume surgeons also achieved lower rates of permanent stoma formation.

However, this evidence belies the likely complexity of this statistical association. Lower mortality among units with high volume may be related to both surgeon and hospital factors\textsuperscript{74}. If individual surgeons perform greater numbers of operations per year, this may help develop their expertise and skill. However, hospital resources such as nursing care, intensive care and radiology support, may perhaps play an even more important role in improving outcomes\textsuperscript{74}. Longitudinal evidence that raising volume within a centre improves results is lacking. One Canadian study reported that increasing annual caseload within a hospital was not associated with significant changes in outcomes, although there may have been a trend towards a reduction in mortality and a rise in length of stay\textsuperscript{77}. There is also no universal definition of ‘high volume’ upon which to base recommendations for practice.

Considered as a whole, there is strong observational evidence of an association between operative volume and outcomes for many surgical procedures, but
understanding of this association and demonstration of causation are lacking.

**Staffing level**

**Nursing staff**  A 2007 systematic review and meta-analysis examined 28 studies meeting inclusion criteria, and reported a statistically and clinically significant association between increased registered nurse staffing levels and a number of improved patient outcomes\(^{78}\). These included reduced mortality for surgical, medical and intensive care patients, as well as reduced failure-to-rescue in surgical patients and reduced length of stay for patients on surgical and intensive care units. The authors did note that other hospital characteristics, such as a particular institution’s commitment to quality of care, were likely to play a causal role in the association.

A number of other studies have examined nurse staffing levels within speciality-specific contexts. West et al undertook a systematic review of nursing levels within intensive care which included 15 studies\(^ {79}\). They found a stronger link between staffing and adverse events than between staffing and mortality rates, but cautioned that the evidence was not yet convincing. A Belgian study of postoperative nursing care after cardiac surgery used sophisticated hierarchical modelling techniques to investigate the relationship between nursing hours and qualifications, and patient outcomes\(^ {80}\). The study reported a significant association between greater nursing hours per patient day, and a higher proportion of registered nurses with a Bachelor’s degree, and lower in-hospital mortality.

While further study may be required to characterise the relationship between nursing levels, their qualifications, and patient outcomes within particular surgical and medical settings, current evidence is strongly suggestive that greater levels of
nursing staff are associated with improved results for patients.

**Medical staff** There is some evidence supporting a relationship between increased numbers of doctors and improved outcomes. Directly applicable to colorectal surgery, the evidence is sparse. Yasunaga et al examined six surgical cancer procedures (including colorectal surgery) in Japan and found that relative to the reference group (low physician-to-bed and nurse-to-bed ratios), hospitals with high levels of both doctors and nurses had lower mortality and lower failure-to-rescue rates. However, isolated high levels of either doctors or nurses was not associated with improved outcomes. In a study of NHS hospitals, Jarman and colleagues found a significant association between adjusted hospital mortality rate and both the hospital doctor-to-bed and general practitioner-to-population ratios. Pronovost and colleagues conducted a systematic review examining the impact of staffing patterns for intensive care on patient outcomes. Their review found that high-intensity staffing (defined as mandatory consultation with an intensivist or a fully closed intensive care unit (ICU) model) was associated with lower mortality.

In other areas, such as obstetrics, there is evidence suggesting a greater number of doctors per capita in the general population may improve outcomes. More broadly, population- or primary care-based studies have shown a link between higher doctor-to-patient ratios and improved outcomes, though higher levels of specialist doctors may be associated with reduced quality measures in some cases.

The cited studies were clearly rather heterogeneous, and examined the level of medical staffing in a number of different ways. For example, Yasunaga et al inves-
tigated physician staffing level, defined as the number of attending and resident physicians per 100 beds\textsuperscript{81}, and Jarman and colleagues examined the ratio of all hospital doctors to beds\textsuperscript{82}. Pronovost and coworkers examined the intensity of input from the critical care team, with various definitions across different studies\textsuperscript{83}. It would be difficult to define a specific staffing strategy to implement on the basis of these studies. Perhaps more importantly, it is currently not clear whether the broad findings reported above are immediately relevant to colorectal surgery.

### 3.5.2 Process factors

In this section, research focused on the relationship between patient management during the three phases of the patient journey (preoperative, operative, and postoperative) and the consequent outcomes will be examined. In addition, the role of standardisation of care and inter-professional communication will be reviewed. Other process factors examined are discussed in appendix B.

**Preoperative assessment and optimisation**

Recent years have seen the gradual introduction of some form of preoperative assessment process across all surgical specialties. In many cases, this is provided in a preoperative assessment clinic. Such a clinic may serve a variety of functions, including: completing administrative tasks, such as checking test results; providing basic medical advice, such as information on smoking cessation and perioperative medication management; individualised perioperative risk stratification; and modification of perioperative risk through specific interventions. No evidence was identified comparing care with and without preoperative assessment, in the broad
sense outlined above. Therefore, a more focused approach was required. The literature was searched to identify evidence relevant to the specific functions of perioperative risk assessment and modification, and a summary of the research is presented below.

Perioperative risk stratification may be performed using more or less sophisticated methods. Accurate risk stratification for morbidity and mortality has proven challenging, and there is a need for further research in this area\textsuperscript{91}. Readily available patient-level information such as age and medical history may allow simple predictions of risk. Van Klei et al showed that model predictions of postoperative myocardial infarction or death based on risk factors garnered from the history alone were not improved by routine use of electrocardiographs (ECGs), a simple bedside investigation\textsuperscript{92}. More complex, but also more expensive, techniques exist, such as cardiopulmonary exercise testing (CPET). Physiological variables determined using CPET may be associated with operative risk of morbidity and mortality\textsuperscript{93}, but it is not widely available within the NHS. It is likely that such techniques may be more broadly adopted if they can be shown to improve outcomes, and reduce costs.

There is evidence that operative risk may be modified in certain settings. A meta-analysis by Chopra et al demonstrated a reduction in cardiac morbidity and length of stay after surgery associated with perioperative statin therapy for patients who are at risk of cardiac events or undergoing high-risk surgery\textsuperscript{94}. However, the evidence for perioperative beta blockade is mixed, and variation in the drug regimens used in trials has made meta-analysis problematic\textsuperscript{95}. Two reviews of preoperative physical therapy, or ‘prehabilitation’, have suggested that this may reduce length of stay and reduce pulmonary complications, but the overall qual-
ity of data is modest\textsuperscript{96,97}. Davies and Wilson reviewed the evidence regarding specific preoperative interventions to raise delivery of oxygen to tissues\textsuperscript{98}. Three randomised controlled trials in which delivery of oxygen was raised to an artificially high level showed significant reductions in mortality in the intervention groups, with relative risk reductions exceeding 66%. A larger, multi-centre trial showed no such benefit, although methodological and patient details, as well as the long 9-year period it took to complete the study, may have reduced its ability to measure differences between study groups. In addition, this intervention is highly expensive and invasive, requiring preoperative admission to an Intensive Care Unit for monitoring and therapy. Its potential benefits seem dramatic, but its financial and resource cost is perhaps prohibitively high in the face of concerns about its true efficacy.

There is some evidence suggesting perioperative risk may be predicted and modified. Preoperative assessment may also serve logistical purposes, including the coordination of care and provision of basic medical advice. High quality evidence of benefit may be lacking, but it may still be considered to represent good practice given the range of functions it serves.

**Operative management**

There is good evidence that the specific surgical approach, such as key-hole or laparoscopic surgery compared with conventional open surgery, has a significant impact on outcomes after colorectal surgery. A number of systematic reviews of trials have shown that the laparoscopic technique is associated with shorter length of stay, possibly reduced complication rates and equivalent mortality compared with the open approach\textsuperscript{99–102}. 
The impact of specific anaesthetic and analgesic approaches on current practice in colorectal surgery is still being defined. This may be due to the significant changes in practice that have occurred over the last 10 years, with increasing adoption of laparoscopic surgery and multimodal enhanced recovery programs. It has been shown that in major open abdominal surgery, epidural analgesia is associated with better pain control and earlier return of gastrointestinal function. Indeed, epidural analgesia has been viewed as an integral component of enhanced recovery programs. However, with increased adoption of minimal access surgery, the role of epidural analgesia is being reassessed. In colorectal surgery, a randomised trial reported longer stay and slower return of gut function in patients receiving an epidural, when compared with a spinal block or systemic opioids via a patient-controlled device. Joshi and colleagues reviewed the evidence covering all analgesic strategies for laparoscopic colorectal surgery, concluding that the risk-benefit profile of epidural analgesia suggested that this should not be routine practice. Instead, the authors recommended a multimodal approach, utilising local anaesthetic infiltration and combination therapy of paracetamol with anti-inflammatory drugs, reserving opioids for breakthrough pain. The only strategy that has not changed is the minimisation of opioid use, which has been shown to delay the return of normal bowel function.

There is a lack of evidence supporting goal-directed fluid therapy during anaesthesia, a specific technique for the optimisation of an individual patient’s circulation. Rocca and Pompei described the important physiological parameters that may be measured with common monitoring systems and evidence associated with each, concluding that the area remains highly controversial.

Therefore the available evidence suggests that a laparoscopic approach, com-
combined with an analgesic strategy to minimise opiate use and optimise pain control, may represent optimal operative practice to reduce length of stay, but there is no evidence that such an approach will have a significant impact on other outcomes such as postoperative mortality.

**Postoperative management**

While the ERP incorporates elements across the entire perioperative pathway, some of the most salient changes to traditional practice included in this approach apply to the postoperative management of patients. These include early oral nutrition and mobilisation, as well as avoidance of routine postoperative nasogastric intubation, or abdominal drainage for surgery above the peritoneal reflection (i.e. non-rectal abdominal surgery)\(^\text{110}\). A number of systematic reviews and meta-analyses have demonstrated shorter hospital stay and reduced complications associated with the ERP, although the quality of evidence upon which these findings are based is typically low or moderate\(^\text{111–113}\).

Outcomes research has suggested that the management of postoperative complications, a specific set of patient care processes, may be an important determinant of surgical performance. It is known that patients experiencing complications are more likely to die after surgery\(^\text{114–116}\). It has also been shown that if a patient develops a postoperative complication, length of stay is increased\(^\text{117}\). Although a higher rate of complications may be associated with a greater number of deaths at the level of individual patients, a number of studies have shown a lack of association between the unit-level rate of complications and overall mortality or failure to rescue rates\(^\text{27,57,61}\). These findings, also discussed in section 2.4 of chapter 2, suggest that the relationship between a complication occurring and a fatal out-
come may not be fixed. It may be inferred that some units are better at managing complications when they occur, minimising the adverse outcomes for their patients, delivering a higher quality of care. One challenge to better understanding the role of complications and their management is accurate measurement. As discussed earlier, complications may be defined according to pre-specified criteria, or according to the treatment required (such as the Clavien-Dindo system\textsuperscript{46,47}). The strength of the Clavien-Dindo approach is that it reflects the severity of the complication, and obviates the need to have an exhaustive list of definitions for all eventualities. Its increasing use may help further knowledge in this area, but at present there is a lack of evidence pertaining to the role of specific practices in complication detection and management.

Overall, the published evidence suggests that postoperative care based upon the ERP, and the effective management of complications, may be important determinants of surgical outcomes, although the specific care processes involved have not been comprehensively defined.

**Standardisation and protocols**

It is challenging to design a study to effectively evaluate whether the standardisation of care and use of protocols has a positive impact on patient outcomes. Typically, care pathways are developed by appraising the existing evidence applicable to a specific condition or procedure, combining recommendations for best practice into a broad package of care. It is arguably unethical to attempt to evaluate the impact of such a pathway in a randomised controlled trial, as those not cared for within the pathway may not be receiving care that is known to be superior. It may be impossible to disentangle the effects of specific evidence based clinical processes.
from changes to other aspects of the organisation of care, such as communication and collaboration. For example, if the sequence of care is pre-planned for the majority of routine cases, protocols may be considered as communication ‘ahead of time’ between the members of the health care team, as team members know what to expect and what should be done at each point in the patient journey. Also, it can be difficult to avoid contamination between intervention and control groups for complex, multi-disciplinary interventions. Thus, evaluating and understanding the impact of standardisation of care is fraught with methodological difficulties.

Notwithstanding such challenges, there is some evidence suggesting that standardising care and utilisation of protocols may improve outcomes. A Cochrane Systematic Review examined the impact of clinical pathways on practice and outcomes\textsuperscript{118}. The definition of clinical pathways was carefully considered, covering specific and locally adapted guidance on patient management including an element of time-specificity for individual steps. The studies included covered a variety of conditions and countries. The key findings were an association between pathway use and reduced complication rates, improved documentation, and likely reduced hospital length of stay (a finding limited by heterogeneity of the studies included). Pathway adoption was associated with reduced hospital resource consumption. Evaluation of the impact on mortality was limited to just 3 studies, with a pooled odds ratio of 0.85 (95% confidence interval of 0.61 - 1.11).

The SURPASS Collaborative Group developed and implemented a comprehensive checklist covering the entire surgical patient journey in 6 hospitals, and compared outcomes with 6 control hospitals\textsuperscript{119,120}. They found an association between implementation of the checklist and a reduction in mortality and complication rate, with a non-significant reduction in length of stay, compared with no
change in control hospitals.

Other studies of pathways and protocols have pointed out that having a protocol does not necessarily translate into implementation and delivery\textsuperscript{121}, and that in certain settings some patient subgroups may stand to benefit more from standardisation than others\textsuperscript{118}.

Considered as a whole, the available evidence suggests that standardisation of care using clinical pathways is likely to improve outcomes.

**Communication and collaboration**

There is some evidence supporting an association between better communication and collaboration, and improved patient outcomes. Davenport et al used a survey to examine the relationship between various organisational factors and outcomes\textsuperscript{122}. They found a significant association between increased risk-adjusted surgical morbidity and lower respondent ratings of communication with attendings (p<0.01) and residents (p=0.08), with a weaker, non-significant association for communication with nurses. Young et al undertook a qualitative study of communication and coordination practices in high and low morbidity and mortality hospitals, describing more regular and highly developed systems for communication in hospitals with better results\textsuperscript{123}. In a follow-on study, a survey based on earlier qualitative work was used to examine if improved collaboration was associated with better outcomes\textsuperscript{124}. A high score for coordination of care was non-significantly associated with lower morbidity, with a weaker association with mortality.

Indirect evidence also supports an association between improved outcomes and greater communication and collaboration. Shortell et al assessed the association
between intensive care unit management and outcomes\textsuperscript{125}. A composite score labelled ‘caregiver interaction’, reflecting leadership, culture, communication, coordination and conflict management, was significantly associated with reduced length of stay, reduced nursing staff turnover, improved staff-rated quality of care and ability to meet the needs of patients’ relatives. The improvement in outcomes associated with adoption of the SURPASS checklist are likely, to some extent, to have involved improved coordination of care\textsuperscript{120}. In an observational study, communications failures were shown to contribute to a significant proportion of process failures in postoperative care\textsuperscript{126}.

Overall, the evidence suggests that good communication and collaboration is an important determinant of certain clinical outcomes, though evidence of an association with mortality is lacking.

3.6 Discussion

This review has shown that a number of structure and process factors may have a significant impact on unit-level outcomes. Meta-analytic evidence suggests that increased operative volume and higher levels of nursing staff are associated with better results. There is also high-level evidence of improvements in performance associated with laparoscopic surgery and standardisation of care, particularly the ERP. Research suggests a link between better quality communication and collaboration and improved outcomes, and the way different units manage postoperative complications may also be important. Changes in surgical approach have rendered earlier data regarding the optimal analgesic strategy less applicable to current practice, with routine use of epidural pain relief increasingly questioned in laparoscopic
surgery. Evidence supporting the benefits of preoperative assessment and optimisation is relatively limited, though this element seems to have become an accepted standard of care, perhaps due to logistic and administrative benefits.

It is likely that some of the factors that have been associated with improved outcomes may be correlates of better performance, rather than causative contributors to improvements in care. In a paper examining the impact of teaching intensity on outcome, Silber et al reported a notable association between teaching intensity and hospital surgical volume, as well as number of beds, technology index, nurse-to-bed ratio and nurse mix\textsuperscript{127}. Unpicking this complicated picture, where there is collinearity between common variables of interest, is one of the key challenges in quality improvement research.

In addition, despite the research that has been summarised in this chapter, we still do not have a detailed understanding of how certain structures and processes affect outcomes. What do high-volume surgeons or units do differently to achieve their results? What are the differences in the ways units manage complications, that mean some units have much lower failure to rescue rates than others? Are the elements of the ERP critical in reducing length of stay, or is the main benefit of this program through standardisation and reduction in variation? Detailed clinical research specifically answering these questions is needed to direct practical efforts to improve the quality of care as it is delivered within the NHS, and in other health care systems.
3.7 Chapter summary

This chapter began by reviewing some of the salient features of structures and processes as measures of performance, before presenting the methods and selected results of a literature review of the evidence linking specific structure and process factors with outcomes, relevant to unit-level performance in colorectal surgery. This review provides essential background for the work presented in this thesis. It has shown that there are still important gaps in the present understanding of surgical performance. For example, we do not understand how a higher volume of surgical cases within a unit affects clinical care, in practical terms, to improve outcomes. A better understanding of the factors underlying such relationships will be key to guiding future improvements in the quality of care provided to patients undergoing colorectal surgery.
Introduction summary and thesis questions

Introduction summary

The preceding three chapters have set out the historical context and research background for the work presented in this thesis. Repeated high-profile failures of professional self-regulation, coupled with changes in expectations of health care within society and rapid growth in the availability of information, have driven a shift towards open scrutiny of health care performance. While undertaking this thesis, the government began publishing consultant- and hospital-level outcome data across a number of specialties, including colorectal surgery, during 2013. This has arguably sharpened the focus on unit-level outcomes, and the need to develop a better understanding of surgical performance more broadly.

The existing research examining unit-level outcomes, and the key structure and process factors shown to have a relationship with outcomes, has been reviewed. This research has shown that unit-level performance is not necessarily consistent across different outcome measures. The importance of the management of complications has been highlighted by studies examining the outcomes of patients who experience a complication after surgery, known as failure to rescue. Within the surgical literature, there is a large volume of evidence supporting a link between higher volumes of surgery, at the unit or consultant level, and lower postoperative mortality rates. There is also good evidence that a higher ratio of nurses to pa-
patients, and standardised care, in particular the Enhanced Recovery Program, are associated with improved outcomes. The evidence in other areas of care organisation is less persuasive, although the quality of communication and collaboration within a unit seems to be important.

While such research has helped advance our understanding of surgical performance, there is a need for further work to help better understand the measurement of performance in colorectal surgery, and to understand how to improve clinical care in daily practice. For example, more research is needed to understand the relationship between operative volume and outcomes in detail. Simply increasing the volume of surgery provided by a unit will not necessarily improve the performance of that particular team. It may be a useful strategy to improve outcomes at the population level, but greater advances in care could perhaps be achieved if the mechanism by which higher volume units achieve better results was understood in detail.

**Thesis questions**

In this context, this thesis seeks to develop a better understanding of unit-level performance in colorectal surgery, with specific focus on trying to answer the following questions:

1. Do patients want to engage in choosing their provider for gastrointestinal cancer surgery?

2. What is the relationship between different outcome measures, varied by definition or patient subgroup, as indicators of unit-level performance after col-
orectal surgery?

3. What particular aspects of the organisation and delivery of care within a colorectal unit are important in determining the outcomes it achieves?

The six original research studies presented in this thesis provide some answers to these questions, and suggest directions for future research to build upon this work.
Part II

Identifying patient preferences
Section overview

This section presents the first original research study in this thesis. This study examines the involvement of gastrointestinal tract cancer patients in provider choice for tests and surgery. It also assesses which aspects of provider performance are important to patients. This study helps clarify the current role of provider choice in the NHS in England, and helps better understand whether the priorities of patients are similar to those of the research community.
Chapter 4

Study 1: Decision-making and information needs about providers of surgical care

4.1 Chapter overview

This chapter describes a patient questionnaire study, designed to investigate the current status of patient provider choice in the NHS. The study also explored the information about providers or surgical teams that patients considered important. As clinicians and health care researchers, it is important to understand patients’ experiences and perspectives in these areas. The results of this study will also help establish whether the areas of care considered important by patients are well aligned with the focus of the research and clinical communities.

4.2 Introduction

In 2009, provider choice was made a formal part of the NHS Constitution\textsuperscript{128}. Over recent years, successive UK governments have made great efforts to facilitate patient choice, for various reasons: to improve efficiency; to reduce waiting times;
to improve provider responsiveness to patients; and potentially to add value that patients may attach to participating in health care choices\textsuperscript{129,130}. Certain of these objectives rely on patients acting as informed consumers, behaving according to rational decision theory to maximise their own interests. However, patients may not make health care decisions in this way for a variety of reasons. They may struggle to obtain and understand complex health care information\textsuperscript{129,131,132}. They may also be unwilling or too ill to take responsibility for such important decisions, preferring to rely on the authority and expertise of their doctor.

Previous work has begun to define patient awareness and experience of choice in the NHS. Approximately half of patients referred for an outpatient appointment by their General Practitioner (GP) are aware of their right to choose their provider\textsuperscript{129,133,134}. Patients and the public consider choice important\textsuperscript{129,135}, and over 80% of participants in two supported choice schemes within the NHS recommended their respective programs\textsuperscript{136,137}. However, data collected on behalf of the Department of Health suggests that a maximum of 50% of patients receiving an outpatient appointment were offered choice, and this may have fallen recently\textsuperscript{133,134}. Reports about whether patients have enough information to make decisions vary widely; recent data shows that 60-89% of respondents reported that they had enough information to help them make provider choices\textsuperscript{129,133}, whereas earlier international work reported that only 35% of UK respondents had sufficient information to choose a hospital\textsuperscript{135}.

Alongside measures to increase patient choice, there has been significant growth in the quantity of information available about providers. In part, this has occurred to help facilitate meaningful patient choice, as patient ‘consumers’ need relevant information upon which to base decisions, although some evidence suggests they
may not use such information even when it is available\textsuperscript{138}. Other benefits include increased transparency and direct effects to drive quality improvement among providers, unrelated to market-driven competition for patients\textsuperscript{139–141}. Although there has been a huge drive to produce information, less attention has been paid to what information patients actually value, and how they might assess and use such information.

This study set out to examine provider choice experiences and preferences, as well as identifying what information was considered important by patients that had undergone gastrointestinal (GI) cancer surgery in England. Previous research examining choice among cancer patients has typically focused on treatments\textsuperscript{142}, rather than providers. Patients with cancer have largely been excluded from studies investigating provider choice. Indeed, patients being referred from primary to secondary care with suspected cancer follow a separate, urgent pathway outside the ‘Choose and book’ system in the NHS which may significantly limit the opportunity to participate in provider choice for these patients. This study aimed to increase understanding of the preferences of these patients to help healthcare providers meet their needs more effectively, and to expand understanding of the experience and role of patient choice in the NHS.

**Study questions**

- What role do GI cancer patients currently experience, and what role would they like, in deciding between providers for their investigations and surgery?

- What provider-level information do these patients consider important when due to undergo surgery, from the data already available within the NHS?
4.3 Methods

4.3.1 Study design

This cross-sectional questionnaire was administered to patients who had undergone surgical treatment for cancer of the upper or lower GI tract in the NHS in England. Patients were recruited via two routes: by postal invites through 5 NHS Trusts; and online via two patient support groups. Invited NHS Trusts each posted 50 surveys to patients they had treated (3 and 2 Trusts included upper or lower GI cancer patients, respectively). Three Trusts were large teaching hospitals (2 in the capital), and two were district general hospitals. Contemporaneously, two patient support groups publicised the study among their members: one lower GI group posted a brief description and web-link on its forum; the other upper GI group sent two emails, two weeks apart, to its members to advertise the study. Invites were distributed in October and November 2014. The study was closed 8 weeks after the last survey was posted.

The study was not designed to measure a specific difference between groups, and therefore no power calculation was appropriate. We aimed to obtain 100 postal responses. With an estimated 40% response rate, this required 250 packs be sent out. It was hoped to recruit a similar number of patients via support groups. No financial incentives were offered to participants.

4.3.2 Questionnaire design

The questionnaire examined two key areas: patients’ self-reported experiences and preferences for involvement in provider choice for tests and surgery; and what in-
formation patients considered important about their provider and surgical team. Participants used the single-item Control Preferences Scale (see table 4.1) to rate their experiences and preferences for involvement in decision making\textsuperscript{143}. The Control Preferences Scale has been widely used to study treatment choice, including among cancer patients\textsuperscript{144}, and the single item version is quick and easy for patients to complete independently. Two decision points were examined. Patients were asked about their involvement in choosing a provider before being diagnosed with cancer, when they had ‘tests to examine their insides’ (decision point 1). They were also asked whether they had been involved in provider choice after they had been diagnosed with cancer, had had further tests to establish their fitness for surgery, and decided to have an operation (decision point 2). These two points were selected because patient may have different experiences and preferences before and after a diagnosis of cancer has been made. At each stage, participants indicated with whom they would like to discuss their decisions.

<table>
<thead>
<tr>
<th>Decision style</th>
<th>Role descriptor</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>I made the decision myself</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>I made the final decision after seriously considering my doctor’s opinion</td>
<td>B</td>
</tr>
<tr>
<td>Shared</td>
<td>My doctor and I shared responsibility for deciding</td>
<td>C</td>
</tr>
<tr>
<td></td>
<td>My doctor made the final decision after seriously considering my opinion</td>
<td>D</td>
</tr>
<tr>
<td>Passive</td>
<td>My doctor made the decision on his / her own</td>
<td>E</td>
</tr>
</tbody>
</table>

Table 4.1: Single-item Control Preferences Scale for decision-making experiences or preferences.

In addition, patients were asked what provider information they considered important when due to undergo surgery. From the publicly available information on NHS Trusts, 23 items were selected, covering a range of structures, processes and
outcomes, at organisational and departmental levels (see table 4.9). Participants were asked how important each item would have been to them once they knew they were going to have surgery. Respondents rated each item on a 5-point Likert scale, from 1 (not important at all) to 5 (critically important). An optional section asked for demographic information, including age, gender, education level and type of cancer. The full protocol is provided in appendix C.

A preliminary postal pilot phase was conducted, with a feedback questionnaire for respondents to complete. Two Trusts each sent out 10 surveys, and 11 (55%) responses were received. This showed that the preliminary protocol was easy to understand, respondents were comfortable with the questions, and it took most respondents less than 15 minutes to complete.

4.3.3 Data analysis

All analysis was conducted using IBM SPSS for Windows, Version 22.0 (IBM Corp., Armonk, USA). Missing data were excluded on a question-by-question basis. At decision point 1, cases were also excluded if respondents indicated that they had seen someone other than their GP initially, as other modes of presentation (such as via the Accident and Emergency department) may have limited the opportunity to participate in provider choice. At both decision points, experience ratings were excluded if other question responses for the same decision point were contradictory. For example, if a participant indicated that their GP had not offered them a choice of provider, yet they rated their experience as ‘My doctor and I shared responsibility for deciding’. This combination of responses was uncommon, and could be prevented using question logic in the online questionnaire system. Of
the postal responses received, only 3 and 9 contradictory responses were identified when being referred for tests or surgery, respectively.

The characteristics of respondents were examined, where provided. It was expected that there may be demographic and/or question response differences between postal and online participants. Therefore, an initial comparison between these groups was performed, before deciding whether to combine their responses for the main data analysis.

Frequencies of responses were determined. The degree of matching between reported experiences and preferences was examined. Cases were excluded from this part of the analysis if they did not include a valid experience rating and an associated preference response. Factors associated with decision-making experiences and preferences were investigated using hierarchical regression. This examined patient factors as fixed effects, allowing for potentially different experiences and preferences among postal and online respondents (included in the model as a random intercept). Patient factors with more than 2 categories were collapsed into fewer, larger groups for this analysis.

Ordinal importance ratings for information items were tested for normality, and appropriate descriptive statistics reported. The relationship between decision-making preferences when being referred for surgery and reported information needs was examined using Spearman’s rank correlation coefficient.

For all tests, a $p$-value of $<0.05$ was considered statistically significant.
4.3.4 Ethical approval

Ethical approval was granted by the South West Research Ethics Committee on 21st July 2014 (reference: 14/SW/1043). All survey responses were anonymous. Postal participants provided implicit consent by completing and returning the survey. Online participants indicated consent after reading the information page before proceeding to the questionnaire.

4.4 Results

4.4.1 Response rates and participant characteristics

463 responses were included in the analysis. 118 (47.2%) valid responses were received from 250 surveys posted. Online response rates could only be calculated for the upper GI cancer support group. 256 (23.1%) responses could be attributed to 1108 emails delivered. Response rates were significantly higher for the postal survey ($p<0.001$, 2-tailed $\chi^2$ test).

Participant characteristics are provided in table 4.2. The majority of respondents were aged between 50 and 69 years (273 of 424 (64.4%)), and most were male (286 of 419 (68.3%)). Recruitment through the upper GI cancer support group’s email list resulted in a sample dominated by this cancer type (334 of 415 (80.5%) responses).

Univariate analysis showed that online participants were younger ($p=0.002$, $\chi^2$ test), more highly educated ($p<0.001$), and had their surgery longer ago ($p<0.001$) than postal respondents. To check the validity of combining online and postal responses, decision-making experiences and preferences were compared. Online re-
| Table 4.2: Respondent characteristics. GI - gastrointestinal. |

<table>
<thead>
<tr>
<th>Recruitment method</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post</td>
<td>118</td>
<td>25.5</td>
</tr>
<tr>
<td>Online</td>
<td>345</td>
<td>74.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;40</td>
<td>13</td>
<td>3.1</td>
</tr>
<tr>
<td>40-49</td>
<td>27</td>
<td>6.4</td>
</tr>
<tr>
<td>50-59</td>
<td>114</td>
<td>26.9</td>
</tr>
<tr>
<td>60-69</td>
<td>159</td>
<td>37.5</td>
</tr>
<tr>
<td>70-79</td>
<td>93</td>
<td>21.9</td>
</tr>
<tr>
<td>≥80</td>
<td>18</td>
<td>4.2</td>
</tr>
<tr>
<td>Missing</td>
<td>39</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sex</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>286</td>
<td>68.3</td>
</tr>
<tr>
<td>Female</td>
<td>133</td>
<td>31.7</td>
</tr>
<tr>
<td>Missing</td>
<td>44</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Secondary school</td>
<td>160</td>
<td>38.3</td>
</tr>
<tr>
<td>Vocational qualification</td>
<td>109</td>
<td>26.1</td>
</tr>
<tr>
<td>University</td>
<td>108</td>
<td>25.8</td>
</tr>
<tr>
<td>Postgraduate degree</td>
<td>41</td>
<td>9.8</td>
</tr>
<tr>
<td>Missing</td>
<td>45</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cancer type</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper GI</td>
<td>334</td>
<td>80.5</td>
</tr>
<tr>
<td>Lower GI</td>
<td>81</td>
<td>19.5</td>
</tr>
<tr>
<td>Missing</td>
<td>48</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time since surgery (years)</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;1</td>
<td>128</td>
<td>31.4</td>
</tr>
<tr>
<td>1-2</td>
<td>153</td>
<td>37.5</td>
</tr>
<tr>
<td>3-5</td>
<td>66</td>
<td>16.2</td>
</tr>
<tr>
<td>6-10</td>
<td>37</td>
<td>9.1</td>
</tr>
<tr>
<td>&gt;10</td>
<td>24</td>
<td>5.9</td>
</tr>
<tr>
<td>Missing</td>
<td>55</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cured by surgery</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>306</td>
<td>74.5</td>
</tr>
<tr>
<td>No</td>
<td>38</td>
<td>9.2</td>
</tr>
<tr>
<td>Unsure</td>
<td>67</td>
<td>16.3</td>
</tr>
<tr>
<td>Missing</td>
<td>52</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Any health care-related harm</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>66</td>
<td>15.8</td>
</tr>
<tr>
<td>No</td>
<td>309</td>
<td>74.1</td>
</tr>
<tr>
<td>Unsure</td>
<td>42</td>
<td>10.1</td>
</tr>
<tr>
<td>Missing</td>
<td>46</td>
<td></td>
</tr>
</tbody>
</table>
respondents reported similar experiences of decision-making ($p=0.660$ and $p=0.534$ at decision points 1 and 2, respectively; Mann-Whitney U Test), but expressed preferences for greater involvement than postal respondents ($p<0.001$ at both decision points). However, both postal and online groups expressed the most frequent preference for shared decision-making, with similar overall distributions of responses. Therefore, all subsequent analysis was performed on combined postal and online responses.

### 4.4.2 Provider choice

Participants indicated whether they were given a choice of provider at both decision points (see table 4.3). 71 of 463 (15.3%) respondents indicated that their initial consultation was not with their GP, with several reporting presenting as an emergency via the Accident and Emergency department.

<table>
<thead>
<tr>
<th>Decision point 1</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>When you saw your General Practitioner (GP) initially, were you given the choice of which hospital or clinic to go to?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>93</td>
<td>20.1</td>
</tr>
<tr>
<td>No</td>
<td>298</td>
<td>64.5</td>
</tr>
<tr>
<td>Saw somebody else</td>
<td>71</td>
<td>15.4</td>
</tr>
<tr>
<td>Missing</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>Total</td>
<td>463</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Decision point 2</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Once you had these results and decided you would have an operation, did your specialist give you a choice of which hospital to go to for surgery?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>66</td>
<td>14.9</td>
</tr>
<tr>
<td>No</td>
<td>355</td>
<td>80.3</td>
</tr>
<tr>
<td>Unsure</td>
<td>21</td>
<td>4.8</td>
</tr>
<tr>
<td>Missing</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>463</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3: Provision of choice of provider at different decision points.
Experiences and preferences

Experience and preference ratings for involvement in decision-making are shown in table 4.4. Patients reported very low levels of involvement in choice of provider, with their clinician deciding where they underwent tests or surgery in 77.0% and 81.8% of cases. In sharp contrast, only 11.4% and 14.0% of patients wanted their doctor to make provider choices for them, at decision points 1 and 2 respectively. Experience ratings were significantly different from preference ratings at both points of the pathway ($p<0.001$ each, 2-tailed $\chi^2$ test).

<table>
<thead>
<tr>
<th>Role</th>
<th>Decision point 1</th>
<th>Decision point 2</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experience</td>
<td>Preference</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td>Experience</td>
<td>Preference</td>
<td>n</td>
</tr>
<tr>
<td>A</td>
<td>40</td>
<td>10.4</td>
<td>21</td>
</tr>
<tr>
<td>B</td>
<td>17</td>
<td>4.4</td>
<td>81</td>
</tr>
<tr>
<td>C</td>
<td>25</td>
<td>6.5</td>
<td>183</td>
</tr>
<tr>
<td>D</td>
<td>6</td>
<td>1.6</td>
<td>50</td>
</tr>
<tr>
<td>E</td>
<td>295</td>
<td>77.0</td>
<td>43</td>
</tr>
<tr>
<td>Missing</td>
<td>2</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>385</td>
<td>391</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4: Frequencies of decision-making role experiences and preferences at different decision points.
See table 4.1 for descriptions of roles A-E; p-value calculated using two-tailed $\chi^2$ test.

Experience-preference matching

The relationship between experience and preference ratings was examined. A detailed breakdown is shown in table 4.5. A summary of the degree of matching between experiences and preferences is provided in table 4.6. At both decision points, the majority of patients reported a preference for greater involvement in decision-making about providers than they had experienced (253 of 371 (68.2%))
and 297 of 423 (70.2%) for tests and surgery, respectively).

<table>
<thead>
<tr>
<th>Decision point 1</th>
<th>Preference</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>A</td>
<td>13</td>
<td>12</td>
<td>13</td>
<td>1</td>
<td>0</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0</td>
<td>12</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0</td>
<td>4</td>
<td>18</td>
<td>1</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>7</td>
<td>52</td>
<td>142</td>
<td>44</td>
<td>40</td>
<td>285</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>20</td>
<td>81</td>
<td>181</td>
<td>47</td>
<td>42</td>
<td>371</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Decision point 2</th>
<th>Preference</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experience</td>
<td>A</td>
<td>12</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>0</td>
<td>19</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>0</td>
<td>6</td>
<td>18</td>
<td>3</td>
<td>1</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>16</td>
<td>65</td>
<td>159</td>
<td>49</td>
<td>58</td>
<td>347</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>28</td>
<td>95</td>
<td>185</td>
<td>55</td>
<td>60</td>
<td>423</td>
</tr>
</tbody>
</table>

Table 4.5: Relationship between decision-making role experiences and preferences at different decision points.
See table 4.1 for description of roles A-E.

**Discussing decisions**

Participants were asked with whom they would like to discuss decisions about choosing a provider (see table 4.7). When being referred for tests, family and GPs were most frequently selected. Once a decision to undergo surgery had been made, family remained the most common response, with similar numbers wanting to discuss choice with a specialist hospital doctor. GPs and specialist hospital nurses were also important, though they were selected less often.
<table>
<thead>
<tr>
<th></th>
<th>Decision point 1</th>
<th>Decision point 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Prefer less involvement</td>
<td>34</td>
<td>9.2</td>
</tr>
<tr>
<td>Preference matched</td>
<td>84</td>
<td>22.6</td>
</tr>
<tr>
<td>Prefer more involvement</td>
<td>253</td>
<td>68.2</td>
</tr>
<tr>
<td>Total</td>
<td>371</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.6: Degree of matching between decision-making role experiences and preferences at different decision points.

<table>
<thead>
<tr>
<th></th>
<th>Decision point 1</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spouse / family</td>
<td>195</td>
<td></td>
</tr>
<tr>
<td>GP</td>
<td>194</td>
<td></td>
</tr>
<tr>
<td>Nobody else</td>
<td>34</td>
<td></td>
</tr>
<tr>
<td>Practice nurse</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>36</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Decision point 2</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spouse / family</td>
<td>223</td>
<td></td>
</tr>
<tr>
<td>Specialist hospital doctor</td>
<td>203</td>
<td></td>
</tr>
<tr>
<td>GP</td>
<td>97</td>
<td></td>
</tr>
<tr>
<td>Specialist hospital nurse</td>
<td>68</td>
<td></td>
</tr>
<tr>
<td>Nobody else</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>29</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.7: Response frequencies indicating with whom respondents would like to discuss decision-making at different decision points.
4.4.3 Factors associated with experiences and preferences

Patient factors were collapsed into the following groups: age $<60$ versus $\geq 60$ years; secondary school or vocational education versus university or postgraduate degree; $<1$ versus $1-2$ versus $\geq 3$ years since surgery; self-rated cure versus other rating; no self-rated health care-related harm versus other rating.

Hierarchical analysis results, showing the significance of the association of all examined factors with experiences or preferences at each decision point, are provided in table 4.8. This analysis showed that reported experiences were not significantly associated with patient gender, education, age, cancer type, time since surgery, cure status or health care-related harm. However, there were statistically significant associations with patients’ reported preferences. At decision point 1, non-cured status was significantly associated with preference for more active involvement in provider choice when being referred for tests ($p=0.031$). At decision point 2, non-cured status and female sex were associated with an expressed desire for more active involvement in provider choice for surgery ($p=0.002$ and $p=0.046$, respectively). Cancer type, patient age, education, health care-related harm and time since surgery were not associated with preferences at either point. The extremes of preferences predicted by the model were examined across both decision points. Patients with the lowest predicted preferences for involvement wanted their doctor to choose in 26.9%-33.3% of cases. Therefore, 66.7%-73.1% of such patients wanted some degree of involvement in provider choice. Conversely, those with the highest preference for involvement only wanted their doctor to decide in 1.6%-4.0% of cases.
Table 4.8: Hierarchical regression results showing the significance of association between individual variables and reported experiences and preferences at different decision points.

* - p<0.05.

### 4.4.4 Information needs

430 responses were analysed (33 excluded returns contained no ratings). Kolmogorov-Smirnov Tests showed that all ratings were non-normally distributed. Mean, median and interquartile ranges are provided in table 4.9. The five highest rated items, in order, were how long patients with respondents’ type of cancer waited for treatment, hospital rates of infection with ‘super-bugs’, the number of operations where a foreign body was left behind each year, annual volume for respondents’ type of cancer operation, and postoperative mortality rates. The five lowest rated items, in order, were how much of the hospital’s waste was sent for recycling, whether the hospital balanced its income and spending, how many beds there were in the hospital, how full the hospital usually was, and how much it cost to park there.
<table>
<thead>
<tr>
<th>Organisation level</th>
<th>Structure</th>
<th>Process</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>How many beds there are in the hospital</td>
<td>2.38</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>How many consultants there are in the hospital</td>
<td>3.20</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>How many nurses, midwives and health visitors there are in the hospital</td>
<td>3.02</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>How many nurses, midwives and health visitors there are per bed in the hospital</td>
<td>3.46</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>How much it costs to park at the hospital per hour</td>
<td>2.68</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>How long it takes me to get to the hospital from my home</td>
<td>2.89</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>How full the hospital usually is</td>
<td>2.66</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>How often staff report things that may have gone wrong in the hospital</td>
<td>3.85</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>How often the hospital checks whether patients are at risk of clots on the legs or lungs while in the hospital</td>
<td>4.12</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Rates of infection with 'super-bugs' in the hospital</td>
<td>4.15</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Whether the hospital successfully balances its income and spending</td>
<td>2.35</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>How satisfied staff are with the care they deliver across the whole hospital</td>
<td>3.90</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Previous patients satisfaction with their care across the whole hospital</td>
<td>3.70</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Whether the whole hospital has an average number of deaths, or more or less than expected</td>
<td>3.59</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>How often patients are readmitted to hospital soon after receiving any form of care provided</td>
<td>3.63</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>How much of the hospital's waste is sent for recycling</td>
<td>2.12</td>
<td>2</td>
</tr>
<tr>
<td>Department level</td>
<td>Structure</td>
<td>Process</td>
<td>Outcome</td>
</tr>
<tr>
<td></td>
<td>How many operations of this type of cancer operation the team performs each year</td>
<td>4.14</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>How many operations of any type were cancelled at the last minute in the hospital over the previous year</td>
<td>3.43</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>How long patients with your type of cancer wait for treatment</td>
<td>4.395</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>How often patients die soon after your type of cancer operation in this hospital</td>
<td>4.14</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>How well previous patients with your type of cancer rated the hospital</td>
<td>3.96</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>How often patients with any form of surgery stay in much longer than expected</td>
<td>3.54</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>The number of operations each year where the doctors and nurses accidentally left something behind</td>
<td>4.15</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 4.9: Participant ratings of importance of information items. Scale from 1 (not important at all) to 5 (critically important). Med - median; IQR - interquartile range; n - valid responses.
4.4.5 Relationship between decision-making and information needs

After applying the Bonferroni correction for multiple comparisons (resulting in a significant $p$-value of $0.05/23=0.002$), there was significant, moderate ($0.1<\rho<0.3$) correlation between decision-making preferences and 8 information items. There were associations with the following hospital-level information items: rates of infection with ‘super-bugs’, hospital-wide death rates, hospital readmission rates and error reporting by staff. There were also associations with surgery-specific information: annual operative volume, post-operative mortality, longer-than-expected stay after any form of surgery, and last-minute cancellation of operations. In all cases, the direction of correlation suggested that patients who wanted more active involvement in decisions rated the above information items as more important than those who preferred a more passive role.

4.5 Discussion

This study has documented a dramatic difference between patients’ current level of involvement in provider choice, and the level of involvement patients indicated they would like. In over 75% of cases, respondents indicated that their clinicians decided where they went to have tests or surgery without any patient input. Only 11.4%-14.0% of patients wanted their doctors to make these decisions for them in this way. When respondents’ experiences and preferences were compared at the individual level, more than two-thirds reported that they would have liked to be more actively involved in provider choice. This suggests that the political
discourse around increasing patient choice and using competition to drive quality improvements has not translated into greater involvement in provider choice for the patients included in this study.

The strengths of this study include the large sample and relatively good response rates obtained\textsuperscript{145–148}. The study has ecological validity as participants were reflecting on real-world experiences of health care. The study also has important limitations. The study cohort was younger than the typical patient population for oesophageal, gastric or colorectal cancer\textsuperscript{149}, as well as being more highly educated than the general population in England and Wales\textsuperscript{150}. Patients rated their experiences and preferences using the single-item Control Preferences Scale. This simple and easily administered tool met the needs of this study, but does not provide a detailed understanding of how patients want to exercise choice. Patients may have expressed certain preferences and opinions when asked in this study, yet they may behave differently at the time of making a particular decision. Specific biases may have affected participants’ responses. The lack of association between time since surgery and participant’s responses suggests that recall bias was not a major problem. The association between reported preferences and cure status may be interpreted as evidence of hindsight bias. However, examination of the extremes of preference predicted by the model showed that this bias did not alter the key finding of the study, that the majority of patients reported a desire to be involved in decision-making about their provider. While the study examined the importance patients attached to information items, it did not assess the nature or adequacy of information actually received. It also did not ask respondents whether they would use the selected information presented to guide their choice of provider. There is evidence that even when relevant information is available,
patients do not seek or use it to make provider decisions. Postal and online responses were analysed together, although online participants, as members of patient support groups, may have self-selected as more active participants in health care. Hierarchical modelling was used to allow for this when investigating factors associated with decision-making experiences and preferences.

The participants in this study expressed a clear preference for being involved in decision-making, with 43.9% to 48.4% of respondents indicating a preference for equal sharing of decisions, although only 6.5% to 6.8% reported experiencing this level of involvement. The validity of comparison with other studies focused on treatment decisions is not clear. With this in mind, the present results are similar to those reported in a recent review of decision-making preferences across a broad range of patient groups. However, in contrast to the present findings, a colorectal cancer-specific review reported that the majority of patients preferred a passive role in decision-making, based on studies published between 2003 and 2008. The discrepancy between these findings may represent a shift in the preferences of patients with GI cancer over time. Expectations about the responsiveness of health care providers may have changed, in line with the political emphasis on choice over the last decade. The high levels of involvement desired in the present study group contrast with reported reluctance among GPs to offer choice in all circumstances, especially when specialist treatment is required.

In the vast majority of cases, patients did not experience the level of involvement in decision-making that they preferred; only 22.6% to 26.0% of respondents reported matching of role preferences and experiences. These figures are much lower than the 63% preference-matching reported in a recent review. As preferences were similar, this difference arose because participants in the present study
perceived much lower rates of involvement in decision-making.

Mismatches between preferences and experiences of decision-making have importance beyond the basic moral imperative to respect patients’ autonomy and to deliver patient-centred care. When patients’ experiences match the level of involvement that they prefer, post-consultation anxiety is significantly reduced, compared with experience-preference role mismatch$^{151}$. Shared decision-making may improve patient satisfaction and reduce conflict over decisions$^{152}$. Qualitative work suggests involvement in decision making is important because it helps patients develop a sense of control, reduce anxiety, and plan for the future, even when there are no real choices$^{153}$. Shared decision-making may be achieved in different ways, depending upon context, although common factors include an interactive exchange, based within a trusting relationship, resulting in a clear decision being made$^{154}$.

Participants most wanted to discuss decisions with their GP and family when being referred for tests, or with their specialist doctor and family when deciding where to have surgery. These findings are consistent with previous similar work$^{129,132,134,136}$. The importance of family in helping patients make decisions should not be neglected. When faced with complex and high-stakes decisions, patients may prefer to base choices on personal experience, intuition and anecdote rather than complicated technical data$^{129,131,139}$. However, patients recognise that their doctors have important knowledge and expertise that are necessary to help them make sense of the decisions they face$^{132}$. The pivotal roles of GPs and hospital specialist doctors are clear, if the degree of mismatch between patients’ experiences and preferences is to be reduced. Efforts should focus not just on improving patient involvement, but also on how clinicians make referral decisions;
many doctors distrust performance data and base referral practices on experience and personal recommendation\textsuperscript{129,155}. Our results highlight that doctors should not make assumptions about which patients want to be involved in provider choice. Education and age were not associated with preferences, and responses covered the full range available. Clinicians should be encouraged to increase involvement, but this should be tailored to the needs of individual patients.

Among the information items examined, participants attached primary importance to data related to their surgery. Others have reported similar findings, when participants were asked what information was important to them or helped them determine provider choice\textsuperscript{129–131,136,137,156}. For example, Boyce et al reported MRSA blood stream infection and mortality rates were the two highest rated information items in their investigation of online information provision and provider choice\textsuperscript{131}. Others reported a similar focus on cleanliness, success rates of treatment and high quality of care\textsuperscript{129,136}. Variable ratings between studies, and the finding that information preferences are not stable over time\textsuperscript{131}, suggest there can be no fixed recommendation about what information should be provided for patients. Information needs are also likely to be highly personal.

Participants who reported a preference for greater involvement in provider choice for surgery attached greater importance to a number of specific information items. These included key surgical information such as operative volume and postoperative mortality. However, this study did not ask what information patients would use to guide their choice of provider. It was first necessary to establish whether patients wanted to be involved in choice, before further work can focus on this in more detail. Some information patients considered important may lack the potential to discriminate between high and low quality providers. For example,
the rate of retained foreign bodies after surgery is very low, and variation between Trusts may be due to chance rather than quality. In addition, hospitals with higher rates may simply be better at recording such events, rather than being any less safe than their peers. Instead of examining how information may be used to inform choice, the present study addressed information needs broadly, and the findings can help better inform all patients, whether or not they want to exercise choice. The information examined is already freely available, and it would be easy to present patients with a summary of the top 5 rated information items. This data could easily be collated by NHS Choices, or by an independent organisation such as Dr Foster Intelligence.

Our results also have implications for policy makers. Very low levels of involvement suggest that patient provider choice cannot currently be an effective driver for quality improvement. If competition through patient choice is to be promoted, much needs to be done to increase patient involvement, and to provide appropriate information and support to facilitate provider choice. There may also need to be appropriate changes to referral systems to provide cancer patients with similar access to provider choice as already exists for benign conditions through ‘Choose and book’.

Given the limitations of this retrospective study, confirmation with prospective data is warranted. The present findings may be generalised with caution to other patient groups within the UK. However, expectations about involvement in health care decisions may be shaped by broad political and socio-cultural factors. Therefore, further work is needed to establish generalisation to health care systems in other countries.

This study has documented an unmet desire for greater involvement in decision-
making regarding provider choice among GI cancer patients. These patients want to discuss these decisions with their family, GPs and hospital specialists. They consider specific information about treatment waiting times, operative volume and postoperative mortality rates to be important, as well as safety information on retained foreign bodies at surgery, and hospital infection rates. Increased involvement in decision making may improve patient satisfaction and reduce anxiety at a difficult time for patients. Limited current involvement makes it unlikely that patient provider choice is an effective lever for quality improvement at present.

4.6 Chapter summary

This patient survey study examined provider choice and information needs among GI cancer patients in the English NHS. The majority of participants reported much lower involvement in decision-making than they reported they would like, and they tended to consider surgery-specific information to be more important than broader hospital-level data. Much work must be done if provider choice is to operate as a driver for quality improvement, as intended by policy makers.
Part III

Investigating colorectal outcomes
Section overview

This section of the thesis presents three original research studies examining mortality rates after colorectal surgery using routinely collected administrative data from the NHS in England. The first chapter provides background information about the database and how it has been used in this thesis. Study 2 explores trends in postoperative mortality over a 14 year period between 1998 and 2012, to better understand some of the potential reasons for the observed changes. Study 3 examines the impact of varying the period of follow-up for mortality upon the identification of individual units as performance outliers. Study 4 assesses the relationship between elective and non-elective outcomes at the institutional level. Together, these studies help develop a more refined understanding of the complexity of unit-level performance in colorectal surgery.
Chapter 5

Using routine administrative data

5.1 Chapter overview

The first part of this chapter explains about the origins, contents, and strengths and weaknesses of the Hospital Episode Statistics (HES) database. The second part then describes how HES data has been used in this thesis, covering the different phases of data extraction, processing and analysis. Specific considerations relevant to the individual studies presented in subsequent chapters are discussed within the relevant methods sections for each study.

5.2 About the HES database

5.2.1 Origins

In 1987, the HES dataset was established to gather information on all NHS activity involving admission to hospital\textsuperscript{157}. Data was initially collated at a regional level, but since 1996 this information has been returned to a central data repository. Over subsequent years, further areas of NHS activity have been added to the database, which now includes details of outpatient clinic and Accident and Emergency department attendances. The contents of this database are described in section 5.2.2.
HES is an electronic database, and information is entered at individual NHS institutions. Typically, organisations employ ‘coders’ who transfer information from paper records to a series of data fields within the computerised database. Locally stored data is then submitted centrally on a quarterly basis, which is organised according to the financial year.

HES data is currently managed centrally by the Health and Social Care Information Centre (HSCIC), a publicly funded body. Patient-level data, such as date of birth, is confidential as it carries the risk of allowing the identification of individual patients. Access to such data may be granted if the patients involved provide consent, or if approval is granted under section 251 of the NHS Act 2006. This allows the common law duty of confidentiality to be set aside for defined medical purposes. Patient identifiable information may be disclosed where anonymised data does not meet the intended requirements, and seeking consent is not practical, such as when the sample size is too large. Release of data is conditional upon appropriate local arrangements for its secure management. If data are to be used for research purposes, approval must also be obtained from the local Research Ethics Committee. Mr Faiz, one of the supervisors of this thesis, holds approval for outcomes and quality of care research using HES data under section 251 of the NHS Act, and approval has also been granted by the London - Queen Square Research Ethics Committee (reference number: 13/LO/1235).

5.2.2 Contents

Within HES, the unit of data collection is the ‘Finished Consultant Episode’ (FCE). This represents a period of care for an individual patient under a specific
consultant, with an associated start and end date. This ‘episode’ may correspond directly with an entire admission to hospital, or in other cases it may represent only part of an admission, when a patient has been transferred from the care of one consultant to another. A single period of admission to hospital is called a ‘spell’, which may consist of one or more FCEs. If a patient is transferred from one hospital to another, it is possible to link the two spells in to one ‘superspell’ using an unique patient identifier.

Each individual HES record contains a large number of data fields. These include patient demographics, admission details, diagnostic codes, and procedure codes with dates, as well as information about the healthcare provider. Diagnoses are recorded using codes from the International Classification of Diseases and Related Health Problems (Tenth Revision) (ICD-10)\(^{158}\), and are linked to the entire episode of care without specific dates. Procedures are recorded according to the OPCS Classifications of Interventions and Procedures, Version 4 (OPCS-4). Updated versions of this classification have been released and adopted within the NHS over subsequent years\(^{159}\). The most recent version, 4.7, was adopted from April 2014. The OPCS manual covers the full range of surgical operations as well as other procedures such as radiological or endoscopic tests or interventions. Within HES, individual OPCS codes are recorded with an associated date of procedure, allowing a sequence of events to be constructed for individual patients.

For outcome analysis, in-hospital deaths are easily identified using HES data. To capture post-discharge deaths, HSCIC can provide HES data extracts that are linked to Office of National Statistics (ONS) records using unique patient-level information. This allows extension of the follow-up period for examining mortality rates beyond the initial period of hospital care.
5.2.3 Data strengths and weaknesses

While there is increasing acceptance of research conducted using administrative data within the health care profession, the results of such studies should be interpreted critically. There are clear benefits associated with using HES data, as well as potential weaknesses, and these are outlined below.

Much of the work that has been done to establish the utility of administrative data, such as HES, has involved comparison with voluntary diseases registers, or other prospective, clinical data collection systems, such as NSQIP in the United States. For large-scale comparison of performance, these are the only available sources of information. Given the clinical focus of specific disease registers, use of such databases for performance analysis is widely accepted by practising clinicians. By comparison, many express distrust of administrative data. In this context, it is not surprising that much work has focused on a comparison between the two. However, technically, comparison of administrative data to disease registers does not demonstrate the scientific validity of administrative data for the purpose of assessing performance in health care. Rather, such comparisons may establish the accuracy and degree of agreement between the respective datasets. It is important to note that voluntary disease registers have their own problems, and cannot therefore be held up as the ‘gold standard’ against which all else should be compared. Research has shown such databases are subject to selection bias, with those submitting data achieving better results than those who do not\textsuperscript{160,161}. Like administrative data sets, they also suffer from problems with coding accuracy\textsuperscript{162}. Therefore, there are two key objectives when comparing administrative databases to disease registers or other sources of data: firstly, to establish similarity or agree-
ment between them; and secondly, to reassure those who distrust administrative
data that they are an acceptable alternative to disease registry data.

**Strengths**

One of the main strengths of HES is that it represents a population-based data
set. All relevant activity within the NHS should be recorded and returned to the
HSCIC. While it cannot claim to achieve 100% capture (few data sets can), HES
represents the most complete data source on in-hospital activity in England. It
may therefore be considered to avoid (or at least minimise) problems of reporting
bias. Studies have shown that HES captures a greater number of relevant cases
when compared with voluntary disease registers for colorectal cancer\textsuperscript{164} or vascular
surgery\textsuperscript{161}.

Various researchers have demonstrated the accuracy of administrative data for
modelling health care outcomes and comparing institutional performance. Aylin et
al compared predictive models based on HES data with those based on cardiac, vas-
cular and colorectal surgical disease registers\textsuperscript{164}. HES-based models yielded similar
measures of discrimination to models based on clinical registry data. A similar US
study compared NSQIP data with the administrative University HealthSystem
Consortium Clinical Database\textsuperscript{165}. That study showed that while models based on
the clinical database had statistically significantly better discrimination, the ab-
bsolute difference in discrimination was small, and models based on administrative
data were a good fit for the observed outcomes\textsuperscript{165}. Holt and coworkers assessed
the accuracy of data in HES against ONS mortality records and retrospective case-
note review for patients undergoing elective abdominal aortic aneurysm repair\textsuperscript{166}.
They showed excellent capture of mortality in HES, and concluded that HES can
be used to compare mortality between institutions.

A study of patient safety indicators using HES data suggested that clinically important events were captured within this administrative database\(^\text{167}\). Nine indicators were identified within HES, and the length of stay and mortality of affected patients were compared with matched controls. For all but one indicator, both outcomes were significantly worse in affected patients. This study supports the validity of using HES data to assess clinical activity within the NHS.

Audit and research has shown that the quality of coding in administrative databases has improved over time. In England, the Audit Commission regularly assesses the accuracy of data contained within HES. Their 2012 report showed that data accuracy had improved between 2007/8 and 2011/12 for both procedures and diagnoses\(^\text{168}\). An older study of Swiss administrative data examined coding of comorbidities assessed by the Charlson Comorbidity Index, compared with casenote review\(^\text{169}\). The authors reported an overall improvement in coding between 1999 and 2003.

Data linkage between HES and ONS mortality records by HSCIC allows analysis of long-term survival. This contrasts sharply with the highly esteemed NSQIP, which only monitors outcomes during the first 30 days after treatment\(^\text{56}\). Previous research has shown that mortality rates after colorectal surgery are elevated for the whole of the first year after surgery\(^\text{170–172}\). It is therefore important to be able to assess death rates beyond the initial in-hospital stay, or the 30\(^{\text{th}}\) postoperative day. ONS-linked HES extracts provide the opportunity to assess survival at any time point after treatment.

HES data is also relatively cheap and easily available. Provided it has an acceptable degree of accuracy, it would arguably be inexcusable not to use this
existing data source to try and develop a better understanding of health care.

**Weaknesses**

One of the principal concerns about HES data relates to the accuracy of coded information. Within individual NHS Trusts, staff who input data, called ‘coders’, are trained how to extract the relevant information from clinical records and enter this information into the electronic database. The capture of accurate clinical data relies on: legible and comprehensive documentation of relevant information in the medical record by health care workers; the diligence and skill of coders in interpreting this information; and the existence of appropriate diagnostic and procedure codes within the database. Clinicians often express concern about the reliability of this process, and the ability of non-clinical staff to capture clinical detail accurately. Researchers have documented problems in the accuracy of coding of comorbidities within administrative datasets\textsuperscript{166,169}, or the precise nature of individual patients’ pathology\textsuperscript{173}. Nonetheless, as described in the previous section, we know that the quality of coding is improving over time\textsuperscript{168}. It has also been shown that administrative datasets perform similarly to clinical databases in modelling healthcare outcomes\textsuperscript{164,165}. From this, it may be inferred that the quality of coding is adequate for such purposes. Clearly, limits in coding accuracy must be remembered when handling administrative data. A relatively old study of coding within HES reported that the first 3 characters of diagnosis and procedure codes were more reliable than full 4-character codes\textsuperscript{174}. It is expected that coding practices in the present day, 20 years after that study, are much more reliable and accurate. Nonetheless, the principle is still relevant: careful use of broad groups of codes may result in a more accurate picture of clinical activity.
Another common concern about administrative data is the lack of specific clinical variables. For example, the HES database does not contain information on the local or systemic status of cancer (known as its stage), or other aspects of a cancer, such as its histological grade. Across a range of types of cancer, features such as stage of disease have been shown to be significant predictors of survival\textsuperscript{175–178}. Despite this lack of disease-specific information, Aylin et al showed that a model of in-hospital mortality after colorectal cancer surgery based on HES data had similar discriminatory power to an alternative model based on clinical registry data which included stage of disease\textsuperscript{164,179}. This finding suggests that for modelling short-term outcomes, the lack of detailed cancer-specific information is unlikely to pose a significant problem. Even with advanced, incurable colorectal cancer, where surgical removal is possible, median survival of between 11 and 30 months may be expected\textsuperscript{180}. In-hospital mortality, or death within the first 30 or 90 days after surgery, should not be due to progressive cancer. Therefore, if short-term outcomes are being examined, the lack of cancer stage data should not compromise model discrimination. However, if administrative data are being used to study medium- or long-term outcomes, the absence of such information may become increasingly important.

More broadly, a lack of regular engagement by clinical staff may be considered a limitation of administrative data. In a survey of NHS hospital consultants, only 21\% of respondents were regularly involved in clinical coding\textsuperscript{181}. However, with a response rate of only 2.8\%, this finding cannot be generalised; it may be that true engagement is much lower. Regardless, this finding has implications for coding accuracy and broader use of HES data. Were engagement levels much higher, this would help assure the accuracy of coded information. It would also contribute
towards wider acceptance of its use for local and national performance assessment and quality improvement.

There are other limits related to the specific contents of administrative databases. Typically, they do not capture data on specific structures and, in particular, processes of care. Within the HES database, diagnosis codes do not have an associated date. It can therefore be difficult or impossible to establish if a specific diagnosis represents an historic illness, or one acquired during an episode of care. Using pseudonymised patient identifiers, it is possible to link episodes and spells of care from different points in time, to try and address this limitation. However, this method relies on the potentially inaccurate inference that a specific diagnosis was newly made during an episode of care, rather than between episodes. If appropriately completed, it would be much more accurate to use a ‘present on admission’ flag for each diagnosis field, as used by the US Centres for Medicare and Medicaid\textsuperscript{182}. This indicator is to be introduced to HES, once a consultation phase has been analysed and guidance issued\textsuperscript{183}.

5.3 Using HES data

This section explains how HES data has been used in this thesis. It covers the steps required to obtain, process and analyse HES data to examine clinical practice.

5.3.1 HES data extraction and processing

As the HES database contains information on all in-patient care, specific criteria must be used to define a manageable cohort of episodes that may be extracted, processed and analysed. An extract may be defined by any single or combina-
tion of variables from those present within the database, such as diagnosis code, procedure code, episode start and end dates, admission type or patient age. For example, in study 2, only patients undergoing a major colorectal resection for colorectal cancer were selected for inclusion. Definition of this cohort involved looking up all OPCS-4 codes for major colorectal resections (see table 6.1) to identify a list of relevant operation codes. In addition, the relevant diagnosis codes were looked up in the ICD-10 manual: C18 (‘malignant neoplasm of the colon’), C19 (‘malignant neoplasm of the rectosigmoid colon’) and C20 (‘malignant neoplasm of the rectum’). The combination of these codes was used to extract only the data that related to patients undergoing a colorectal resection for bowel cancer.

HES data extracts must be processed before appropriate analysis can be performed to examine specific characteristics or outcomes. The processing required depends upon the requirements of the particular study. However, certain processing steps are common to all the HES-based studies reported in this thesis. These include:

- Removing duplicate episodes.
- Removing cases with invalid codes for patient sex.
- Identifying the operation of interest from the multiple operations or procedures recorded within each individual episode, and recording the associated date.
- Removing additional episodes relating to the same admission or spell.
- Where patients have had more than one eligible operation, selecting the first of these and discarding subsequent procedures. Patients may have more than
one eligible procedure within one episode, or across multiple episodes.

An illustration of the number of episodes and patients included at each of the above steps is provided in table 5.1. For this example, the baseline cohort was defined as all major colorectal resections for colorectal cancer in adults between 1st April 1998 and 31st March 2012. During that period, 4836 (1.8%) patients had two or more eligible procedures.

<table>
<thead>
<tr>
<th>Processing step</th>
<th>Episodes</th>
<th>Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline cohort</td>
<td>269 077</td>
<td>262 702</td>
</tr>
<tr>
<td>Remove duplicates</td>
<td>268 589</td>
<td>262 702</td>
</tr>
<tr>
<td>Remove invalid sex codes</td>
<td>268 533</td>
<td>262 646</td>
</tr>
<tr>
<td>Remove additional episodes relating to same admission</td>
<td>267 598</td>
<td>262 646</td>
</tr>
<tr>
<td>Remove additional episodes relating to eligible resections after a patient’s first major resection</td>
<td>262 646</td>
<td>262 646</td>
</tr>
</tbody>
</table>

Table 5.1: Basic HES data processing steps and illustrative numbers.

Once these preliminary steps have been completed, the specific variables of interest for a particular study may be derived. In this thesis, variables were usually grouped for analysis as categorical variables. These commonly included:

- **Age** To preserve anonymity, the HES extract does not provide the date of birth of individual patients. Instead, age in years is provided at the beginning and end of each episode, in the ‘startage’ and ‘endage’ fields. For all studies in this thesis, ‘startage’ was used for age-based analysis.

- **Charlson Comorbidity Index** ICD-10 codes indicating diagnoses included in the Charlson Comorbidity Index (CCI) may be identified using the codes described by Quan et al\textsuperscript{184}. In this thesis, the original weights for each comorbidity were used to calculate the total score\textsuperscript{185}.  

• **Urgency of admission** There are 17 codes for the HES variable ‘admimeth’, describing the urgency and source of admission to hospital. For studies in this thesis, codes were grouped into elective and non-elective categories. Elective admissions were identified using codes 11, 12 and 13 for ‘Elective: from waiting list’, ‘Elective: booked’ and ‘Elective: planned’, respectively. Non-elective admissions were defined as any other admission type, such as those with an emergency admission code, or those where the patient was admitted from another hospital. All episodes with an obstetric admission code were recorded as non-elective, as gastrointestinal surgery performed during such an admission can be assumed to be unplanned. Only episodes coding admission from a high security psychiatric hospital or from an unknown source were excluded.

Other variables, such as operation type, surgical approach, year of procedure and Trust codes were processed as described in individual studies.

### 5.3.2 HES data analysis

Once a defined extract has been processed, it is ready for analysis. In this thesis, all statistical analysis was performed using IBM SPSS Statistics for Windows, Versions 20.0 or 22.0 (IBM, Armonk, New York, USA). In study 2, all analysis was performed on unadjusted outcomes. A description of the simple techniques used is provided in that chapter. In studies 3 and 4, more sophisticated analysis was performed, in a similar way for each study. Multiple regression models were used to identify risk-factors for specific outcomes, and risk-adjusted outcomes were derived for comparison of performance between institutions. The techniques used
are described in the following parts of this section.

**Identifying risk factors**

Risk-factors for outcomes were identified using regression models. As postoperative mortality is a two-way outcome (a patient is either alive or dead at a particular point in time), binary logistic regression was used. Categorical independent variables were assessed individually for association with the dependent outcome variable, to identify any preliminary relationship, before variables were entered into multiple regression models to allow assessment of the individual impact of specific variables. Results of such regressions are reported as odds ratios (OR) for individual variables, with 95% confidence intervals (CI) and \( p \)-values.

**Testing model fit**

The goodness-of-fit of individual logistic regression models was assessed using the Receiver Operating Characteristic (ROC) Curve function in IBM SPSS. This function calculates the area under the ROC curve, also known as the c-statistic. This provides an indication of the ability of the model to accurately predict the observed outcomes. A value of 0.5 indicates that the model is no better than chance, whereas a value of 1.0 indicates perfect discrimination. A c-statistic less than 0.7 is considered to reflect poor model fit, 0.7 to 0.8 indicates moderate fit, and values above 0.8 represent good model discrimination.\(^{164}\)

**Adjusting institutional outcomes**

To facilitate comparison of outcomes between providers, case-mix adjustment was performed using multiple regression models generated when analysing risk factors
for mortality, as described in section 5.3.2. The SPSS® program can be instructed to record model predictions for the dependent variable for individual cases. For binary outcomes such as death, the model generates a probability of dying between 0 and 1 for each case within a particular dataset. At the provider level, these probabilities may be added together to create an expected number of deaths for the cohort of patients under study. The observed total number of deaths may also be calculated. The ratio of observed-to-expected (O/E) deaths may be derived, providing an indication of whether a particular provider had more or fewer deaths than expected by the regression model, given the specific case-mix of that institution’s patients.

This O/E ratio may be used to compare providers directly. Alternatively, an adjusted number of deaths may be calculated as follows:

\[
\text{Adjusted deaths} = \frac{\text{Observed}}{\text{Expected}} \times \text{Cohort death rate} \times \text{Provider caseload}
\]

**Comparing institutional outcomes**

Institutional performance for different patient cohorts was compared using two techniques. Firstly, case-mix adjusted institutional outcomes were displayed graphically on funnel plots, and units with extremes of outcome were identified to assess the consistency of relative institutional performance across cohorts. This visual method allows the reader to assess and assimilate a large amount of information quickly and easily. While all institutional data is displayed on a funnel plot, its design naturally directs attention towards a relatively small number of units at either end of the spectrum of performance, away from the majority of units whose performance is centrally distributed. Therefore, this technique was combined with a second approach. O/E mortality ratios for each provider’s individual subgroups
were assessed for normality using the Kolmogorov-Smirnov test and graphical examination of results, and appropriate correlation tests were then performed. Results of such analyses were reported with Pearson’s r or Spearman’s rho ($\rho$), and the associated $p$-value.

Funnel plots were created using the ‘Funnel plot for proportions’ template, available online from the Public Health Observatories website\textsuperscript{186}. This utilises a Poisson distribution to calculate second and third standard deviation control limits from the cohort mean event rate. This facilitates testing of the null hypothesis that all units have the same event rate, approximated by the cohort mean. With large, population-based samples, this approximation is likely to be a very close estimate of the true population mean. For each institution with a particular sample of patients, there is likely to be error in the associated estimate of that institution’s true outcome rate, the magnitude of which is related to the size of the sample. Assuming the null hypothesis is true, that the true institutional outcome rate is the same as the population outcome rate, 95% of institutional samples should fall within two standard deviation control limits of the cohort mean, and 99.7% should be within the third standard deviation control limit. Therefore, if a particular unit’s event rate is outside the second standard deviation control limit, the likelihood of such a result occurring by chance alone, provided the null hypothesis is true, is less than 5%, or $p<0.05$. If a unit’s results are outside the third standard deviation control limit, the probability that this has occurred by chance is $p<0.003$. In most studies within this thesis, individual units were considered performance outliers and the null hypothesis was rejected if their outcomes were outside third standard deviation control limits. This threshold for rejecting the null hypothesis is more stringent than the common convention of $p<0.05$. This was chosen for two reasons. It
allowed for some variation in outcomes due to uncontrolled case-mix. Within the studies in this thesis, it also resulted in a more manageable number of outliers for identification across cohorts, as there were often a large number of outliers outside second standard deviation control limits.

5.4 Chapter summary

This chapter has covered the source of HES data, what information the database contains, and how the data has been extracted, processed and analysed in this thesis. It has also briefly outlined its key strengths and weaknesses. Specific details of how HES data has been used for particular studies are reported in the relevant chapters.
Chapter 6

Study 2: Mortality after colorectal cancer surgery between 1998 and 2012

6.1 Chapter overview

This chapter examines trends in postoperative mortality after colorectal cancer surgery between 1998 and 2012. Simple, unadjusted analysis is presented to describe the changes in detail, and provide some understanding of the role of salient developments during this period: the introduction and widespread adoption of laparoscopic surgical techniques; and the national roll out of bowel cancer screening. Potential changes in complication management over time are also explored. This broad examination of HES data helps provide some insights into the factors that may be important in determining the outcomes of care for patients. It also helps to contextualise the subsequent, focused studies presented in this thesis.
6.2 Introduction

There have been many changes in colorectal practice over the last twenty years. Laparoscopic surgery is now widely practiced\textsuperscript{187–190}. The Enhanced Recovery After Surgery (ERAS) protocol, popularised by Kehlet\textsuperscript{191}, has increasingly become the standard of care\textsuperscript{192,193}. Besides these changes in direct patient care, the rise of the Multi-Disciplinary Team Meeting (MDTM) has changed the way clinical decisions are made for cancer patients\textsuperscript{194,195}. Bowel cancer screening has also become commonplace, with national programs in a number of countries\textsuperscript{196–199}. In addition, in England, the care pathway for patients with suspected or confirmed cancer has been subject to performance measurement since the introduction of wait time targets, announced in the NHS Cancer Plan in 2000\textsuperscript{200}.

This study set out to examine trends in postoperative mortality after colorectal surgery in detail, to try and better understand the role of specific developments in care over recent years. A fall in postoperative mortality has previously been documented by other researchers\textsuperscript{24}. This study used administrative data to examine the uptake of laparoscopic colorectal surgery, and the outcomes of patients who were or were not operated on using this approach. It also considered the outcomes of patients eligible to take part in the national Bowel Cancer Screening Program (BCSP) alongside those not directly affected by this program. Lastly, it sought to examine the results of patients who have or have not experienced a complication of surgery. Examination of these various groups will help establish a better understanding of the changes in colorectal surgery during the 14-year period under study.
Study questions

- How have recent falls in postoperative mortality after colorectal surgery affected different patient subgroups, defined by: sex; urgency of procedure; surgical access route; eligibility for screening; or the presence of a complication requiring surgical re-intervention?

- What roles have the introduction of laparoscopic surgery and bowel cancer screening for 60-69 year olds had in reducing postoperative mortality after colorectal cancer surgery?

- What role has the management of major postoperative complications that require surgical reintervention had in reducing postoperative mortality for these patients?

6.3 Methods

This study used HES data to determine the number of patients undergoing a colorectal resection for cancer in the NHS in England, as specified in the inclusion criteria below. The associated postoperative mortality rates were also derived. The study examined laparoscopy rates, patient ages, and surgical failure to rescue rates, to allow assessment of the results of subgroups according to these variables. An overview of the process of data extraction, processing and analysis is described in chapter 5. Study-specific methods are presented within this section.
6.3.1 Inclusion and exclusion criteria

For this study, records were extracted from HES using a combination of OPCS-4 and ICD-10 codes. Patients were included if they underwent a primary colorectal resection, as defined by the OPCS-4 codes in table 6.1, in association with a primary ICD-10 diagnosis of colorectal cancer (C18-20). Patients aged 18 years or more undergoing an eligible procedure between 1 April 1998 and 31 March 2012 were included. If a patient underwent more than one relevant procedure during the study period, only the first resection was included.

<table>
<thead>
<tr>
<th>Procedure name</th>
<th>OPCS-4 code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panproctocolectomy</td>
<td>H04</td>
</tr>
<tr>
<td>Total colectomy</td>
<td>H05</td>
</tr>
<tr>
<td>Extended right hemicolecotmy</td>
<td>H06</td>
</tr>
<tr>
<td>Right hemicolecotmy</td>
<td>H07</td>
</tr>
<tr>
<td>Transverse colectomy</td>
<td>H08</td>
</tr>
<tr>
<td>Left hemicolecotmy</td>
<td>H09</td>
</tr>
<tr>
<td>Sigmoid colectomy</td>
<td>H10</td>
</tr>
<tr>
<td>Other excision of colon</td>
<td>H11</td>
</tr>
<tr>
<td>Subtotal colectomy</td>
<td>H29</td>
</tr>
<tr>
<td>Excision of rectum</td>
<td>H33</td>
</tr>
</tbody>
</table>

Table 6.1: OPCS codes for colorectal procedures selected for inclusion.

6.3.2 Definition of variables and outcomes

A number of variables were derived to allow comparison of groups over time. The variables are listed below.

- **Year of procedure** was aligned to the financial year, from 1 April to 31 March.

- **Urgency of admission** was coded as elective or non-elective.
• **Sex.**

• **Use of laparoscopy** was derived from the following OPCS codes: Y75.1-4, Y75.8-9, Y76.3, Y76.5, Y76.8-9, Y50.8 and Y52.8. These codes include operations converted from laparoscopic to the open approach.

• **Age.**

• **Surgical failure to rescue** within 28 days of the index procedure, during the same admission, was identified using the OPCS-4 codes shown in table 6.2.

<table>
<thead>
<tr>
<th>Procedure group</th>
<th>OPCS-4 code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small bowel operation</td>
<td>G49, G51, G53.2, G58, G61, G63.4, G69, G71, G72, G73, G75, G78.4</td>
</tr>
<tr>
<td>Colorectal operation</td>
<td>H04, H05, H06, H07, H08, H09, H10, H11, H29, H13, H17, H33</td>
</tr>
<tr>
<td>Stoma procedure</td>
<td>G60, G74, G75, H14, H15</td>
</tr>
<tr>
<td>Operation for deep sepsis</td>
<td>T34, T45, T46.3, T46.8-9, Y22</td>
</tr>
<tr>
<td>Operation for superficial sepsis</td>
<td>S47.2, S47.4, S47.6, S47.8-9, T31.5</td>
</tr>
<tr>
<td>Operation for bleeding</td>
<td>J69, J70.1, J70.8-9, J72.2, J72.4, J72.8-9, T30.1</td>
</tr>
<tr>
<td>Operation for adhesions</td>
<td>T41.2-3, T41.5, T41.8-9, T42.3</td>
</tr>
<tr>
<td>Other procedure</td>
<td>T30.2-4, T30.8-9, T41.4, T42.4, Y29, Y31, Y32, Y50.2, Y70.1-2, Y75</td>
</tr>
</tbody>
</table>

Table 6.2: OPCS codes used to identify surgical failure to rescue procedures.

The outcome examined was 90-day postoperative mortality. This outcome was selected as it was considered that 90-day mortality may provide a better reflection of the outcomes of complication management than 30-day mortality, and complication management was one of the areas being examined in this study. The difference between 30- and 90-day mortality is examined in the next study in this thesis, presented in chapter 7.
6.3.3 Statistical analyses

This section explains how the analysis was performed, to assess trends in mortality over time for the various patient groups examined. Data was aggregated by financial year, split into subgroups according to urgency of admission and patient sex for all analyses. Further subgroups were examined according to surgical access route (laparoscopic or open), patient age, and whether or not patients had undergone a surgical failure to rescue procedure. Total numbers of cases and the number of deaths for each group were determined. From these numbers, unadjusted mortality rates were calculated. It was also possible to derive the proportions of patients undergoing laparoscopic surgery or a surgical failure to rescue procedure from this data. These results were tabulated (provided in appendix D), illustrated graphically, and assessed for changes over time using modelling techniques described below. For all tests, a $p$-value of <0.05 was considered statistically significant.

Laparoscopy

Relative to the total number of cases performed in a particular year, the proportion of procedures performed with an associated minimal access code was derived for elective and non-elective male and female patient groups. Annualised 90-day post-operative mortality rates were then determined by urgency of admission, operative approach (open or laparoscopic) and patient sex.

Screening

The national BCSP was rolled out in 2006, for members of the population aged 60-69. By 2012, patients aged 69 years when invited to participate in the first wave of
screening in 2006, would have been 75 years old. To allow comparison of separate screened and non-screened patient groups, it was decided to exclude patients aged 70 to 79 from this section of the analysis only (all patient age groups were included in all other analysis). Therefore, for screening, the outcomes of patients aged less than 60 years, 60-69 years, and 80 years or more were studied, across elective and non-elective groups.

Complication management

Patients experiencing a complication were identified using the OPCS-4 codes listed in table 6.2 in secondary procedure fields in HES, performed within 28 days of the index procedure, during the same admission. The rates of such procedures were determined on an annual basis, across elective and non-elective settings. Mortality results over time were compared among those who did or did not undergo such an operation.

Trend analysis

Year-on-year trends were examined using the ‘Curve estimation’ function in IBM SPSS. In each case, financial year was the independent variable, and the event rate (90-day mortality, laparoscopy or surgical failure to rescue) was the dependent variable. The optimal modelling of trends over time can be complex. While many researchers use linear or exponential modelling, these may not always be the best fitting model type\textsuperscript{201}. However, for the purposes of this study, modelling was performed to test the null hypothesis that there was no change over time. The pros and cons of specific model types, and the accuracy of fit that they provided, was a lesser concern. Nonetheless, a basic consideration of the key differences between
these model types was used to aid model selection. Linear models estimate that the absolute magnitude of change is the same every year. If the time period is extended sufficiently, a linear model could, for example, predict negative mortality rates, or laparoscopy rates over 100%. Exponential models assume that the relative magnitude of change is the same each year. Such models can therefore be asymptotic to zero. Exponential models are a convenient alternative to linear models, that do not predict impossible, negative mortality rates\textsuperscript{201}. Therefore, for this study, mortality rates were examined using this approach. Previous researchers have reported exponential adoption of new techniques once they have passed a certain ‘tipping point’\textsuperscript{202}. Therefore, laparoscopy rates were also modelled exponentially, as the study covered the period during which laparoscopic colorectal surgery became widely practised. The true trend in surgical failure to rescue rates over time was likely to be complex. This rate could not exceed 100%, and a continuous (linear) or accelerating (exponential) rise over time would not have been clinically desirable. For simplicity, a linear model was used to determine whether to reject the null hypothesis, that rates did not change over time, accepting this model type may not have provided the optimal representation of clinical practice.

6.4 Results

In the NHS in England between 1998 and 2012, 205 095 elective and 56 780 non-elective procedures were included in the study. By the 90\textsuperscript{th} postoperative day, 10 258 (5.0%) and 11 130 (19.6%) patients had died after elective and non-elective surgery, respectively. Annual numbers of elective and non-elective cases
and postoperative mortality rates are shown in table 6.3.

<table>
<thead>
<tr>
<th>Year</th>
<th>Elective Cases</th>
<th>Elective D90 (%)</th>
<th>Non-elective Cases</th>
<th>Non-elective D90 (%)</th>
<th>Total Cases</th>
<th>Total D90 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998-9</td>
<td>13888</td>
<td>954 6.9</td>
<td>4858</td>
<td>993 20.4</td>
<td>18746</td>
<td>1947 10.4</td>
</tr>
<tr>
<td>1999-0</td>
<td>14108</td>
<td>881 6.2</td>
<td>4886</td>
<td>1064 21.8</td>
<td>18994</td>
<td>1945 10.2</td>
</tr>
<tr>
<td>2000-1</td>
<td>13632</td>
<td>834 6.1</td>
<td>4535</td>
<td>960 21.2</td>
<td>18167</td>
<td>1794 9.9</td>
</tr>
<tr>
<td>2001-2</td>
<td>12740</td>
<td>815 6.4</td>
<td>4422</td>
<td>894 20.2</td>
<td>17162</td>
<td>1709 10.0</td>
</tr>
<tr>
<td>2002-3</td>
<td>13387</td>
<td>789 5.9</td>
<td>4322</td>
<td>897 20.8</td>
<td>17709</td>
<td>1686 9.5</td>
</tr>
<tr>
<td>2003-4</td>
<td>13190</td>
<td>716 5.4</td>
<td>4226</td>
<td>868 20.5</td>
<td>17416</td>
<td>1584 9.1</td>
</tr>
<tr>
<td>2004-5</td>
<td>13364</td>
<td>764 5.7</td>
<td>4321</td>
<td>902 20.9</td>
<td>17685</td>
<td>1666 9.4</td>
</tr>
<tr>
<td>2005-6</td>
<td>14638</td>
<td>741 5.1</td>
<td>4143</td>
<td>818 19.7</td>
<td>18781</td>
<td>1559 8.3</td>
</tr>
<tr>
<td>2006-7</td>
<td>14705</td>
<td>712 4.8</td>
<td>3823</td>
<td>757 19.8</td>
<td>18528</td>
<td>1469 7.9</td>
</tr>
<tr>
<td>2007-8</td>
<td>15358</td>
<td>710 4.6</td>
<td>3617</td>
<td>671 18.6</td>
<td>18975</td>
<td>1381 7.3</td>
</tr>
<tr>
<td>2008-9</td>
<td>16199</td>
<td>675 4.2</td>
<td>3447</td>
<td>639 18.5</td>
<td>19646</td>
<td>1314 6.7</td>
</tr>
<tr>
<td>2009-10</td>
<td>16297</td>
<td>642 3.9</td>
<td>3473</td>
<td>641 18.5</td>
<td>19770</td>
<td>1283 6.5</td>
</tr>
<tr>
<td>2010-1</td>
<td>16979</td>
<td>567 3.3</td>
<td>3518</td>
<td>555 15.8</td>
<td>20497</td>
<td>1122 5.5</td>
</tr>
<tr>
<td>2011-2</td>
<td>16610</td>
<td>458 2.8</td>
<td>3189</td>
<td>471 14.8</td>
<td>19799</td>
<td>929 4.7</td>
</tr>
<tr>
<td>Totals</td>
<td>205095</td>
<td>10258 5.0</td>
<td>56780</td>
<td>11130 19.6</td>
<td>261875</td>
<td>21388 8.2</td>
</tr>
</tbody>
</table>

Table 6.3: Number of resections and 90-day deaths by year and urgency of admission.
D90 - deaths within the first 90 postoperative days.

6.4.1 Laparoscopy

In this section, laparoscopy rates over time will be presented and examined first, before reviewing the changes in mortality after laparoscopic and open surgery.

Laparoscopy rates

The proportion of laparoscopic procedures performed each year is shown in figure 6.1. The associated annual counts of operations (and deaths) are provided in table D.1 in appendix D. Laparoscopy rates rose significantly among all groups examined (model results shown in table 6.4).
Figure 6.1: Rates of laparoscopic surgery against year, by urgency of admission and sex of patient.

<table>
<thead>
<tr>
<th>Urgency of admission</th>
<th>Patient sex</th>
<th>R-squared</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elective</td>
<td>Male</td>
<td>0.951</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.954</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Non-elective</td>
<td>Male</td>
<td>0.935</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.772</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 6.4: Exponential model results for laparoscopy rates over time.
Mortality rates according to surgical approach

Postoperative 90-day mortality rates after open and laparoscopic surgery are shown in figure 6.2, with elective and non-elective cohorts displayed separately. Regardless of the urgency of the procedure, mortality after open surgery fell over time, although the fall was less marked in the non-elective cohort (see table 6.5 for model results). Death rates after laparoscopic surgery during the first half of the study period showed wide fluctuation before stabilising and demonstrating a steady downward trend. This stabilisation corresponds temporally to rising numbers of cases undergoing laparoscopic surgery, with over 500 elective laparoscopic procedures in male and female groups from 2005-6. Lower volumes of non-elective surgery were found, with annual laparoscopic cases rising over 20 for the whole of England for the first time in 2004-5. Considered from these respective time points, mortality after laparoscopic surgery fell significantly for all except non-elective male patients.

<table>
<thead>
<tr>
<th>Urgency of admission</th>
<th>Operative approach</th>
<th>Patient sex</th>
<th>R-squared</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elective</td>
<td>Open</td>
<td>Male</td>
<td>0.886</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.850</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Laparoscopic *</td>
<td>Male</td>
<td>0.948</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.923</td>
<td>0.001</td>
</tr>
<tr>
<td>Non-elective</td>
<td>Open</td>
<td>Male</td>
<td>0.647</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.614</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Laparoscopic †</td>
<td>Male</td>
<td>0.349</td>
<td>0.123</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.646</td>
<td>0.016</td>
</tr>
</tbody>
</table>

Table 6.5: Exponential model results for postoperative mortality rates over time, by urgency of admission, surgical approach and sex of patient. * - mortality rates examined from 2005; † - mortality rates examined from 2004.
Figure 6.2: 90-day mortality rates against year, by surgical approach and sex of patient. Elective and non-elective results shown separately. Lap - laparoscopic approach.
6.4.2 Screening

90-day mortality rates for patients in the age groups examined are shown in figure 6.3. Annual count data are provided in tables D.2 and D.3 in appendix D. Among patients aged 60-69, who were eligible for screening once it was introduced nationally in 2006, postoperative death rates fell significantly after both elective and non-elective surgery (see table 6.6 for model results). Among the oldest patients, aged 80 years and over, mortality also fell significantly irrespective of the urgency of their colorectal operation. Postoperative mortality fell significantly for patients under 60 years of age undergoing colorectal surgery on an elective basis, but there was no significant fall for these patients after non-elective surgery.

The decline in postoperative mortality among 60-69 year old patients since the introduction of screening in 2006 occurred within two broader contexts. Firstly, mortality among this patient group was already falling before screening was rolled out. Secondly, there have been marked reductions in mortality among non-screening age patient groups.

6.4.3 Complication management

This section first examines the rate of surgical re-intervention for a complication, as defined in the methods section, before assessing the temporal trend in mortality according to whether or not such a procedure was performed.

Surgical failure to rescue rates

The annual proportion of patients undergoing a surgical failure to rescue procedure during the index admission is shown in figure 6.4. Annual count data are provided
Figure 6.3: 90-day mortality rates against year, by age group and sex of patient. Elective and non-elective results shown separately. Ages given in years.
<table>
<thead>
<tr>
<th>Urgency of admission</th>
<th>Age group (years)</th>
<th>Patient sex</th>
<th>R-squared</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elective</td>
<td>&lt;60</td>
<td>Male</td>
<td>0.613</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.745</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>60-69</td>
<td>Male</td>
<td>0.885</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.781</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>80+</td>
<td>Male</td>
<td>0.736</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.911</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Non-elective</td>
<td>&lt;60</td>
<td>Male</td>
<td>0.210</td>
<td>0.099</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.080</td>
<td>0.328</td>
</tr>
<tr>
<td></td>
<td>60-69</td>
<td>Male</td>
<td>0.391</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.519</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>80+</td>
<td>Male</td>
<td>0.690</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.429</td>
<td>0.011</td>
</tr>
</tbody>
</table>

Table 6.6: Exponential model results for postoperative mortality rates over time, by urgency of admission, age group and patient sex.

in table D.4 in appendix D. The rate of reintervention rose significantly for all groups except non-elective males (see table 6.7). Rates were consistently lower in women than in men, regardless of the urgency of the index procedure.

<table>
<thead>
<tr>
<th>Urgency of admission</th>
<th>Patient sex</th>
<th>R-squared</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elective</td>
<td>Male</td>
<td>0.635</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.751</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Non-elective</td>
<td>Male</td>
<td>0.087</td>
<td>0.305</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>0.348</td>
<td>0.026</td>
</tr>
</tbody>
</table>

Table 6.7: Linear model results for surgical failure to rescue rates over time.

**Mortality rates according to reintervention or not**

The 90-day postoperative mortality rate for patients that did or did not undergo a surgical failure-to-rescue procedure is shown in figure 6.5, across both elective and non-elective groups. Mortality fell significantly for all patients that did not undergo a failure to rescue reintervention (see table 6.8). The death rate after surgical
Figure 6.4: Rates of surgical failure to rescue procedures against year, by urgency of admission and sex of patient. FTR - failure to rescue.
reintervention also fell significantly for elective colorectal resections regardless of patient sex. The mortality rate fell for male patients undergoing reintervention after a non-elective colorectal resection, but there was no fall for non-elective female patients.

<table>
<thead>
<tr>
<th>Urgency of admission</th>
<th>Surgical FTR</th>
<th>Patient sex</th>
<th>R-squared</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elective</td>
<td>No reintervention</td>
<td>Male</td>
<td>0.883</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.892</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Reintervention</td>
<td>Male</td>
<td>0.775</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.588</td>
<td>0.001</td>
</tr>
<tr>
<td>Non-elective</td>
<td>No reintervention</td>
<td>Male</td>
<td>0.726</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.700</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Reintervention</td>
<td>Male</td>
<td>0.349</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female</td>
<td>0.138</td>
<td>0.191</td>
</tr>
</tbody>
</table>

Table 6.8: Exponential model results for postoperative mortality rates over time, by urgency of admission, surgical failure to rescue and sex of patient. FTR - failure to rescue.

### 6.5 Discussion

This study has examined mortality rates after colorectal surgery in England between 1998 and 2012. During that time, the rate of laparoscopic surgery rose markedly to approximately 50% for elective patients, with a lesser rise to about 10% for non-elective patients. Postoperative mortality fell for all but one patient group, irrespective of the surgical approach, patient sex or urgency of procedure. Bowel cancer screening was rolled out nationally in 2006 for patients aged 60-69 years. While mortality fell for this age group after 2006, the results show that mortality also reduced for screening-age patients before the introduction of screening, between 1998 and 2006. In addition, there were significant falls in mortality for
Figure 6.5: 90-day mortality rates against year, by surgical failure to rescue and sex of patient. Elective and non-elective results shown separately.
patients unaffected by screening, aged 80 years and over or below 60 years of age. The rate of surgical reintervention for a complication (surgical failure to rescue) rose for 3 of 4 patient groups examined. The attendant mortality rates fell for all but one group, regardless of whether or not a patient underwent a reintervention for a major complication, their sex or urgency of admission.

This detailed exploration of the changes in postoperative mortality may shed some light on the potential role of certain important developments in colorectal surgery over recent years. The decline in mortality was almost universal, across the different groups examined by urgency, patient sex, operative approach, screening eligibility and major complication. This observation suggests that improvements in care have benefited nearly all patients undergoing colorectal surgery for cancer, regardless of laparoscopic techniques, bowel cancer screening, or improvements in complication management. Developments in each of these areas may have made an important contribution to an overall improvement in care, but none appear to have had a dominant effect.

The limits of HES data have been described earlier in this thesis (see section 5.2.3). Improvements in the accuracy of coding over time are particularly pertinent to this study. This may have had a significant, but difficult to quantify, effect on the results presented. It is likely that some proportion of the rise in surgical failure to rescue rates presented is a direct reflection of improved coding, rather than a true increase in reintervention rates. However, the reported rise in laparoscopy rates appears to be an accurate reflection of clinical practice, closely matching laparoscopy rates reported in contemporary National Bowel Cancer Audit Program reports\textsuperscript{187,194,203}. This suggests that recording of this particular information was accurate within the HES database. For the analysis presented in this chapter,
the accuracy of coding was likely to be adequate to assess broad trends in care and its outcomes. Systematic changes in coding would have had to affect one subgroup more than another to have influenced trends differentially between the subgroups studied, and this is considered unlikely. This study is also limited by its specific examination of the results of patients who underwent surgical treatment for cancer. Longitudinal changes in the stage of disease and the selection of individual patients to undergo an operation may have influenced the results achieved. It is of note that overall operative volume did increase during the study period, as shown in table 6.3, with a marked increase in the annual number of elective operations performed. There was however a large drop in the number of non-elective procedures performed. These changes warrant further exploration with more detailed information on cancer stage.

This study fits within a broader literature documenting improvements in outcomes across a range of cancer types. International epidemiologic studies of bowel and other cancers have reported falling mortality rates over time\cite{23,204,205}. Research specifically examining the outcomes of colorectal surgery has shown a steady improvement in outcomes across a number of countries\cite{24,206–208}.

While this observational study cannot specifically determine why there has been such a global improvement in the outcomes of colorectal cancer surgery during the period examined, possible explanations may be proposed. As mentioned above in the limitations section, case selection may have changed during the study period. Technical developments in cross-sectional imaging may have improved the accuracy of staging investigations that are usually performed before a patient undergoes surgery. This may have enabled clinicians to better select the most appropriate patients for surgery. For example, better staging information may
help surgeons avoid operating on patients with small but widespread metastatic disease, that may not have been detected in the past, and were likely to have poor short-term outcomes. While surgeons may have become better at selecting cases appropriately, in oncological terms, the results presented suggest that surgeons have not simply become more reluctant to operate on patients that are high risk due to their age. Table D.3 shows a marked rise in the number of patients aged 80 years and over undergoing elective surgery, while the associated mortality rates fell. Fully unravelling the complexity of case selection would require greater information on clinical variables that are not present in HES. However, the general rise in volume over subsequent years suggests that any changes have not resulted in a dramatic fall in the proportion of patients being offered surgery.

The near-universal nature of the fall in mortality suggests that broad quality of care factors may have been important. Better detection and management of comorbidities, such as ischaemic heart disease and diabetes, may have improved the ability of patients affected by these conditions to survive surgery, regardless of its urgency, the surgical approach or the presence of a complication. Improvements in diagnostic and therapeutic imaging and endoscopic services may have contributed to better outcomes for all patients. Developments in intensive care may similarly have benefited patients broadly. Detailed assessment of these potential factors is challenging. For example, technical developments in imaging and endoscopy have often been incremental, representing evolution over a diffuse time period, rather than a clear transition from one period to another. Assessment of the impact of comorbidity over time may be confounded by contemporaneous improvements in coding. It is likely that multiple improvements in the many different areas of clinical practice upon which patients undergoing surgical treat-
ment for cancer depend - surgical, medical, anaesthetic, radiological, endoscopic, oncological - have combined to produce the documented fall in mortality.

The results presented suggest that greater advances appear to have been made in elective care than in non-elective surgery. 90-day mortality after elective surgery fell significantly for all subgroups examined. In comparison, the death rate after non-elective surgery did not fall significantly for 4 of 14 subgroups: laparoscopically treated males; male and female patients under 60 years of age; and female patients undergoing a surgical failure to rescue reintervention. Two of these findings may have been significantly influenced by small sample sizes. The annual number of non-elective males undergoing laparoscopic surgery only rose above 100 from 2009-10, and the number of non-elective female patients undergoing a failure to rescue procedure averaged 75 across the study period. In addition, the use of laparoscopic surgery in the non-elective setting is still very limited, and likely heavily influenced by case selection biases. Annual volumes were higher for non-elective patients aged under 60 years of age, so the finding that postoperative mortality for this group did not fall may be more reliable. It has been reported that younger patients with cancer tend to present with more advanced disease\textsuperscript{211,212}. When faced with surgery on a non-elective basis, these patients may represent a particularly challenging group where improvements in outcomes are very difficult to achieve. More detailed study with clinical data is warranted.

Overall, the fact that postoperative mortality rates fell for nearly every subgroup examined, regardless of surgical approach, eligibility for screening, or reintervention for a complication, suggests that quality of care for patients undergoing surgery for colorectal cancer has improved broadly. This background provides important information for future quality improvement. Attention should be paid
to the full range of factors that influence patient care, rather than focusing on single technical innovations, such as laparoscopic surgery. The accrual of small improvements across many areas may perhaps yield the greatest benefit for patients. These findings should be borne in mind when assessing clinical care, and when considering the nature of interventions to improve practice in the future.

6.6 Chapter summary

The study presented in this chapter has examined broad trends in the postoperative mortality after colorectal cancer surgery between 1998 and 2012. It found that mortality rates fell for nearly all patient groups examined, regardless of the urgency of the procedure (elective or non-elective), the mode of operation (laparoscopic or open), whether patients were eligible for screening or not, or whether a patient experienced a complication requiring surgical reintervention or not. These findings may be interpreted to suggest that broad improvements in care quality, across a range of clinical domains relevant to all patients undergoing surgery, may have been key drivers in improving surgical outcomes. Specific technical innovations, such as the advent of laparoscopic surgery, may perhaps have played a lesser role. High, or improving, performance may therefore depend upon the effective coordination of all aspects of care, rather than a narrow focus on a single, or limited number, of areas of surgical practice.
Chapter 7

Study 3: Population-based cohort study comparing 30- and 90-day institutional mortality rates after colorectal surgery

7.1 Chapter overview

This chapter presents the methods and results of a study using HES data to examine differences between department-level 30- and 90-day mortality after colorectal surgery in English NHS Trusts. It is important to understand how variation in the definition of a particular outcome may influence the assessment of unit-level performance.

7.2 Introduction

Thirty-day mortality has conventionally been used to reflect the outcome of surgical care. Published 30-day and 1-year mortality rates after colorectal surgery range from 3.0% to 4.9%, and 8.8% to 12.4% respectively\textsuperscript{170–172}. High mortality
beyond 30 days highlights the importance of considering alternative periods for defining institutional performance and outcome reporting.

It is unknown whether lengthening follow-up for analysis of postoperative deaths to 90 days results in identification of a different group of units with outlying results compared with 30-day mortality. The relationship between high or low mortality rates at 90 days and death rates at 180 and 365 days has also not been studied in the literature. This study examines the relationship between mortality metrics, varied by the time period included, during the first year after colorectal surgery between 2001 and 2007, using English HES data.

This study is important in helping understand how the specifics of performance measurement may affect our understanding and interpretation of unit-level results. As has been previously discussed in section 2.2 of chapter 2, the apparent objectivity of outcomes belies the fact that the definition of specific outcome measures, such as postoperative mortality, may be subject to conscious or unconscious biases, when more than one definition is available. By examining the difference between 30- and 90-day mortality, we can better understand the impact of using different definitions on our assessment of performance.

Study question

- Is there any difference between 30- and 90-day mortality in the assessment of relative unit-level performance after colorectal surgery?
7.3 Methods

The methods used in this study are described in chapter 5. In summary, an extract of HES data was defined according to the inclusion criteria below, before processing and ‘cleaning’ steps to remove duplicate entries, invalid variables and identify the first eligible resection for individual patients included. The HES extract was ONS linked, providing the date of death of patients where applicable.

7.3.1 Inclusion and exclusion criteria

In this study, patients aged 18 years or more undergoing elective or emergency colorectal resection between April 2001 and February 2007 were selected for inclusion. The specific OPCS-4 codes used to define the extract are provided in table 7.1. Patients treated in NHS Trusts that only performed colorectal surgery on a selected patient population (for example, children or women only), or less than 10 times during the study period were excluded from the study.

7.3.2 Definition of variables and outcomes

For the comparison of cohorts and case-mix adjustment of outcomes, the following variables were derived and grouped:

- **Age** in years was obtained from the ‘startage’ data field, coded into four groups: <55, 55-69, 70-79, and ≥80.

- **Charlson Comorbidity Index** was grouped as follows: 0, 1-2, 3-5, and ≥6.

- **Sex**.
• **Primary colorectal diagnosis** was determined using the first diagnosis field in the HES database. ICD-10 codes were inspected and classified as benign or malignant.

• **Urgency of operation** was coded as elective or non-elective.

• **Operative procedure** was grouped by anatomical site as shown in table 7.1.

• **Use of laparoscopy** was coded as open or laparoscopic, derived using OPCS-4 codes Y058, Y752 and Y714. These include operations converted from laparoscopic to open.

• **Year of procedure** was coded according to the calendar year of the operation.

Outcome was defined as all-cause mortality at various time-points after surgery:
30, 90, 180 and 365 days.

7.3.3 Statistical analysis

The characteristics of the patients dying within each time period examined were assessed for similarity using the $\chi^2$ test. Multiple logistic regression was performed to identify risk-factors for death and calculate risk-adjusted mortality rates which were displayed on funnel plots. O/E mortality rates for different follow-up periods were assessed for correlation using the appropriate parametric or non-parametric correlation coefficient, according to the distribution of the data. For all tests, a $p$-value of $<0.05$ was considered statistically significant. For further details on how the analysis was performed, please refer to section 5.3.2 in chapter 5.

7.4 Results

7.4.1 Cohort characteristics

Between 1st April 2001 and 28th February 2007, 171 688 patients underwent a primary colorectal resection in 153 NHS Trusts. 112 557 (65.6%) operations were performed on an elective basis, and 107 780 (62.8%) for malignant disease. 5 566 (3.2%) procedures were performed laparoscopically. The cohort characteristics are presented in table 7.2.

7.4.2 Mortality rates

Of 171 688 patients, 14 537 (8.5%) died within 30 days of surgery. At 90 days, 19 466 (11.3%) had died, increasing to 23 942 (13.9%) and 31 782 (18.5%) at 180 and
30- vs 90-day mortality
172

Age (years)

Sex
CCI

Admission
Diagnosis
Resection

Approach
Year

<55
55-69
70-79
≥80
Male
Female
0
1-2
3-5
≥6
Elective
Non-elective
Malignant
Benign
Total/subtotal
Right
Left
Rectal
Open
Laparoscopic
2001
2002
2003
2004
2005
2006
2007

All patients (%)
(n=171 688)
31 921 18.6
54 880 32.0
53 305 31.0
31 582 18.4
86 971 50.7
84 717 49.3
51 854 30.2
77 048 44.9
14 908 8.7
27 878 16.2
112 557 65.6
59 131 34.4
107 780 62.8
63 908 37.2
8 303 4.8
59 084 34.4
31 164 18.2
73 137 42.6
166 122 96.8
5 566 3.2
21 558 12.6
28 539 16.6
28 485 16.6
28 650 16.7
29 442 17.1
30 081 17.5
4 933 2.9

0-30
(n=14 537)
604
2 649
5 294
5 990
6 972
7 565
4 222
5 731
2 000
2 584
4 274
10 263
7 408
7 129
925
5 254
2 379
5 979
14 380
157
906
537
459
428
392
423
392
1
2
2
2
2
2

(%)
4.2
18.2
36.4
41.2
48.0
52.0
29.0
39.4
13.8
17.8
29.4
70.6
51.0
49.0
6.4
36.1
16.4
41.1
98.9
1.1
13.1
17.5
16.9
16.7
16.5
16.7
2.7

Mortality interval (days)
0-90 (%)
0-180
(n=19 466)
(n=23 942)
937 4.8
1 293
3 805 19.5
4 960
7 126 36.6
8 871
7 598 39.0
8 818
9 508 48.8
11 750
9 958 51.2
12 192
5 152 26.5
5 719
7 546 38.8
9 198
2 549 13.1
2 959
4 219 21.7
6 066
6 351 32.6
8 594
13 115 67.4
15 348
10 869 55.8
14 476
8 597 44.2
9 466
1 141 5.9
1 282
7 320 37.6
9 367
3 128 16.1
3 757
7 877 40.5
9 536
19 224 98.8
23 633
242 1.2
309
604 13.4
3 197
387 17.4
4 204
264 16.8
3 992
229 16.6
4 005
230 16.6
3 953
228 16.6
3 948
524 2.7
643
2
3
3
3
3
3

(%)
5.4
20.7
37.1
36.8
49.1
50.9
23.9
38.4
12.4
25.3
35.9
64.1
60.5
39.5
5.4
39.1
15.7
39.8
98.7
1.3
13.4
17.6
16.7
16.7
16.5
16.5
2.7

0-360
(n=31 782)
1 951
7 202
11 700
10 929
15 705
16 077
6 436
12 366
3 666
9 314
13 123
18 659
21 240
10 542
1 480
12 815
4 889
12 598
31 341
441
309
605
351
331
204
153
829

4
5
5
5
5
5

(%)

6.1
22.7
36.8
34.4
49.4
50.6
20.3
38.9
11.5
29.3
41.3
58.7
66.8
33.2
4.7
40.3
15.4
39.6
98.6
1.4
13.6
17.6
16.8
16.8
16.4
16.2
2.6

Table 7.2: Characteristics of entire patient cohort, and patients dying at specified post-operative intervals.
CCI - Charlson Comorbidity Index; p-value calculated using χ2 test

p

<0.001

0.034

<0.001

<0.001

<0.001

<0.001

0.052

0.998


365 days, respectively.

### 7.4.3 Factors associated with mortality

On single variable analysis all covariates were significantly correlated with mortality, therefore all variables were entered into multiple regression models for each postoperative time period. Results are shown in table 7.3. Across periods, increased risk of death was statistically significantly associated with increasing age, increasing comorbidity score, non-elective surgery, and having a benign diagnosis. A lower risk of mortality was significantly associated with female gender, operations other than total/subtotal colectomy, laparoscopic surgery, and later year of operation. Model fit was satisfactory, with c-statistics ranging from 0.758 to 0.809.

### 7.4.4 Relationship between mortality time periods

Funnel plots of risk-adjusted mortality rate against departmental caseload for individual NHS Trusts were created for mortality for 0-30, -90, -180, -365 day time periods (figures 7.1 to 7.4). A High or Low Mortality Unit (HMU / LMU) was defined as any unit with risk-adjusted 90-day mortality rates above or below third standard deviation control limits, respectively. HMU / LMUs were depicted using different markers to facilitate identification across time periods. 8 HMUs and 12 LMUs were identified (see figure 7.2).

At 30 days postoperatively, four units had mortality rates above the third SD (i.e. 99.7%) control limit (see figure 7.1), and all of these were HMUs at 90 days. Of the remaining four HMUs, three were close to the 95% control limit (two above,
<table>
<thead>
<tr>
<th>Age (years)</th>
<th>0-30 days</th>
<th>0-90 days</th>
<th>0-180 days</th>
<th>0-365 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 55</td>
<td>Ref</td>
<td>Ref</td>
<td>Ref</td>
<td>Ref</td>
</tr>
<tr>
<td>55-69</td>
<td>3.22</td>
<td>2.94</td>
<td>3.53</td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td>5.07</td>
<td>4.62</td>
<td>5.53</td>
<td></td>
</tr>
<tr>
<td>70-79</td>
<td>7.04</td>
<td>6.45</td>
<td>7.69</td>
<td></td>
</tr>
<tr>
<td>80-89</td>
<td>14.15</td>
<td>12.94</td>
<td>15.46</td>
<td></td>
</tr>
<tr>
<td>≥ 90</td>
<td>14.15</td>
<td>12.94</td>
<td>15.46</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sex</th>
<th>Male</th>
<th>Ref</th>
<th>Ref</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>0.87</td>
<td>0.84</td>
<td>0.91</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CCI</th>
<th>0</th>
<th>Ref</th>
<th>Ref</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2</td>
<td>2.39</td>
<td>2.26</td>
<td>2.53</td>
<td></td>
</tr>
<tr>
<td>3-4</td>
<td>5.32</td>
<td>4.92</td>
<td>5.76</td>
<td></td>
</tr>
<tr>
<td>≥ 6</td>
<td>3.83</td>
<td>3.55</td>
<td>4.14</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Admission</th>
<th>Elective</th>
<th>Ref</th>
<th>Ref</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-elective</td>
<td>4.45</td>
<td>4.27</td>
<td>4.63</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Malignant</th>
<th>Ref</th>
<th>Ref</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benign</td>
<td>3.70</td>
<td>3.49</td>
<td>3.92</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Resection</th>
<th>Total/subtotal</th>
<th>Ref</th>
<th>Ref</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right</td>
<td>0.58</td>
<td>0.54</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>Left</td>
<td>0.55</td>
<td>0.50</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Rectal</td>
<td>0.64</td>
<td>0.59</td>
<td>0.70</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Approach</th>
<th>Open</th>
<th>Ref</th>
<th>Ref</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laparoscopic</td>
<td>0.55</td>
<td>0.47</td>
<td>0.65</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>2001</th>
<th>Ref</th>
<th>Ref</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>1.01</td>
<td>0.94</td>
<td>1.08</td>
<td>0.781</td>
</tr>
<tr>
<td>2003</td>
<td>0.98</td>
<td>0.92</td>
<td>1.05</td>
<td>0.617</td>
</tr>
<tr>
<td>2004</td>
<td>0.94</td>
<td>0.88</td>
<td>1.01</td>
<td>0.068</td>
</tr>
<tr>
<td>2005</td>
<td>0.90</td>
<td>0.84</td>
<td>0.97</td>
<td>0.003</td>
</tr>
<tr>
<td>2006</td>
<td>0.91</td>
<td>0.85</td>
<td>0.98</td>
<td>0.008</td>
</tr>
</tbody>
</table>

Table 7.3: Multiple logistic regression results for mortality at each specified post-operative time interval.
Figure 7.1: Funnel plot of 30-day mortality against caseload.
90-day outliers marked: HMU - high mortality unit; LMU - low mortality unit.
Figure 7.2: Funnel plot of 90-day mortality against caseload. 90-day outliers marked: HMU - high mortality unit; LMU - low mortality unit.
Figure 7.3: Funnel plot of 180-day mortality against caseload.
90-day outliers marked: HMU - high mortality unit; LMU - low mortality unit.
Figure 7.4: Funnel plot of 365-day mortality against caseload. 90-day outliers marked: HMU - high mortality unit; LMU - low mortality unit.
one below) and one unit was within control limits, just below average for mortality. Of the 12 LMUs identifiable at 90 days, six were below 99.7% control limits for 30-day mortality. Of the remaining six LMUs, all had mortality below the 95% control limits at 30 days. Two further units had mortality results lower than three SD from the mean at 30 days but were not identified as LMUs at 90 days.

Funnel plots for intermediate mortality at 180 and 365 days were examined (figures 7.3 and 7.4). Increased variation in departmental results was seen over time, with increasing numbers of outliers beyond 95% and 99.7% control limits noted (see table 7.4). At 180 days, all 90-day HMUs were above third SD control limits, with no other units identified with mortality above this threshold (figure 7.3). Eleven 90-day LMUs were more than three SD below average for mortality. The 12th LMU had below average mortality, within two SD of the national average. A further four units not identified as 90-day LMUs had mortality rates below 99.7% control limits. At 365 days, seven of eight 90-day HMUs had mortality above 99.7% control limits, while the eighth HMU’s mortality was between 95% and 99.7% control limits (figure 7.4). Eight of 12 LMUs had mortality rates below 99.7% control limits. Of the remaining four LMUs, three were more than two SD below average for mortality, and one was within 95% control limits.

O/E mortality rates for each of the postoperative time periods were examined for normality. The Kolmogorov-Smirnov test suggested that 30-, 90- and 180-day mortality ratios were normally distributed (lower limit of significance reported as $p=0.200$ for each), but 365-day mortality was not ($p=0.010$). Graphical inspection of O/E mortality rates showed an approximately normal distribution, but with one high outlying unit at 90 days that became a higher outlier over subsequent periods. After exclusion of this unit, repeat Kolmogorov-Smirnov testing suggested O/E
mortality rates for all periods were parametrically distributed. Non-parametric Spearman correlation coefficients for mortality periods of all units, and parametric Pearson’s correlation after exclusion of the outlying unit were very similar. Therefore, it was decided to include all units in the assessment of correlation between time periods and perform a non-parametric statistical test. Spearman’s correlation coefficients are reported in table 7.5. Overall there was good correlation for mortality rates across time periods. 0 to 90- and 0 to 180-day O/E mortality ratios had the strongest overall correlation with other follow-up periods.

<table>
<thead>
<tr>
<th>Postoperative period (days)</th>
<th>0-90</th>
<th>p</th>
<th>0-180</th>
<th>p</th>
<th>0-365</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-30</td>
<td>0.924</td>
<td>&lt;0.001</td>
<td>0.853</td>
<td>&lt;0.001</td>
<td>0.735</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>0-90</td>
<td></td>
<td></td>
<td>0.957</td>
<td>&lt;0.001</td>
<td>0.860</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>0-180</td>
<td></td>
<td></td>
<td></td>
<td>0.933</td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.5: Spearman’s correlation coefficient for observed-to-expected mortality across different postoperative time periods for all units. 
p - 2-tailed significance test results.
7.5 Discussion

The present results confirm previous findings of a significant burden of mortality beyond 30 days after colorectal surgery\textsuperscript{170–172}. Mortality after 365 days was 18.5\%, more than twice the 8.5\% death rate at 30 days. In addition, extending follow-up from 30 to 90 days resulted in double the number of units identified as high mortality outliers, including all 30-day high outliers. High outlying mortality at 90 days was associated with high mortality rates at 180 and 365 days postoperatively. Low outlying mortality at 90 days was also associated with low outlier status across time periods.

The findings of this study must be considered in light of the known strengths and weaknesses of HES data, as described in section 5.2.3. There are also specific limitations relevant to this study. Elective and non-elective patients may represent different populations which should arguably be studied separately. However, the regression model included an adjustment for urgency of admission, and this technique has been used in other similar studies of postoperative outcomes\textsuperscript{22,24,57}. The study examined all-cause mortality, and follow-up over longer periods may have resulted in inclusion of increasing numbers of deaths due to unrelated factors. While the lack of data on stage of disease for cancer patients may not be a particular concern during short-term follow-up, this unmeasured confounder may have had a significant influence on outcomes towards the end of the first year after surgery. There have been changes in routine colorectal practice since 2007, the end of the period examined in this study. These include increased use of laparoscopy\textsuperscript{187,213} and the increasing implementation of Enhanced Recovery Programs\textsuperscript{192,193}. Some have shown no association between these changes and death.
rates\textsuperscript{102,112,113,214}, while others have reported an association between laparoscopy and reduced mortality\textsuperscript{170,215}. However, to influence the central findings of this study, these changes would need to have had a differential effect on death between 30 and 90 days after surgery.

The mortality rates presented are high, with wide variation and a steady reduction over time. Non-elective colorectal surgery is known to be associated with a higher death rate than elective treatment\textsuperscript{62,172,216}. This study included a relatively large proportion of patients treated on a non-elective basis that significantly influenced the average mortality rates, differentiating this study from others reporting 30-day and 1-year mortality\textsuperscript{170–172}. Wide variation in outcomes has been discussed previously in chapters 1 and 2. The range of institutional performance documented in the present study reinforces the findings of other studies of colorectal outcomes\textsuperscript{20,22,24,57}. Falling mortality rates after colorectal cancer surgery have been examined in some detail in chapter 6. The present study included a more diverse range of patients, covering benign and malignant disease of the colon and rectum. There may be important differences in patients with these different diagnoses that could have a differential effect on their outcomes over time, and this could be subject to further study. For a discussion of the broad changes that may have influenced postoperative mortality after colorectal surgery, please refer to the discussion section of chapter 6.

A greater number of mortality outliers were identified 90 days after surgery than at 30 days. The underlying reasons are not clear. Previous research has suggested that 90-day mortality captures more deaths than in-hospital mortality\textsuperscript{217}. Therefore, post-discharge deaths, with or without readmission, are likely to contribute to 90-day death rates. This does not, however, explain changes
in relative performance. The development of HMU status after 30 postoperative days may relate to a number of factors, including unmeasured case-mix variables, complication management, and follow-up practices. Unmeasured factors, such as socio-economic status, may influence short-term outcomes in uncertain ways, perhaps altering help-seeking behaviour after discharge, for example. It is known that complications, such as deep vein thrombosis, may present after 30 days\textsuperscript{218}. We also know that a significant number of readmissions after colorectal surgery occur after this period\textsuperscript{219}. Severe complications that present late (or take a long period to reach a final resolution) may only be captured when the time period is extended to 90 days. Intensive post-discharge follow-up with ready access to specialist care may affect outcomes by modifying the course of complications when they occur. 90-day mortality may provide some reflection of the ability of a hospital to appropriately detect and manage complications. It is important to note that the management of severe postoperative complications is often multidisciplinary, requiring input from a range of hospital services, including radiology and critical care. If 90-day mortality significantly reflects the incidence and management of complications, this metric may perhaps be strongly influenced by factors and resources outside the direct influence of the responsible consultant. In addition, the ability to successfully manage complications may be associated with the volume of cases treated within a unit\textsuperscript{220}. While these factors may be important, more work is needed to better understand why a unit may become an outlier as the follow-up period is extended from 30 to 90 days.

One unit had below average mortality at 30 days, but became a high outlier by the 90th postoperative day. Unadjusted mortality for this unit was initially low (2.3% at 30 days), rising to within 0.5% of the national average at 365 days.
It had a low risk patient cohort that was younger and less comorbid, with more women, when compared with the national cohort. It also performed relatively more elective procedures for malignant disease. The risk-adjustment model would calculate a lower-than-average expected mortality for such a cohort. Therefore the O/E ratio for this unit was significantly greater than 1, resulting in this unit’s HMU status.

Previous publications examining the relationship between mortality rates across different time periods for non-colorectal specialties have shown good correlation when follow-up is cumulative\textsuperscript{221,222}. One colorectal paper assessed correlation of mortality results between non-overlapping time periods (0-30 days, 30-days to 1-year, and 1-5 years) and also found good correlation, with Spearman’s $\rho=0.62$ and $\rho=0.52$ for correlation of 0-30 day mortality with 30-day to 1-year, and 1-5 year mortality, respectively\textsuperscript{171}. The authors suggested that this represented consistent quality of care throughout the follow-up period studied.

In the present study, extending follow-up from 30 to 90 days after surgery resulted in the identification of twice the number of units with high outlying mortality, including all 30-day high outliers. It also increased the number of units with low outlying mortality. 90-day follow-up allows a longer time period for the complications of surgery to become manifest. Consequently, 90-day mortality may also, perhaps, provide a better reflection of different institutions’ abilities to detect and successfully manage complications. More detailed work is needed to better understand the differences observed to fully appreciate the relative merits of 30-versus 90-day mortality for assessing departmental level performance. However, the present results suggest that adopting 90-day mortality will not result in failure to detect 30-day high outliers, and may identify more outliers that could poten-
potentially have problems with quality of care.

7.6 Chapter summary

This study has explored how varying the follow-up period after colorectal surgery may influence department-level mortality results. A prolonged, 90-day period of follow-up resulted in the detection of a greater number of units with outlying performance, although further work is needed to better understand the observed findings. The threshold for identification of outliers must be considered carefully. A balance must be struck between the appropriate identification of outliers where genuine problems with care may exist, and adopting metrics which identify a greater number outliers and fewer genuine quality of care issues. It may be considered that, as a performance measure, 90-day mortality may provide a broader assessment of the quality of care for colorectal patients, given the likely importance of complication management and post-discharge care incorporated in this measure. This should be considered, when deciding how to assess the performance of colorectal units generally.
Chapter 8

Study 4: Outlier identification in colorectal surgery should separate elective and non-elective service components

8.1 Chapter overview

This chapter describes a study using HES data to explore the relationship between elective and non-elective mortality rates after colorectal surgery at the departmental level. The previous study examined how the definition of a particular outcome measure influenced the assessment of unit-level performance for the same group of patients. This study assesses the relationship between performance measures across different patient groups treated within the same unit.

8.2 Introduction

It is well established that emergency colorectal surgery carries a significantly greater risk of death or complication than elective surgery\textsuperscript{22,24,213,223}. However,
it has not been established whether colorectal units with high outlying mortality rates for elective surgery also have outlying death rates for non-elective surgery. If outlier status is closely related between elective and non-elective surgery, one may be used as a proxy measure of the other. However, if the department-level outcomes of these patient cohorts differ, it may be necessary to examine both groups to obtain a comprehensive picture of the performance of a particular unit. Internationally, this may have implications for organisations and registries that provide feedback of outcomes to hospitals.

The relationship between unit-level elective and non-elective outcomes may also help better understand what are the key determinants of performance in these areas of colorectal practice. For example, a unit may be well organised for the reliable and effective delivery of planned surgery, with efficient pathways of care. However, the same unit may perhaps struggle to cope with unanticipated or out-of-hours demands. Such a unit may have good elective outcomes, with average or poor non-elective results. Differences in relative performance may provide the opportunity for closer study of individual units to develop the mechanistic understanding of the determinants of surgical outcomes. This contributes to the central objectives of this thesis to understand the complexity of surgical performance in greater detail, and how high performance may be achieved.

The present study examines institutional mortality outlier status after elective and non-elective colorectal surgery between 2006 and 2012, across all English colorectal units within the NHS using HES data. This will help inform performance monitoring practices and direct future quality improvement research.
Study question

- Is there any difference between elective and non-elective mortality in the assessment of relative unit-level performance after colorectal surgery?

8.3 Methods

The methods used in this study are very similar to those used in study 3, and are described in detail in the earlier chapter on using administrative data (chapter 5). The process of data extraction, processing and analysis was the same, with study-specific details outlined in the following parts of this chapter. While the data and techniques were similar, the data was used to answer a different research question in the present study.

8.3.1 Inclusion and exclusion criteria

All patients aged 18 years or more, admitted to English NHS hospitals for colorectal surgery between April 2006 and March 2012, were selected for inclusion using the OPCS codes provided in table 8.1. Patients were excluded if they received treatment at Trusts that did not provide routine elective and emergency colorectal surgery to an unselected patient population. For example, women’s health and cardiothoracic subspecialist Trusts were excluded, as were tertiary referral units with no routine acute admissions service.
Table 8.1: Operations and OPCS codes selected for inclusion, coded into groups for analysis.

8.3.2 Definition of variables and outcomes

For this study, the variables were defined and grouped similarly to those described in study 3, on page 169 of chapter 7, with the following differences:

- **Operative procedure** was defined as shown in table 8.1.

- **Laparoscopic approach**, including cases converted to open, was defined using the following codes: Y50.8, Y52.8, Y75.1-4, Y75.8-9, Y76.3, Y76.5, Y76.8 and Y76.9. Different codes were used in this study, compared with study 3, due to the introduction of updated versions of the OPCS coding system over subsequent years.

- **Year of procedure** was defined from 1st April to 31st March, rather than by calendar year.

The outcome of interest for this study was 90-day all-cause mortality. The
results of patients treated in Trusts that merged during the study period were analysed under their merged organisation.

8.3.3 Statistical analysis

The analysis performed was very similar to that conducted in study 3. For all tests, a $p$-value of $<0.05$ was considered statistically significant. A detailed explanation of the analysis methods is provided in section 5.3.2 of chapter 5.

8.4 Results

8.4.1 Cohort characteristics

195 118 patients treated at 147 NHS Trusts met inclusion criteria. The cohort characteristics are shown in table 8.2. Of 131 902 elective patients, 4 999 (3.8%) died in the first 90 postoperative days. For non-elective surgery, 11 706 of 63 286 (18.5%) patients did not survive beyond 90 days. There were statistically significant differences between the elective and non-elective patient cohorts for all measured characteristics.

8.4.2 Risk-factors for death at 90 days

All variables had significant associations with death on simple analysis and were entered into multiple regression models. Table 8.3 shows regression results. The only notable difference between elective and non-elective cohorts related to the association between patient sex and death. Model fit was satisfactory ($c$-statistics range 0.756-0.826).
### Table 8.2: Patient characteristics for whole cohort, and elective and non-elective subgroups.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Elective</th>
<th>Non-elective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;55</td>
<td>37%</td>
<td>35%</td>
</tr>
<tr>
<td>55-69</td>
<td>43%</td>
<td>34%</td>
</tr>
<tr>
<td>70-79</td>
<td>29%</td>
<td>29%</td>
</tr>
<tr>
<td>≥80</td>
<td>18%</td>
<td>18%</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>51%</td>
<td>54%</td>
</tr>
<tr>
<td>Female</td>
<td>49%</td>
<td>46%</td>
</tr>
<tr>
<td>CCI</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>26%</td>
<td>19%</td>
</tr>
<tr>
<td>1-2</td>
<td>45%</td>
<td>31%</td>
</tr>
<tr>
<td>3-5</td>
<td>15%</td>
<td>10%</td>
</tr>
<tr>
<td>≥6</td>
<td>14%</td>
<td>15%</td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malignant</td>
<td>59%</td>
<td>71%</td>
</tr>
<tr>
<td>Benign</td>
<td>41%</td>
<td>29%</td>
</tr>
<tr>
<td>Resection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total/subtotal</td>
<td>6.8%</td>
<td>5.3%</td>
</tr>
<tr>
<td>Right</td>
<td>36%</td>
<td>33%</td>
</tr>
<tr>
<td>Left</td>
<td>15%</td>
<td>16%</td>
</tr>
<tr>
<td>Approach</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open</td>
<td>76%</td>
<td>68%</td>
</tr>
<tr>
<td>Minimal access</td>
<td>24%</td>
<td>32%</td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006-7</td>
<td>16%</td>
<td>16%</td>
</tr>
<tr>
<td>2007-8</td>
<td>17%</td>
<td>17%</td>
</tr>
<tr>
<td>2008-9</td>
<td>17%</td>
<td>17%</td>
</tr>
<tr>
<td>2009-10</td>
<td>17%</td>
<td>17%</td>
</tr>
<tr>
<td>2010-11</td>
<td>18%</td>
<td>18%</td>
</tr>
</tbody>
</table>

**CCI** - Charlson Comorbidity Index; *p*-values calculated using χ² test.
<table>
<thead>
<tr>
<th>Age (years)</th>
<th>&lt;55</th>
<th>55-69</th>
<th>70-79</th>
<th>≥80</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>Ref</td>
<td>2.87</td>
<td>6.00</td>
<td>11.88</td>
</tr>
<tr>
<td>95% CI</td>
<td>2.66 - 3.11</td>
<td>5.52 - 6.43</td>
<td>11.00 - 12.84</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sex</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>1.01</td>
<td>0.86</td>
</tr>
<tr>
<td>95% CI</td>
<td>0.97 - 1.06</td>
<td>0.83 - 0.89</td>
</tr>
<tr>
<td>p</td>
<td>0.55</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CCI</th>
<th>0-1</th>
<th>2</th>
<th>3-5</th>
<th>≥6</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>Ref</td>
<td>2.42</td>
<td>4.93</td>
<td>5.50</td>
</tr>
<tr>
<td>95% CI</td>
<td>2.30 - 2.54</td>
<td>4.62 - 5.26</td>
<td>5.16 - 5.86</td>
<td></td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Admission</th>
<th>Elective</th>
<th>Non-elective</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>4.32</td>
<td>0.94</td>
</tr>
<tr>
<td>95% CI</td>
<td>4.15 - 4.50</td>
<td>0.89 - 1.00</td>
</tr>
<tr>
<td>p</td>
<td>0.04</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>Malignant</th>
<th>Benign</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>3.30</td>
<td>3.08</td>
</tr>
<tr>
<td>95% CI</td>
<td>3.15 - 3.45</td>
<td>2.85 - 3.33</td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Resection</th>
<th>Total/subtotal</th>
<th>Right</th>
<th>Left</th>
<th>Rectal</th>
<th>Open</th>
<th>Minimal access</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>Ref</td>
<td>0.65</td>
<td>0.58</td>
<td>0.62</td>
<td>Ref</td>
<td>0.55</td>
</tr>
<tr>
<td>95% CI</td>
<td>&lt;0.001</td>
<td>0.60 - 0.69</td>
<td>0.54 - 0.63</td>
<td>0.58 - 0.66</td>
<td>Ref</td>
<td>0.52 - 0.58</td>
</tr>
<tr>
<td>p</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>Ref</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Approach</th>
<th>Open</th>
<th>Minimal access</th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>Ref</td>
<td>0.55</td>
</tr>
<tr>
<td>95% CI</td>
<td>Ref</td>
<td>0.52 - 0.58</td>
</tr>
<tr>
<td>p</td>
<td>Ref</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>OR</td>
<td>Ref</td>
<td>0.94</td>
<td>0.91</td>
<td>0.87</td>
<td>0.76</td>
<td>0.70</td>
</tr>
<tr>
<td>95% CI</td>
<td>Ref</td>
<td>0.89 - 1.00</td>
<td>0.86 - 0.96</td>
<td>0.82 - 0.92</td>
<td>0.71 - 0.81</td>
<td>0.66 - 0.75</td>
</tr>
<tr>
<td>p</td>
<td>Ref</td>
<td>0.04</td>
<td>0.01</td>
<td>0.03</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>
8.4.3 Department-level performance in elective and non-elective surgery

Funnel plots of department-level case-mix adjusted 90-day mortality against total caseload are shown in figures 8.1 to 8.4. In figures 8.1 and 8.2, units with mortality more than 3 standard deviations above or below the mean (defined as high or low mortality units) for elective surgery are depicted with different markers, as shown on the legend. In figures 8.3 and 8.4, outliers for non-elective surgery are marked.

![Funnel plot of elective 90-day mortality against caseload.](image)

Figure 8.1: Funnel plot of elective 90-day mortality against caseload. Elective outliers marked: HMU - high mortality unit; LMU - low mortality unit.

For elective surgery, figure 8.1 shows that there were 3 high and 7 low mortality units. For non-elective surgery, figure 8.3 shows that there were 2 high and 3 low mortality units. Inspection of these plots reveals that units with high outlying
Figure 8.2: Funnel plot of non-elective 90-day mortality against caseload. Elective outliers marked: HMU - high mortality unit; LMU - low mortality unit.
Figure 8.3: Funnel plot of non-elective 90-day mortality against caseload.
Non-elective outliers marked: HMU - high mortality unit; LMU - low mortality unit.
Figure 8.4: Funnel plot of elective 90-day mortality against caseload.
Non-elective outliers marked: HMU - high mortality unit; LMU - low mortality unit.
Elective vs non-elective surgery do not necessarily also have high mortality for non-elective surgery (figures 8.1 and 8.2). All three elective high mortality outliers were within 95% control limits of the cohort mean for non-elective mortality. Similarly, units with low outlying elective mortality did not necessarily have very low non-elective death rates. None of the 7 elective low mortality units were low mortality outliers for non-elective surgery. These units did, however, appear to tend to have lower than average non-elective mortality, with four units between 95 and 99.7% control limits below the mean, and all but one of these units having below average death rates.

Examination of figures 8.3 and 8.4 shows that neither of the 2 high mortality units for non-elective surgery were high outliers for elective surgery. Both of these units had elective mortality rates within 95% control limits. Non-elective low mortality outliers also had average elective mortality results.

Kolmogorov-Smirnov testing suggested that overall department-level O/E mortality ratios for elective and non-elective surgery were normally distributed (lower limit of significance reported as \( p=0.200 \)). Graphical assessment was consistent with this, with a bell-shaped distribution and one high outlier for elective mortality. Therefore, Pearson’s correlation coefficient was determined, showing moderate correlation between elective and non-elective O/E mortality rates (\( r=0.503, p<0.001 \)).

### 8.5 Discussion

In this study, none of the colorectal units with high outlying mortality after elective surgery were also high outliers after non-elective surgery, and vice-versa for non-
elective high mortality units. Low mortality outlier status was also not closely associated between elective and non-elective colorectal practice. These results suggest that separate examination of elective and non-elective outcomes is required to provide a complete assessment of an individual unit’s performance.

The key strengths and weaknesses of studies based on HES data are described in section 5.2.3 of chapter 5. As a population-based study, the findings of this paper should be representative of colorectal surgery in similar health care systems. However, generalisation to other surgical specialties may be limited by differences in patient, disease and physiological characteristics. In addition, this observational study cannot explain why the reported differences have arisen.

The death rates presented in this study are comparable with published data on mortality after emergency and elective colorectal surgery\textsuperscript{24,62,223}, although more deaths will inevitably be captured by extending the follow-up period beyond the in-hospital stay or 30th postoperative day\textsuperscript{217}. The finding that the units included in this study did not perform at the same level for elective and non-elective surgery reinforces the earlier findings of Ingraham et al within the ACS NSQIP group of hospitals\textsuperscript{62}. However, although the present study found only moderate correlation between department-level adjusted mortality rates for elective and non-elective surgery, the cited study found no significant association. This may be due to important differences in care organisation. England has a comprehensive care system, where an individual unit’s elective and non-elective patient cohorts are largely determined by geography. In the United States, socio-economic variations in healthcare coverage may have a complicated and unknown effect on the characteristics and outcomes of elective and non-elective cohorts treated at a particular hospital. In addition, the ACS NSQIP group of hospitals are self-selected partic-
participants in a comprehensive quality improvement program, perhaps systematically differentiating them from other American hospitals or other health care systems internationally.

The differences in performance between patient cohorts may to some extent be explained by the different pathways followed by elective and non-elective patients, as well as differences in their demand for specific resources within a hospital. For example, the acute physiological changes associated with non-elective presentation and the higher risk of adverse outcomes in these patients\textsuperscript{22,62} may result in greater need for critical care input. There may also be more demand for radiological investigations and treatment among these patients. Certainly, there is some evidence that emergency general surgical outcomes may be influenced by the local level of these non-surgical resources\textsuperscript{228}.

Elective and non-elective groups may also present different challenges to the organisation and coordination of care within a unit. Elective care may perhaps be more effectively programmed and coordinated ahead of time, with each of the various perioperative phases following a pre-planned sequence of steps that can be readily predicted. The need to respond to unanticipated events may be lower than for non-elective surgery. It is known that non-elective patients are more likely to experience a postoperative complication, such as a heart attack or chest infection. Non-elective care is therefore more unpredictable and may place greater demands upon the team’s ability to respond to unexpected and rapidly evolving clinical events. This may make the quality of the relationship, and freedom of communication, between nursing and medical staff relatively more important in the non-elective setting. In this way, the way care is organised and the level of particular local resources, may perhaps have a differential impact on the outcomes
of elective and non-elective patients.

The findings of this study have implications for outcome reporting, which is becoming increasingly important internationally, with publication of hospital- and / or consultant-level outcomes in countries such as the United States and England\textsuperscript{224–226}. For colorectal surgery, there are three options for outcome reporting when considering the urgency of operation: elective only; non-elective only; or both elective and non-elective cohorts. This study has shown that scrutiny of elective outcomes alone (current practice in England) results in an incomplete picture of performance. Units with outlying mortality for non-elective patients are not identified, as elective and non-elective outlier status and outcomes are not closely associated. Non-elective surgery is a significant part of routine colorectal surgical practice with a high mortality rate and should not be neglected. The same arguments apply to the reporting of non-elective outcomes alone. To provide a complete appraisal of the performance of a colorectal service, both elective and non-elective outcomes must be examined.

Future research examining the outcomes of both cohorts in parallel may help further current understanding of the key determinants of outcome after colorectal surgery. If a unit has average performance for elective patients and a high mortality rate after non-elective surgery, examination of local practices may help doctors and researchers identify specific aspects of care unique to the non-elective pathway that are closely linked with patients’ outcomes.

Overall, a lack of close association between elective and non-elective department-level outcomes suggests that appraisal of the performance of a colorectal unit requires scrutiny of both of these areas of practice. The differences reported may reflect the different demands that these patient groups make upon the organisation
of care and the resources available within individual hospitals. Further research should investigate this area of clinical practice in greater detail to identify the reasons for the findings reported.

8.6 Chapter summary

In this study, departmental-level elective and non-elective colorectal surgical mortality rates were not closely associated. Thorough appraisal of the overall performance of a colorectal unit should include examination of both elective and non-elective outcomes. Relative differences in unit-level performance across these cohorts provides an opportunity for future, in-depth study to better understand the key determinants of outcome for these patient groups.
Part IV

Investigating colorectal structures and processes
Section summary

This section of the thesis presents two studies designed to better understand how colorectal units achieve their results. The first, study 5, uses publicly available data on organisations and units within the NHS, to assess for differences in care between colorectal units at opposite ends of the spectrum of performance for postoperative length of stay. The second, study 6, involved the development of a semi-structured telephone interview to try and examine the key structures and processes of care in these units. Qualitative and quantitative analysis of the results are reported in chapters 10, 11 and 12. This work has contributed towards a better understanding of how high levels of performance can be achieved, in practical terms of the clinical care delivered by front-line staff, and points the way for future research to build upon the findings reported.
Chapter 9

Study 5: Comparing short and long length of stay units using existing data

9.1 Chapter overview

This chapter presents a study designed to explore for potential differences in care between colorectal units according to their length of stay using existing publicly available data. This work was undertaken in collaboration with Professor Paul Aylin and Dr Alex Bottle, of Dr Foster Intelligence.

9.2 Introduction

The original studies presented in this thesis have, up to this point, focused primarily on examining the outcomes of colorectal surgery. However, outcomes cannot easily be used to inform improvements in the quality of care in clinical practice. This study set out to try and better understand differences in performance level among colorectal units, stratified by length of stay, using data that is already collected and made freely available within the NHS in England.
Study question

- Can existing data in the NHS be used to define measurable structure, process and outcome differences between colorectal units with high or low outlying length of stay results?

9.3 Methods

9.3.1 Unit selection

In this section, the choice to select units by length of stay is explained, and the inclusion and exclusion criteria are specified. The method of determining length of stay results is also described.

Choosing length of stay

A range of outcome measures for colorectal surgery may be derived from routinely collected administrative data. These include length of stay, readmission rate, 30-day mortality, re-operation rates, abdomino-perineal excision rates and surgical failure-to-rescue. Length of stay was selected as the outcome measure for this study for a number of reasons:

- The length of stay in hospital is an important determinant of resource use and thereby an indicator of quality.

- All patients undergoing surgery contribute to departmental length of stay measures. In marked contrast, postoperative mortality has fallen significantly, and now only affects a relatively small proportion of most patient cohorts.
• There is less potential stigma or professional sensitivity to variation in length of stay, compared with outcomes such as post-operative mortality.

Unit inclusion and exclusion criteria

20 colorectal units within the NHS in England were selected for inclusion in this study. This sample comprised 10 units each with the highest and lowest outlying length of stay results. This purposive sample was chosen to try and maximise potential differences between the study units.

Units were excluded as follows: if they only treated a highly selected patient population (for example, units in women’s health or paediatric Trusts); if they were private health care providers; or if they had a low volume of operations, with fewer than 20 eligible procedures during the 2 year study period.

Obtaining outcomes

Unit-level risk-adjusted length of stay results were determined for adult patients undergoing an elective colonic resection between January 2011 and December 2012.

Results were retrieved from the Dr Foster Intelligence ‘Real-Time Monitoring, version 8’ (RTMv8) web-based tool. Dr Foster Intelligence is an independent company that routinely obtains and processes HES data. The RTMv8 web interface provides performance data across a broad range of areas of hospital care, that can be extensively customised according to the user’s needs. It calculates risk-adjusted outcomes using a number of covariables, previously described by Bottle and Aylin\(^ {229} \). The outcomes determined include a measure of in-hospital stay, called ‘long length of stay’ (LLoS). LLoS is a binary outcome determined for each patient by comparing individual length of stay with the 75th centile length of stay.
for the entire national cohort defined by the user. A patient is defined as having a ‘long’ stay if their stay exceeds this cut-off. Once a particular cohort and outcome has been specified, the RTMv8 tool generates output that includes the risk-adjusted frequency of LLoS for a provider, standardised to a national average of 100, with 95% confidence intervals. These confidence intervals allow the user to determine if a particular unit has above or below average LLoS results at a significance level of $p < 0.05$. This tool was used to create a table of the LLoS results of all units in the NHS in England meeting the inclusion and exclusion criteria. (A detailed explanation of how this was done is provided in appendix E.) From 179 hospitals, 10 units at each end of the spectrum of performance with outlying results were selected.

9.3.2 Structure, process and outcome data for comparison

Data was collated on selected units and Trusts from a number of sources, including the Health Protection Agency, NHS England, Dr Foster Intelligence, the Picker Institute, and individual Trusts’ websites. Certain data points were selected because previous research had demonstrated an association with surgical outcomes. For example, Symons et al found an independent association between both a higher proportion of critical care beds and greater use of computed tomography (CT) scans per bed, and lower mortality after emergency general surgery\textsuperscript{228}. A relationship between higher nurse staffing levels and improved outcomes has also been demonstrated previously\textsuperscript{78}. In addition, there is increasing acceptance that hospital-wide outcomes, such as Hospital Standardised Mortality Ratio (HSMR), may be a useful indicator of performance. The first inquiry into care at the Mid
Staffordshire NHS Foundation Trust was triggered as a result of a higher-than-average HSMR, and clearly documented wide-ranging problems with care in that institution\textsuperscript{15}. Other data were selected to ensure a broad range of structures, process and outcomes were examined, according to availability from the existing routine data sets.

A list of the 56 data items is provided in table 9.1. Where appropriate, data was adjusted for the size of the Trust before comparison between units. For example, to compare staffing levels, numbers of staff or operations cancelled, absolute figures for these variables were divided by the number of beds in the Trust. Certain variables were derived by combining information from multiple sources. For example, to assess the level of data submission to the colorectal cancer surgery specialist register, the National Bowel Cancer Audit Program, numbers of operations reported in that database were compared with numbers identified from the HES database (‘NBOCAP case capture (%)’). Comparisons were performed between groups of units using the independent samples Mann-Whitney U test. As the sample size was small, \( p<0.10 \) was considered statistically significant.

\textbf{9.3.3 Ethical considerations}\\
Professor Aylin and Dr Bottle, at Dr Foster Intelligence, have permission from the Confidential Advisory Group under Section 251 of the NHS Act 2006 (formerly Section 60 Approval from the Patient Information Advisory Group) to hold confidential NHS data and analyse them for research purposes. They also hold ethical approval for such work from the South East Research Ethics Committee (ref: 10/H1102/25).
<table>
<thead>
<tr>
<th>Structures (n=11)</th>
<th>Outcomes (n=29)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical staff per bed</td>
<td>Financial surplus (%)</td>
</tr>
<tr>
<td>Consultants per bed</td>
<td>Financial comp. income vs total (%)</td>
</tr>
<tr>
<td>Nurses / midwives per bed</td>
<td>MRSA rate</td>
</tr>
<tr>
<td>Managers per bed</td>
<td>C. difficile apportioned rate</td>
</tr>
<tr>
<td>Managers per medical staff</td>
<td>C. difficile total rate</td>
</tr>
<tr>
<td>Managers per consultant</td>
<td>E. coli rate</td>
</tr>
<tr>
<td>Managers per nurse / midwife</td>
<td>SHMI</td>
</tr>
<tr>
<td>Percent consultant medical staff</td>
<td>HSMR</td>
</tr>
<tr>
<td>Ratio support staff to medical</td>
<td>Deaths after surgery</td>
</tr>
<tr>
<td>Ratio support staff to consultants</td>
<td>Deaths in low risk conditions</td>
</tr>
<tr>
<td>Ratio critical care to all beds</td>
<td>NBOCAP death rate</td>
</tr>
<tr>
<td>Process (n=16)</td>
<td>HES death rate</td>
</tr>
<tr>
<td>NRLS reporting rate</td>
<td>Staff survey:</td>
</tr>
<tr>
<td>Percent cancer waits within 62 days</td>
<td>- Satisfaction w/quality (%)</td>
</tr>
<tr>
<td>Staff survey response rate</td>
<td>- Work pressure</td>
</tr>
<tr>
<td>Friends and Family Test response rate</td>
<td>- Work stress (%)</td>
</tr>
<tr>
<td>Inpatient survey response rate</td>
<td>- Witnessed harm (%)</td>
</tr>
<tr>
<td>Cancer survey response rate</td>
<td>- Good comm. management</td>
</tr>
<tr>
<td>Urgent operations cancelled per bed</td>
<td>- Job satisfaction</td>
</tr>
<tr>
<td>Elective operations cancelled per bed</td>
<td>- Overall engagement</td>
</tr>
<tr>
<td>CTs per bed</td>
<td>Friends and Family Test score</td>
</tr>
<tr>
<td>MRIs per bed</td>
<td>Inpatient survey:</td>
</tr>
<tr>
<td>USSs per bed</td>
<td>- Doctors rating</td>
</tr>
<tr>
<td>CTs per consultant</td>
<td>- Nurses rating</td>
</tr>
<tr>
<td>MRIs per consultant</td>
<td>- Ops and procedures rating</td>
</tr>
<tr>
<td>USSs per consultant</td>
<td>- Overall rating</td>
</tr>
<tr>
<td>NBOCAP case capture (%)</td>
<td>Cancer survey:</td>
</tr>
<tr>
<td>NBOCAP death capture (%)</td>
<td>- Involvement in decisions</td>
</tr>
<tr>
<td></td>
<td>- Operation explanation</td>
</tr>
<tr>
<td></td>
<td>- Trust in doctor</td>
</tr>
<tr>
<td></td>
<td>- Trust in ward nurse</td>
</tr>
<tr>
<td></td>
<td>- Treated with dignity and respect</td>
</tr>
</tbody>
</table>

Table 9.1: Data compared between short and long stay units.
NRLS - National Reporting and Learning System; CT - Computed Tomography; MRI - Magnetic Resonance Imaging; USS - Ultrasound Scan; NBOCAP - National Bowel Cancer Audit Program; MRSA - Methicillin-Resistant Staphylococcus Aureus; C. difficile - Clostridium difficile; E. coli - Escherichia coli; SHMI - Summary Hospital-level Mortality Indicator; HSMR - Hospital Standardised Mortality Ratio.
9.4 Results

Descriptive characteristics for the 20 units compared by length of stay are provided in table 9.2. There was no statistically significant difference in the number of cases performed in units with short or long length of stay results \((p = 0.433, \text{ independent samples } T \text{ test, equal variances not assumed})\).

<table>
<thead>
<tr>
<th>LLoS performance</th>
<th>Participation</th>
<th>LLoS results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR</td>
<td>CI</td>
</tr>
<tr>
<td>Short</td>
<td>Yes</td>
<td>23.5</td>
</tr>
<tr>
<td>Short</td>
<td>Yes</td>
<td>31.7</td>
</tr>
<tr>
<td>Short</td>
<td>Yes</td>
<td>33.5</td>
</tr>
<tr>
<td>Short</td>
<td>Yes</td>
<td>38.0</td>
</tr>
<tr>
<td>Short</td>
<td>Yes</td>
<td>38.4</td>
</tr>
<tr>
<td>Short</td>
<td>Yes</td>
<td>41.9</td>
</tr>
<tr>
<td>Short</td>
<td>Yes</td>
<td>47.3</td>
</tr>
<tr>
<td>Short</td>
<td>Yes</td>
<td>48.3</td>
</tr>
<tr>
<td>Short</td>
<td>Yes</td>
<td>49.3</td>
</tr>
<tr>
<td>Short</td>
<td>Yes</td>
<td>49.6</td>
</tr>
<tr>
<td>Long</td>
<td>Yes</td>
<td>150.2</td>
</tr>
<tr>
<td>Long</td>
<td>Yes</td>
<td>154.7</td>
</tr>
<tr>
<td>Long</td>
<td>No</td>
<td>154.9</td>
</tr>
<tr>
<td>Long</td>
<td>Yes</td>
<td>159.6</td>
</tr>
<tr>
<td>Long</td>
<td>No</td>
<td>162.0</td>
</tr>
<tr>
<td>Long</td>
<td>No</td>
<td>167.7</td>
</tr>
<tr>
<td>Long</td>
<td>No</td>
<td>173.5</td>
</tr>
<tr>
<td>Long</td>
<td>No</td>
<td>173.9</td>
</tr>
<tr>
<td>Long</td>
<td>Yes</td>
<td>196.9</td>
</tr>
<tr>
<td>Long</td>
<td>Yes</td>
<td>275.9</td>
</tr>
</tbody>
</table>

Table 9.2: Length of stay results and participation status of all 20 units approached.
RR - Relative Risk; CI - 95% Confidence Intervals; n - total cases.

Of the 56 data items compared between units with high and low outlying LLoS results, only one was statistically significantly different at \(p<0.10\) (see table 9.3).
Units with the shortest length of stay had a higher response rate to the cancer patient survey when compared with long stay units.

<table>
<thead>
<tr>
<th>Data item</th>
<th>LLoS</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Cancer survey: response rate (%)</td>
<td>69.5</td>
<td>62.5</td>
</tr>
</tbody>
</table>

Table 9.3: Significant difference between low and high LLoS units. Median values provided.

9.5 Discussion

Of the large number of data items compared between colorectal units with the shortest and longest length of stay after elective colonic surgery in England, only one was statistically significantly different across these two groups. Short stay units had significantly higher response rates to the cancer patient survey. This finding should be interpreted in the context of the large number of significance tests performed (56 in total), and the decision to adopt a significance level of $p < 0.10$. Given these parameters, one may expect to find 5 or 6 significant differences between these groups by chance alone. Therefore, the single significant difference reported is likely to represent a chance finding due to multiple comparisons.

This study has a number of important limitations. Units were selected for inclusion using outcomes derived from the HES database, the strengths and weaknesses of which have been discussed previously (see chapter 5). Importantly, the majority of the data compared between units stratified by LLoS was aggregated at the Trust level. Local variation of structures, processes and outcomes in the colorectal department could not be captured, as such data is not publicly available.
In addition, the sample size was small, with only 20 units included in the analysis. Therefore, this study may have been prone to type II error, where a true difference in the measured variables has not been detected.

In this study, the range of selected data items did not help explain how the included units achieved their respective levels of performance. For example, units with the shortest length of stay in England did not have significantly different medical or nursing staffing levels per bed, compared with long length of stay units. Similarly, the high performing units examined did not have higher operative volume than those with low performance. There were also no differences in the proportion of critical care beds, use of imaging tests, health care acquired infection rates, or postoperative death rates.

The lack of an association between the selected units’ performance and the structure, process and outcome data examined may have arisen for one of two reasons. Firstly, there may have been no true relationship between the individual data points and length of stay after colorectal surgery. In this case, even with a much larger sample, the study would have found no difference between short and long stay units. Secondly, a true relationship did exist which this study failed to detect due to methodological considerations. For example, the sample size was small and may have lacked the statistical power to detect the underlying relationship. Alternatively, a true relationship may not have been detected because of the lack of granularity among the data points examined. For example, staffing levels, imaging use and health care acquired infection rates were aggregated across whole Trusts. Unit-level variation in each of these areas could not be captured. This may have masked any true association between the data examined and unit performance.
The lack of a gross association between variables such as operative volume or nurse staffing levels, and postoperative length of stay after elective surgery is an interesting finding. The relationship between operative volume and postoperative mortality has been repeatedly shown in a number of studies across a range of surgical specialties, as discussed in chapter 3. The extremes of length of stay performance may perhaps depend upon rather different aspects of clinical care and team functioning than postoperative mortality. Further work should explore whether the lack of association documented in the present study is also present after emergency surgery, or in other surgical specialties.

On the basis of the findings presented, the existing, routinely collected data was not able to define measurable structure, process or outcome differences between colorectal units stratified by length of stay. Therefore, to understand how these units achieved their different levels of performance, alternative data must be collected.

### 9.6 Chapter summary

This chapter describes a study designed to try and understand how units with the highest and lowest length of stay performance after elective colonic surgery in England achieved their results. It used a selection of routinely collected structure, process and outcome data within the NHS to try and define measurable differences between units stratified by their outcomes. Within the limitations of the sample size and data sources, no significant differences were found between the included units. Other, novel data, collected at the unit level, may be required to understand how, in practical terms, these units achieved their respective levels of performance.
This is the focus of study 6, presented in chapters 10, 11 and 12.
Chapter 10

Study 6: Understanding length of stay in elective colorectal surgery - methods and qualitative analysis

10.1 Chapter overview

A central objective of this thesis was to develop a better understanding of how colorectal units achieve their results, in practical terms of the care they deliver. This chapter presents the methods and qualitative results of a study designed to try to achieve this aim using a semi-structured telephone interview. The following chapter will present an exploratory analysis of quantitative data derived from the interview.

10.2 Introduction

The study described in the previous chapter demonstrated how existing structure, process and outcome data were not able to define the key elements of care associated with high performance. Therefore, original data needs to be collected specifically to understand this area of surgical care. The study presented in this
and the following chapters was designed to try to better understand the link between a specific outcome, length of stay after elective colorectal surgery, and the broad range of department-level structures and processes discussed in chapter 3.

A small number of studies have made similar attempts to assess the organisation of care, and demonstrate an association with outcomes, with limited success. In the United States, researchers within the VA system conducted two studies to validate the use of risk-adjusted postoperative outcomes as measures of care quality\textsuperscript{32,230}. In the first study, 20 of a total of 44 medical centres were selected for site visits\textsuperscript{32}. Five units were sampled from each end of the spectrum of performance for risk-adjusted mortality and morbidity. The site visit team conducted individual and group interviews, ‘walk rounds’ of the surgical intensive care unit and surgical floor, direct observation of work in these areas for 1-2 hours, and an exit meeting with the chief of surgery. Seven quality of care domains were assessed and reported. There were only 2 statistically significant associations with unit performance level: the mean technology score in the surgical intensive care unit, and the overall rating of care quality by the site visit team, were higher in units with the best performance. The second study used structured implicit chart review to assess quality of care, with a sample of 739 surgical cases from all 44 VA hospitals. This study failed to show any association between risk-adjusted outcomes and the overall quality of care judged by chart review.

In 2007, three more papers were published based on further work within the VA system, describing a process designed to develop a survey measure of structures and processes, and test for an association with risk-adjusted postoperative morbidity and mortality\textsuperscript{30,31,231}. The survey was based on the qualitative analysis of 44 face-to-face interviews\textsuperscript{231}. The questionnaire was circulated to 123 VA
hospital sites, and results were analysed for 90 sites that returned surveys. Univariate analysis showed significant associations between 14 variables measured by the survey and institutional risk-adjusted morbidity; fewer associations were found between survey measures and risk-adjusted mortality, with 4 significantly correlated variables. Factors associated with higher morbidity included having a larger number of operating rooms, having a higher percentage of returns reporting being short-staffed for nursing staff and aides, and a lower percentage of patients with the same anaesthetist during the pre-, intra- and post-operative operative phases of surgery. The final paper described the development of a hierarchical model to predict patient morbidity outcomes. University affiliation was reported as a dominant factor, significantly associated with increased risk of post-operative morbidity.

In 2009, Bradley et al described a strategy for understanding variation in performance which they labelled as the ‘positive deviance’ approach. While the present study was not based upon this approach, there are a number of similarities. Both assume that good or best practice already exists within the health care community. Both use purposive sampling, based upon an accepted measure of performance that has shown significant variation between units. This variation is a prerequisite to show that the metric has discriminant potential. The positive deviance approach then involves using qualitative methods to understand care at both ends of the spectrum of performance. The results of this work may inform the subsequent development of a quantitative assessment of individual units. The quantitative measure can then be applied to a larger sample of units and statistical tests performed, to determine the generalisation of the qualitative findings. This approach has been successfully used to understand and improve door-to-balloon
time in Acute Myocardial Infarction (AMI)\textsuperscript{232–234}, and to identify hospital strategies associated with lower Risk Standardised Mortality Rates after AMI\textsuperscript{235,236}.

Despite such work, the current understanding of how individual units achieve high performance is limited. The surgical studies cited above variously found some association between technology, staffing levels, continuity of care, university affiliation, and the resulting level of performance. The later medical studies suggested timely coordination of acute care and feedback on performance were important. However, it is clear that further work is required to investigate areas such as these in more detail, in colorectal practice, outside the American health care system.

Study question

• What are the key structures and processes in the organisation and delivery of care within a colorectal unit that determine its length of stay results?

10.3 Methods

10.3.1 Overall design

This study utilised a cross-sectional semi-structured interview to assess selected structures and processes in a sample of English colorectal units. Twenty units were selected for inclusion, ten each with the highest or lowest outlying risk-adjusted length of stay results in England, as described in study 5 (see chapter 9). Broad interviews, examining a range of areas within a colorectal unit, were conducted with a surgeon and nurse in each unit. The resulting data was intended for both qualitative and quantitative analysis (presented in the present and the following
chapters, respectively) to better understand how the organisation of care within a unit determines its length of stay results.

10.3.2 Choice of methodology

A variety of methods may be used to assess the structures and processes of surgical units. These include administrative datasets, questionnaires, interviews, chart review, site visits and direct observation. The most readily accessible data source, administrative data, has already been examined in the previous chapter. This type of data did not discriminate the key elements of care associated with short or long stay after elective colonic surgery.

For this study, a semi-structured telephone interview was chosen. It was considered that this approach was well suited to the aims of the study - to understand how care organisation within a unit influenced aggregate outcomes. Semi-structured, conversational interviews between a research clinician and members of the clinical team represent an effective approach for the collection of rich data in the areas of care selected for examination, such as standardisation of care, leadership and inter-professional communication. (The selection of themes is described in detail later in this methods section.)

A semi-structured interview also presents specific methodological advantages over the alternatives. These include the ability to combine open-ended questions that explore the perspective and understanding of the interviewee, with focused questions and scoring systems to try and obtain a quantitative result. Interviews tend to have a much higher response rate than questionnaires. Telephone interviews are also significantly cheaper than site visits. They are, however, not without
disadvantages. These include greater cost than questionnaires, in both administration and data analysis, and the potential bias introduced by the interviewer.

This approach has been used successfully to develop a deeper understanding of organisational performance in the manufacturing industry. Bloom and van Reenen described several important strategies they adopted to minimise bias in the data gathered using this technique. These included: selecting interviewees that have insight into the structure of an organisation as well as experience of its day-to-day functioning; blinding interviewers and interviewees to organisational performance status; and the use of open rather than closed questioning. They suggested techniques to evaluate bias by repeat interview or adoption of an alternative method to assess the same characteristics of the organisation under scrutiny.

Given its relative advantages, a semi-structured telephone interview, based on the methodology described by Bloom and van Reenen, was selected as the preferred method for assessing structures and processes in colorectal units in this study.

### 10.3.3 Unit selection

Units were selected for inclusion as described in study 5 (see chapter 9). To recap, a purposive sample of 20 colorectal units with the highest and lowest length of stay results in the NHS in England were identified. If a unit did not agree to participate and complete the interview within 2 months of initial approach, it was substituted for another unit with the next highest or lowest outlying results.
10.3.4 Interview development

Themes for the interview

The objective of the interview was to focus on examining department-level factors that may be important in determining unit-level outcomes. Furthermore, the goal was, as far as possible, to derive a generalisable understanding that could be applied to other fields of surgery. Thus, the interview was primarily designed to assess generic factors, such as standardisation of care and communication within the team, with more limited consideration of selected, specific clinical processes relevant to elective colorectal care.

The overall design of the study was based broadly upon Lilford’s expansion of Donabedian’s structures, processes and outcomes paradigm (see figure 1.1 on page 48). As surgical care is a largely linear process, it was considered that the organisation of care within a unit could be assessed according to each consecutive phase of the patient journey. Thus, factors influencing surgical performance could be classified as pre-operative, operative, or post-operative. There may also be overarching themes common to all stages of the patient journey. Other frameworks describing the organisation of care were also reviewed\textsuperscript{69,239}. Combining the themes from these frameworks yielded a list of factors that were considered important in the overall performance of a colorectal unit. These were mapped to the specific phases of the patient pathway or their global relevance, shown in tables 10.1 and 10.2 respectively.

While focusing on broad themes that could be relevant to any surgical speciality, it was also important to make the interview schedule colorectal-specific and relevant to respondents’ experiences. Therefore, questions examining general fac-
Table 10.1: Patient pathway factors important in determining the performance of a colorectal unit.

<table>
<thead>
<tr>
<th>Pre-operative</th>
<th>Operative</th>
<th>Post-operative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deception</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comorbidities</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Complication management**

- Minimise physiological insult
- Multidisciplinary input e.g. nursing
- Daily goals
- Enhanced Recovery Programme
- Routine management protocols

**Pre-operative**

- Minimise physiological insult
- Multidisciplinary input e.g. nursing
- Daily goals
- Enhanced Recovery Programme
- Routine management protocols

**Operative**

- Minimise physiological insult
- Multidisciplinary input e.g. nursing
- Daily goals
- Enhanced Recovery Programme
- Routine management protocols

**Post-operative**

- Minimise physiological insult
- Multidisciplinary input e.g. nursing
- Daily goals
- Enhanced Recovery Programme
- Routine management protocols
<table>
<thead>
<tr>
<th>Structures</th>
<th>Processes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Resources</strong></td>
<td><strong>Leadership</strong></td>
</tr>
<tr>
<td>• Equipment</td>
<td>Quality awareness and improvement</td>
</tr>
<tr>
<td></td>
<td>Standardisation</td>
</tr>
<tr>
<td>– Quality and quantity</td>
<td>Communication and collaboration</td>
</tr>
<tr>
<td>– Diagnostic, therapeutic and monitoring</td>
<td>Supervision and delivery of care</td>
</tr>
<tr>
<td>• Personnel</td>
<td>Culture</td>
</tr>
<tr>
<td>– Medical (surgical, anaesthetic, allied specialties; consultants, trainees)</td>
<td>• Staff valued</td>
</tr>
<tr>
<td>– Nursing</td>
<td>• Openness</td>
</tr>
<tr>
<td>– Support staff</td>
<td>Information management</td>
</tr>
<tr>
<td>• Building and space</td>
<td></td>
</tr>
<tr>
<td>– Theatres, ward, recovery, ITU, clinic, offices</td>
<td></td>
</tr>
<tr>
<td>– Adequacy, quality, inter-relationship</td>
<td></td>
</tr>
<tr>
<td><strong>Demand</strong></td>
<td><strong>Rota organisation</strong></td>
</tr>
<tr>
<td>• Patient need matches resources</td>
<td><strong>Leadership</strong></td>
</tr>
<tr>
<td>• Emergency <em>versus</em> elective care</td>
<td>Quality awareness and improvement</td>
</tr>
<tr>
<td><strong>Rota organisation</strong></td>
<td>Standardisation</td>
</tr>
<tr>
<td>• Staff shortages</td>
<td>Communication and collaboration</td>
</tr>
<tr>
<td>• In-hours <em>versus</em> out-of-hours</td>
<td>Supervision and delivery of care</td>
</tr>
<tr>
<td>• Adequate cover on-site</td>
<td>Culture</td>
</tr>
<tr>
<td>– Seniority</td>
<td>• Staff valued</td>
</tr>
<tr>
<td>– Numbers of staff</td>
<td>• Openness</td>
</tr>
</tbody>
</table>

Table 10.2: Over-arching factors important in determining the performance of a colorectal unit.
tors were tailored to provide a colorectal focus. In addition, a limited number of more specific questions were added. For example, participants were asked whether reducing length of stay was a priority in their department, and how the interviewee considered their unit performed relative to others in the NHS in England. It was also decided to include some fact-gathering closed questions at the end of the interview to obtain more detailed specific structural information, such as the number of colorectal surgeons in the unit.

Initial protocol

On the basis of the themes identified above, a preliminary interview schedule was created. The interview schedule started by summarising the objectives of the study, confirming consent and reminding participants that all responses were confidential. After this, an opening question asked interviewees to describe their role in the colorectal department, to help initiate the conversation and establish rapport. The main part of the interview then asked interviewees about 11 general structure and process factors, and patient management within the three phases of perioperative care. Five general structural factors were covered: equipment levels; personnel levels; building layout; out-of-hours cover; and the adequacy of departmental resources for local demand. Six broad process factors were examined in the preliminary protocol: leadership and culture; quality assessment and improvement; standardisation; communication and collaboration; supervision and escalation of care; and specific measures to reduce length of stay. Patient pathway-specific questions covered: pre-operative patient preparation; operative management priorities; and postoperative management including detection of complications. After this, participants were asked how they thought low length of stay could be
achieved. The final part asked a series of closed questions before the interview was concluded.

**Interview scoring**

As the study was designed to generate both qualitative and quantitative data, a scoring system was devised alongside the preliminary interview protocol. A 7-point ordinal scale was created to allow rating of interviewees’ answers to individual questions against exemplar responses, that were established as anchors for scores of 1, 3, 5 and 7. The exemplars were chosen to represent a very broad range of potential responses, considered to cover both ends of the spectrum of practice that could potentially be found within the NHS in England. The scoring system was refined iteratively during the piloting process before the final protocol was created, to allow interviews to be scored as they were conducted. Details of the qualitative analysis of interview content are provided later in this methods section.

**Identification of relevant evidence**

After achieving the broad overview of the factors considered relevant in determining department-level performance in colorectal surgery, as described above, a literature review was undertaken to identify relevant research. The methods and results of this review are described in chapter 3. Research relevant to each question was used to ensure that all parts of the preliminary interview merited inclusion, and to revise questions as required.
Interview piloting and refinement

After refinement of the interview protocol based on the relevant evidence identified, the interview schedule was piloted and developed further with 3 expert interviewees and 9 peers within the research department. The changes made to the protocol are described in appendix F. Through this process, a final interview protocol was created (see appendix G). This consisted of a short introduction, an opening question to establish the background of the interviewee, eleven semi-structured interview questions, one unscored open question and seven short, closed questions. The themes covered are provided in table 10.3.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>Equipment</td>
</tr>
<tr>
<td></td>
<td>Staffing</td>
</tr>
<tr>
<td>Process - clinical</td>
<td>Pre-operative assessment</td>
</tr>
<tr>
<td></td>
<td>Operative details</td>
</tr>
<tr>
<td></td>
<td>Routine postoperative management</td>
</tr>
<tr>
<td></td>
<td>Detection and management of complications</td>
</tr>
<tr>
<td>Process - organisational</td>
<td>Standardisation</td>
</tr>
<tr>
<td></td>
<td>Communication and collaboration</td>
</tr>
<tr>
<td></td>
<td>Leadership and culture</td>
</tr>
<tr>
<td></td>
<td>Attitudes to safety and adverse events</td>
</tr>
<tr>
<td></td>
<td>Outcomes assessment and feedback</td>
</tr>
</tbody>
</table>

Table 10.3: Areas of care organisation examined in the interview.

10.3.5 Interviewee selection

Within each hospital, a senior colorectal surgeon and senior colorectal ward nurse were selected for interview, ideally each of whom had worked at the Trust for 5 years or more. In accordance with Bloom and van Reenen’s suggested methodology, it was considered important to select staff with some seniority, but also with
experience of working directly with patients in the colorectal unit. Selection of two members of staff from different professional groups was intended to build a triangulated picture of practice in each unit. In addition, significant differences between surgeon and nurse ratings of care, if encountered, was considered a potentially important area that this study could explore. In most Trusts, after registering the study, the Research and Development or Audit department initiated contact with the colorectal department, who in turn identified suitable members of staff to participate in the study.

10.3.6 Interview conduct

The telephone interviews were conducted in July and August 2013, typically lasting between 30 and 40 minutes. Ben Byrne conducted all interviews. Using an interviewer with clinical experience as a surgical trainee within the NHS helped establish rapport with the participants quickly, maximising the quality of data obtained. Interviews were recorded where participants provided consent. Interviewees were not advised of their unit’s length of stay performance.

10.3.7 Data analysis

Audio recordings were transcribed verbatim by a professional transcription company (PageSix Transcription Services, Richmond, UK). Thematic qualitative analysis of interview transcripts was performed to identify similarities and differences in the organisation and delivery of care between the study units. This analysis was intended to define key, common areas of clinical practice associated with the contrasting levels of performance achieved by the units sampled. Initial
analysis was performed by coding the detail of each transcript. Intermediate and higher level categories were developed as connections emerged between coded sections. This resulted in a hierarchical tree of super- and sub-ordinate themes. Initial coding was performed independently by Ben Byrne and Anna Pinto on three randomly selected interviews. After meeting to review emerging themes, two further interviews were randomly selected and independently coded. A further meeting to review and consolidate codes was convened. There was good agreement between researchers at both meetings. All remaining analyses were completed by Ben Byrne. Coding was performed using NVivo 10 for Windows (QSR International Pty Ltd, Melbourne, Australia).

In addition, in-hospital mortality and 28-day readmission rates were retrieved for selected units from Dr Foster’s RTMv8 tool. Along with operative caseload and laparoscopy rates, these variables were compared between study units and the national cohort using appropriate parametric (independent samples t-test) or non-parametric (Mann-Whitney U test) tests, according to the distribution of the data, tested using the Kolmogorov-Smirnov test. For these tests, a $p$-value of $<0.05$ was considered statistically significant.

10.3.8 Ethical considerations

The Imperial College London and Imperial College Healthcare NHS Trust Joint Research Compliance Office advised that the interviews met criteria as a service evaluation and were exempt from ethical review. Local project approval processes were followed at participating hospitals. All participants were provided with an information sheet and returned a signed consent form before participating in the
10.4 Results

This results section initially examines the relationship between participation in the study and outcomes for the 20 units identified to take part in the study. Thereafter, the descriptive characteristics and qualitative analysis of interviews for the ten units with the shortest length of stay will be reported.

10.4.1 Participation

Colorectal units at 10 Trusts at each end of the LLoS spectrum were contacted. There were unanticipated difficulties in engaging Trusts and clinicians in the study. Of the 20 units initially approached, 15 took part. All 10 units with short length of stay agreed to participate, and only 5 of 10 with long stay took part (see table 10.4). High LLoS units were statistically significantly less likely to participate than low LLoS units \( (p = 0.033, \text{Fisher’s Exact test}) \). Attempts were made to substitute non-participating units, and 2 further hospitals with high LLoS were approached. However, neither took part.

<table>
<thead>
<tr>
<th>Participation status</th>
<th>Frequency of LLoS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>Low</td>
</tr>
<tr>
<td></td>
<td>High</td>
</tr>
<tr>
<td>No</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>0.033</td>
</tr>
</tbody>
</table>

Table 10.4: Participation against length of stay results for 20 units initially approached.

\( p \)-value calculated using 2-tailed Fisher’s Exact test.

A good sample of units with short length of stay was achieved, with all selected
units participating. However, the sample of units with long stay was inadequate. Given this, it was decided to undertake qualitative analysis of only those interviews conducted with short length of stay units. This analysis was intended to define areas of common practice across these high performing units, to try and develop the present understanding of the key determinants of length of stay after elective colonic surgery. It was also considered important to identify areas of variation in practice, which could therefore be considered of low importance in delivering short postoperative stay.

10.4.2 Descriptive unit characteristics

LLoS results and other descriptive characteristics of the 10 units included in this analysis are provided in table 10.5. Though not statistically significant, study units had above average laparoscopy rates \( (p=0.081, \text{t-test}) \), and lower than average in-hospital mortality \( (p=0.250, \text{Mann-Whitney U test}) \) and 28-day readmission rates \( (p=0.081, \text{t-test}) \).

10.4.3 Qualitative results

Qualitative thematic analysis identified three key themes in the organisation of care across the short stay units included in the study, as described in the following subsections.

Define and standardise clinical processes

Nine of the ten short stay hospitals had adopted a formalised patient pathway based upon the ERP, with the tenth having piloted the ERP and adopted many
<table>
<thead>
<tr>
<th>Site</th>
<th>LLoS results Cons. FTE ERP</th>
<th>Laparoscopy rates (%)</th>
<th>Mortality RR</th>
<th>Readmission RR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RR CI n</td>
<td>Cons. est. HES</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>23.5 10.1-46.2 158</td>
<td>3 2011 50 46</td>
<td>84.7</td>
<td>109.1</td>
</tr>
<tr>
<td>2</td>
<td>31.7 10.2-74.0 68</td>
<td>2.2 2010 70 58</td>
<td>77.2</td>
<td>49.1</td>
</tr>
<tr>
<td>3</td>
<td>48.3 23.1-88.8 91</td>
<td>3 2008 49 47</td>
<td>101.0</td>
<td>20.6*</td>
</tr>
<tr>
<td>4</td>
<td>49.6 30.3-76.7 193</td>
<td>5 2006 85 85</td>
<td>146.9</td>
<td>105.7</td>
</tr>
<tr>
<td>5</td>
<td>33.5 13.4-69.0 99</td>
<td>4 2007 80 79</td>
<td>0.0</td>
<td>91.9</td>
</tr>
<tr>
<td>6</td>
<td>38.0 15.2-78.3 91</td>
<td>4 2013 70 77</td>
<td>0.0</td>
<td>79.9</td>
</tr>
<tr>
<td>7</td>
<td>38.4 14.0-83.6 75</td>
<td>4 2008 60 48</td>
<td>96.3</td>
<td>122.2</td>
</tr>
<tr>
<td>8</td>
<td>49.3 32.8-71.3 261</td>
<td>6 N/A 33 31</td>
<td>73.2</td>
<td>60.4*</td>
</tr>
<tr>
<td>9</td>
<td>41.9 21.6-73.2 137</td>
<td>3.5 2008 60 59</td>
<td>142.1</td>
<td>83.0</td>
</tr>
<tr>
<td>10</td>
<td>47.3 25.2-80.9 129</td>
<td>3 2011 40 36</td>
<td>46.7</td>
<td>70.1</td>
</tr>
<tr>
<td>Study av.</td>
<td>40.2 - 130 3.8 -</td>
<td>60 57 76.8</td>
<td>79.2</td>
<td></td>
</tr>
<tr>
<td>Non-study av.</td>
<td>100.1 - 117 -</td>
<td>47 109.9 98.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 10.5: Summary characteristics and LLoS results of participating units.
RR - Relative Risk; CI - 95% Confidence Intervals; n - total cases; Cons. FTE - number of colorectal consultants in Full Time Equivalents; ERP - year Enhanced Recovery Program introduced; Cons. est. - consultant surgeon estimate; HES - derived from Hospital Episode Statistics; * - mortality or readmission rates significantly below 100 at p<0.05.
components of this care package. Accordingly, the units standardised a number of clinical processes relating to the pre-, intra- and post-operative phases of the patient journey. This study was not designed to comprehensively assess all details of patient care. Nonetheless, extensive preoperative counselling, and early postoperative mobilisation and oral intake were of clear importance across all study units.

I think it’s [the ERP is] great for the nursing staff, because they know where they are, [and] what they need to be doing with their patients, which is really good.

Site 6, nurse

I think it [the ERP] works really well . . . because it helps every single member of staff and it helps the patient because . . . they know what is expected of them . . .

Site 1, nurse

. . . It’s been our experience that one of the big barriers for patients going through an Enhanced Recovery Program is pre-defined concepts about what they’re going through, and it’s impossible to have a colonic resection and be out of hospital within two days, ‘My dad was in for a week,’ etc. So we’re very keen to sow the seeds of early discharge, if all goes well, right from the start.

Site 9, surgeon
While practice within units was largely uniform, some areas of care demonstrated wide variation between different units. Consultant-estimated laparoscopy rates ranged from 33% to 85%, and postoperative analgesic strategies varied from routine epidural usage, to active avoidance of epidurals, use of local anaesthetic infiltration devices, or patient controlled opiate analgesia.

**Organise team to deliver care reliably and efficiently**

Across the study units, consultants and nurses adopted a lead role in direct care provision on a day-to-day basis. There was less reliance on medical trainees to lead clinical care. Appropriately trained nurses undertook preoperative patient counselling and risk assessment, and, in some hospitals, led postoperative ward care.

...The Enhanced Recovery ...gives us the wherewithal to be a sort of nurse-led sort of protocol. We have a protocol set by all three consultants, but we sort of dictate when lines are taken out, when people start to drink, etc, etc.

Site 1, nurse

Nurses are far better at following protocol than doctors are, and they tend to react better...doctors think they know better, if you see what I mean? So if you have nurse-led discharge, if you have nurse-led Enhanced Recovery Programs, then things tend to happen in a much more structured fashion than if you leave it up to the doctors.

Site 2, surgeon
Such nurse-led care required close consultant support. Other hospitals described how consultant presence on the ward was required to maintain efficient patient flow along the pathway.

...The issue I have with the ERP ...[is] unless we have a consultant actually seeing the patient daily and driving it forward, they [patients] tend to get delayed in their discharge. So, actually, it’s consultant-based care and ward management.

Site 3, surgeon

Besides the allocation of clinical tasks to specific members of the team, the majority of hospitals in the study organised patients on wards by speciality or urgency of admission.

Monitor and respond to deviations from the norm

At every hospital site, nurses used an observation-based early warning score system, with associated protocols for escalation of care, to detect deterioration in their patients. Beyond this, a number of nurses described a more sophisticated approach, combining patients’ symptoms and signs with previous clinical experience, to help detect complications before physiological deterioration.

...From an experienced nurse perspective, you just know when there is something wrong. ...You don’t have to particularly look at the NEWS [early warning score] shot, you can look at a patient and sort of see clinically they’re not quite right. ... Then obviously we’d inform
whoever we felt was appropriate depending upon what their NEWS result was.

Site 1, nurse

In this Trust we have a PAR [early warning score] scoring system, so all our patients have four-hourly observations performed. Within that we also have a two-hourly rounding program. ... So it’s normally either within the observation round or through the two-hourly rounding that problems can arise, for example, if someone hasn’t passed wind, hasn’t been out to the toilet post bowel surgery, that would be triggered when you’re asking your rounding questions. So it’s normally at that point that I write, ‘He might have an ileus.’ You’d let the team know, listen to bowel sounds, etc.

Site 4, nurse

The nursing staff will communicate any concerns, worries, directly to the consultant ... Most importantly, they don’t have to reach physiological parameters for a nurse to be allowed to phone the consultant, in the hope that the nursing staff spot that someone is unwell before they become physiologically unwell.

Site 5, surgeon

Across the studied units, several nurses described reporting concerns directly to consultants, bypassing more traditional patterns of escalation within the medical team.
We’ve got the Early Warning Score that we use, so if a patient scores anything above a three ... we escalate it, and that goes to the F1, reg [registrar]. But, with the Enhanced Recovery, we tend to side-step them and go straight to the consultant.

Site 5, nurse

If there’s any deviation [postoperatively] or things aren’t going as planned, then obviously the relevant consultants tend to get called and phoned anyway, as a rule.

Site 4, surgeon

... We all work together as a team, and any member of my staff could contact any of the consultants at home and would feel happy to do that. And I would do that without hesitation if I felt that a patient was being mis-managed, and I have done in the past.

Site 1, nurse

This relied on excellent relationships between nurses and consultants, with open communication and respect for nurses’ opinions and judgements.

The study sites also reported frequent senior medical input. Most consultants saw their patients from 2 to 3 times per week, to every day. In all but one site, the daily ward round was conducted by at least a registrar, if not a consultant. Some sites reported routine review two or three times per day.

Resources for responding to complications differed between the study sites. All had access to emergency theatres, intensive care and interventional radiology,
but the level of support varied. Emergency theatres were often shared between specialties. A number of interviewees commented that while the intensive care team were very supportive and helpful, they were often stretched. Some sites had 24-hour access to interventional radiology through an on-call system, while others only had routine access during normal working hours.

10.5 Discussion

This study was designed to try and identify key structures and processes in the organisation of care within a colorectal unit, that determined whether it achieved high or low postoperative length of stay. There were unexpected problems with recruiting study sites. High length of stay units were significantly less likely to participate than short stay units. However, a good sample of short stay units was achieved, with all 10 units approached agreeing to participate. Qualitative analysis of data from these 10 units yielded important insights into how excellent length of stay results may be delivered. The next paragraph reviews the key features of clinical practice, before subsequent discussion of care organisation among these units.

In this study, there were a number of areas of clinical practice that were common across the 10 units with the shortest length of stay in the country. All of these units based perioperative care upon the ERP. The most salient features of the ERP reported by interviewees were careful and early setting of patient expectations before surgery, as well as early oral intake and early mobilisation afterwards. During their postoperative stay, patients were regularly reviewed by senior medical staff. This involved daily registrar review as a minimum, with consultant review from
2-3 times per week to every day. In addition, all units used an early warning score system to support the detection of postoperative complications. Besides these common areas, there were some notable variations in clinical practice between the included units. Laparoscopy rates varied from 31% to 85%, and the approach to postoperative analgesia also varied widely.

Beyond these immediate features of clinical practice, there were a number of areas of care organisation that were common to the high performing units studied. Individual units paid careful attention to the delivery of care in accordance with the ERP, to make sure care was provided as intended. This was achieved either through consultant- or nurse-led and delivered care, in daily practice on the ward. Without this, it seemed that delays could occur during the various phases of postoperative recovery, resulting in increasing length of stay. In addition, nurses were empowered to exercise their clinical judgement to identify problems early, rather than relying solely on the early warning score system. This was facilitated by strong working relationships with consultants, and close, direct communication between consultants and ward nurses.

These findings suggest that while certain components of clinical care (such as the ERP) may be pre-requisites for achieving excellent length of stay results, others (such as laparoscopy and analgesic strategy) may perhaps be less important than previously considered. Beyond these specific aspects of clinical care, the study units demonstrated many shared features in the organisation and delivery of care. Across the 10 short stay units, the ERP was implemented with careful consideration and active day-to-day progression on the ward. The study units also described close monitoring for postoperative problems, and excellent interprofessional communication. In these ways, the organisation of care seemed critical
in achieving excellent length of stay among the study units.

This study has a number of specific strengths and weaknesses. The pros and cons of HES data, upon which the selection of units for study was based, have previously been discussed in section 5.2.3 of chapter 5. Notable strengths of this study include the conduct of all interviews by a surgical trainee. This helped establish rapport with interviewees, who provided rich opinions in relatively short interviews. In addition, double-coding of the first 5 interviews with a non-clinical health care researcher, who has extensive experience with qualitative methods, ensured the rigor of this analysis.

The study also has a number of limitations. The units included represent a small sample of the population of colorectal units, and further work is planned to develop and confirm these initial findings in units across a range of performance. Interview data may be subject to social desirability response bias, where participants alter how they represent their work in response to the social pressures and context of such an interview. While interviewees were not informed of their organisation’s performance, their description of local practice may have been positively biased. Only one surgeon and one nurse were interviewed at each site; different data may have been gathered from other members of the same organisations. Lastly, the data collected in these interviews will have been influenced by the themes presented in the semi-structured interview protocol. Double coding with an independent researcher helped ensure that the analysis was appropriately grounded in the data collected, but it is possible that alternative themes may have emerged from open interviews.

Few studies have used similar techniques to understand high performing institutions. Bradley et al employed their positive deviance approach to improve
door-to-balloon times for patients admitted with acute ST-elevation myocardial
infarction\textsuperscript{233,234}. Their qualitative methods were more rigorous, including 11 site
visits and 122 interviews, yet the themes identified were similar to the present
study. For example, 2 themes related to standardised protocols, and one to col-
laborative teamwork\textsuperscript{234}. A smaller study of stroke thrombolysis identified similar
themes\textsuperscript{241}. Comparable results between these earlier studies and the presented
work support the validity of the methodology, data and analysis used in this study.
Semi-structured telephone interviews represent a cheaper and quicker source of
data than site visits and in-depth interviews.

This study highlights the importance of organisational factors in determining
high performance. The clinical care provided by the study institutions was founded
upon well-known, evidence-based practices within colorectal surgery. The ERP
was necessary but not sufficient to achieve the best results. Beyond the ERP,
the units all carefully managed how this care package was actually delivered and
implemented.

\ldots All the obvious stuff for this [reducing length of stay] has been said
and proven a million times before. Enhanced Recovery does work,
laparoscopic surgery does work, goal-driven senior management does
work \ldots so all these things are fairly obvious \ldots Next stage I think now
is to get people to actually follow the things that you know work.

Site 2, surgeon

In this study, ward nurses had a key role in complication detection and manage-
ment. Excellent inter-professional communication was developed through frequent
and senior medical presence on the ward. Strong relationships helped empower nurses to seek definitive senior assessment and treatment for their patients early, bypassing intermediary assessments by more junior team members. The important contribution of nurses in these areas is supported by meta-analytic evidence of an association between increased registered nurse staffing and lower mortality and failure to rescue rates in surgical patients. 

Increasing reliance on consultants and nurses to deliver care represents a paradigm shift in the organisation of the clinical team. While this may be an effective strategy to improve patient care, it may meet resistance from certain stakeholders. Greater direct care delivery by consultants and nurses would need appropriate support and resources. Reduced responsibility for clinical care among trainees may delay the acquisition of the clinical skills necessary to become a consultant. However, greater contact between consultants and trainees may conversely be considered an opportunity to improve skills transfer. Ultimately, it perhaps needs to be remembered that the purpose of health care is to meet patient’s needs and care should be organised accordingly within the resources available.

This study highlights the merits of exploring novel or under-utilised methods for the study of health care performance. While much has been learned by analysing poor institutional performance, the recurrence of scandals within the NHS, such as the Bristol Royal Infirmary Inquiry and Mid Staffordshire Inquiries, suggests that this approach has not delivered the intended improvements in care. Perhaps change may be achieved by providing a positive vision for health care, using lessons learned from peers who have achieved excellent results within the same system. Future research using similar methods may help clarify whether similar factors are associated with high performance defined using other outcomes, and in other fields.
of medicine.

The colorectal units in this study achieved excellent length of stay results, founded upon the ERP. However, an ERP alone may not be enough. The majority of direct clinical care was delivered and led by consultants and nurses, resulting in efficient delivery of the ERP and excellent unit-level outcomes. Well-supported nursing staff and frequent medical review helped identify and resolve care problems promptly. Laparoscopy rates were not significantly different from the national average. An appropriately supported shift in responsibility for care delivery towards nurses and consultants, and away from trainees, may be an effective strategy for the improvement of colorectal outcomes, though further work is required to generalise the study findings more widely.

10.6 Chapter summary

This chapter has described the development of a novel semi-structured interview, designed to try and better understand, in practical terms, how colorectal units are organised to achieve high performance. There were difficulties engaging Trusts and units with the lowest levels of performance, even though they were not informed of their outlier status. In marked contrast, all 10 high performing units participated, and the data from interviews with these units was analysed qualitatively. The results highlight the importance of careful implementation and monitoring of best practice, although work is needed to establish the generalisation of these findings to the broader population of colorectal units in England.
Chapter 11

Study 6: Understanding length of stay in elective colorectal surgery - exploratory quantitative analysis

11.1 Chapter overview

This chapter presents an exploratory quantitative analysis of all interviews conducted in study 6.

11.2 Introduction

The overall objective of this work is to provide a broad, generalised understanding of the determinants of unit-level performance in colorectal surgery, underpinned by a quantitative association between care organisation and clinical outcomes. Previous research in this area was discussed at some length in the introduction section of the previous chapter. The studies most pertinent to the present chapter will be summarised here.

In 1997, Daley et al published a study within the VA system that examined for an association between seven quality of care domains and risk-adjusted mortality
and morbidity\textsuperscript{32}. When high and low outlier units were compared, units with the lowest morbidity and mortality had higher technology and equipment scores, and higher overall quality of care ratings. There was no significant association between performance and the other 6 domains examined. Published in 2007, a later questionnaire study, also among VA hospitals, showed a number of univariate associations between morbidity and questionnaire-based measures of structures and processes\textsuperscript{31}. Low morbidity units had fewer general surgery operating rooms, lower reports of being short staffed for nurses or aides, and better continuity of perioperative anaesthetic care. Multivariate analysis revealed university affiliation as a dominant factor, significantly associated with increased postoperative morbidity\textsuperscript{30}.

Outside the surgical literature, Bradley et al demonstrated an association between questionnaire-based measures of care processes and lower door-to-balloon times for patients undergoing intervention for a heart attack\textsuperscript{232}. Better performance was associated with simplifying the process of activating the team, allowing pre-hospital activation, and real-time feedback on performance.

While these studies have made an important contribution to the literature, it is still not clear how a surgical unit should organise and deliver care to achieve the best results. Association between university affiliation or technology scores and clinical outcomes does not directly explain what high performing units actually do differently, compared with their peers. Further work is required to define the underlying reasons behind such associations. A deeper understanding of these relationships may help deliver greater improvements in care quality for all.

The previous chapter described the development and use of a semi-structured telephone interview, designed to better understand how colorectal units achieve high performance. The present chapter presents an exploratory analysis of the
quantitative data obtained from all interviews conducted with 15 colorectal units, 10 with short length of stay and 5 with long stay. This analysis is intended to help determine whether the interview protocol may be used to help understand the determinants of the results of care within a colorectal unit across the entire spectrum of performance seen.

**Study question**

- Can the key structures and processes in the organisation and delivery of care within a colorectal unit be measured quantitatively with a semi-structured interview?
- How do the measured structures and processes among these units relate to unit-level performance?

**11.3 Methods**

The methods and sampling strategy used to collect the data for the present analysis were described in chapter 10. The statistical analysis of interview data is presented within this section.

**11.3.1 Statistical analysis**

The semi-structured interview protocol, provided in appendix G, contained 11 scored questions. Each question was accompanied by a 7-point ordinal rating system. Points 1, 3, 5 and 7 were anchored using exemplar responses. This data was collected for all interviews conducted.
Descriptive statistics and the distribution of individual and total question scores were examined. Distributions were assessed both graphically, using frequency histograms, and statistically, using the Shaprio-Wilk test for normality. The distribution of responses informed the subsequent use of parametric or non-parametric statistical tests.

Two principal subgroup analyses were performed. Interview responses were examined separately by professional group, and by performance of the local unit. Comparisons between these subgroups were made with parametric or non-parametric tests as appropriate. A final subgroup analysis, examining scores by both profession and unit performance was performed, with descriptive results and an assessment of the correlation between doctors’ and nurses’ scores at units with short or long stay.

For all tests, a $p$-value of $<0.05$ was considered statistically significant.

11.4 Results

One nurse and one consultant surgeon were interviewed at each of the 15 participating units, resulting in a total of 30 interviews. It is important to remember that higher interview scores were attached to characteristics of care organisation considered desirable \textit{a priori}. This results section will begin by examining the distribution of interview scores and some descriptive statistics, before presenting the subgroup analysis performed.
11.4.1 Distribution of scores

Frequency histograms for individual question responses and the interview total are provided in appendix H, with most appearing to show an approximately normal distribution. Assessment with the Shapiro-Wilk test suggested that only the responses to the question on safety attitudes and the overall total were normally distributed ($p=0.099$ and $p=0.500$, respectively). However, the assessment of the distribution of individual question results should be considered in light of the small sample size and discrete nature of the data. Given these results, both parametric and non-parametric descriptive data are reported in each section.

11.4.2 Overall scores

A summary of individual question and total scores for all interviews conducted is provided in table 11.1. The range of responses received varied by question. Variation was most limited for the question focused on preoperative preparation of patients for surgery. Variation was greatest for the question that examined the leadership within the colorectal department. The average total score was 54.3, of a possible maximum of 77, with a standard deviation of 6.06.

11.4.3 Scores by professional group

Descriptive statistics for question responses among doctors and nurses that took part in the study are provided in table 11.2. There was no statistically significant difference in the total interview score between doctors and nurses at individual institutions (paired samples T test $p = 0.164$). Although not statistically significant, the mean interview total was higher (suggestive of reporting better care organisa-
Interview quantitative analysis

<table>
<thead>
<tr>
<th>Question theme</th>
<th>Parametric</th>
<th>Non-parametric</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Level and quality of equipment</td>
<td>5.10</td>
<td>5</td>
</tr>
<tr>
<td>Level of medical and nursing staff</td>
<td>3.47</td>
<td>3</td>
</tr>
<tr>
<td>Preoperative preparation of patients</td>
<td>6.20</td>
<td>6</td>
</tr>
<tr>
<td>Operative management</td>
<td>4.70</td>
<td>5</td>
</tr>
<tr>
<td>Routine postoperative management</td>
<td>5.70</td>
<td>6</td>
</tr>
<tr>
<td>Organisation to detect postoperative problems</td>
<td>5.37</td>
<td>5</td>
</tr>
<tr>
<td>Standardisation of care</td>
<td>5.13</td>
<td>5</td>
</tr>
<tr>
<td>Interprofessional communication and coordination</td>
<td>5.03</td>
<td>5</td>
</tr>
<tr>
<td>Clinical leadership and culture</td>
<td>4.40</td>
<td>5</td>
</tr>
<tr>
<td>Attitudes to safety and adverse incidents</td>
<td>4.50</td>
<td>4</td>
</tr>
<tr>
<td>Collection and feedback of performance data</td>
<td>4.70</td>
<td>5</td>
</tr>
<tr>
<td>Total</td>
<td>54.3</td>
<td>54</td>
</tr>
</tbody>
</table>

Table 11.1: Parametric and non-parametric descriptive statistics for interview scores.
SD - standard deviation; IQR - interquartile range.

The relationship between the paired interview scores across different units was examined. There was weak, non-significant correlation between doctors’ and nurses’ total scores (Pearson’s correlation coefficient, $r = 0.204$; $p = 0.465$). A scatter plot of doctors’ versus nurses’ total interview scores for all 15 units interviewed is shown in figure 11.1. Within the study sample, if a doctor’s interview returned a higher than average score, suggesting reporting of better practice, the associated interview with a nurse in the same unit was not likely to be significantly above average.

11.4.4 Scores by performance status

Interview question scores and the overall total for units with short and long stay are provided in table 11.3. There was no statistically significant difference in score
<table>
<thead>
<tr>
<th>Question theme</th>
<th>Doctors</th>
<th>Nurses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parametric</td>
<td>Non-parametric</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Level and quality of equipment</td>
<td>5.33</td>
<td>1.40</td>
</tr>
<tr>
<td>Level of medical and nursing staff</td>
<td>3.13</td>
<td>1.19</td>
</tr>
<tr>
<td>Preoperative preparation of patients</td>
<td>6.47</td>
<td>0.52</td>
</tr>
<tr>
<td>Operative management</td>
<td>5.13</td>
<td>0.99</td>
</tr>
<tr>
<td>Routine postoperative management</td>
<td>5.87</td>
<td>0.83</td>
</tr>
<tr>
<td>Organisation to detect postoperative problems</td>
<td>5.67</td>
<td>0.62</td>
</tr>
<tr>
<td>Standardisation of care</td>
<td>5.53</td>
<td>0.83</td>
</tr>
<tr>
<td>Interprofessional communication and coordination</td>
<td>4.87</td>
<td>1.69</td>
</tr>
<tr>
<td>Clinical leadership and culture</td>
<td>4.40</td>
<td>1.64</td>
</tr>
<tr>
<td>Attitudes to safety and adverse incidents</td>
<td>4.13</td>
<td>1.36</td>
</tr>
<tr>
<td>Collection and feedback of performance data</td>
<td>5.20</td>
<td>0.78</td>
</tr>
<tr>
<td>Total</td>
<td>55.7</td>
<td>5.80</td>
</tr>
</tbody>
</table>

Table 11.2: Parametric and non-parametric descriptive statistics for interview scores for surgeons and nurses. SD - standard deviation; IQR - interquartile range.
between units with short or long stay (independent samples T test, $p = 0.591$). With this in mind, there was a tendency for units with short length of stay to have lower overall interview totals, suggestive of poorer care organisation, than those with long stay (53.9 vs 55.2, respectively).

### 11.4.5 Scores by profession and performance

The average total interview score for doctors and nurses at short and long stay units is provided in table 11.4. Given the small sample size, with only 10 short and 5 long stay units included, statistical significance tests were not performed. Among surgeons interviewed, the overall interview total was lower and suggested reporting of poorer care organisation in short stay units than in long stay units. This may reflect genuine differences in care organisation, or perhaps consultants
<table>
<thead>
<tr>
<th>Question theme</th>
<th>Short stay</th>
<th></th>
<th>Long stay</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parametric</td>
<td>Non-parametric</td>
<td>Parametric</td>
<td>Non-parametric</td>
</tr>
<tr>
<td></td>
<td>Mean  SD</td>
<td>Median  IQR   Range</td>
<td>Mean  SD</td>
<td>Median  IQR   Range</td>
</tr>
<tr>
<td>Level and quality of equipment</td>
<td>4.95 1.40</td>
<td>5  3  2-7</td>
<td>5.40 1.17</td>
<td>5.5  2  4-7</td>
</tr>
<tr>
<td>Level of medical and nursing staff</td>
<td>2.71 1.26</td>
<td>3  2  2-6</td>
<td>3.80 1.48</td>
<td>3.5  3  2-6</td>
</tr>
<tr>
<td>Preoperative preparation of patients</td>
<td>6.35 0.59</td>
<td>6  1  5-7</td>
<td>5.90 0.57</td>
<td>6  0  5-7</td>
</tr>
<tr>
<td>Operative management</td>
<td>4.95 1.05</td>
<td>5  2  2-7</td>
<td>4.20 1.40</td>
<td>4  3  2-6</td>
</tr>
<tr>
<td>Routine postoperative management</td>
<td>5.75 0.91</td>
<td>6  1  4-7</td>
<td>5.60 1.08</td>
<td>6  1  3-7</td>
</tr>
<tr>
<td>Organisation to detect postoperative problems</td>
<td>5.35 0.81</td>
<td>5  1  3-7</td>
<td>5.40 0.52</td>
<td>5  1  5-6</td>
</tr>
<tr>
<td>Standardisation of care</td>
<td>4.67 0.91</td>
<td>5  1  3-7</td>
<td>5.20 1.32</td>
<td>5.5  2  3-7</td>
</tr>
<tr>
<td>Interprofessional communication and coordination</td>
<td>4.75 1.68</td>
<td>5  3  2-7</td>
<td>5.60 1.17</td>
<td>6  1  3-7</td>
</tr>
<tr>
<td>Clinical leadership and culture</td>
<td>4.15 1.50</td>
<td>4.5  2  1-7</td>
<td>4.90 1.45</td>
<td>5  2  2-7</td>
</tr>
<tr>
<td>Attitudes to safety and adverse incidents</td>
<td>4.50 1.36</td>
<td>4  2  2-7</td>
<td>4.50 1.43</td>
<td>5  3  2-6</td>
</tr>
<tr>
<td>Collection and feedback of performance data</td>
<td>4.70 1.38</td>
<td>5  2  2-7</td>
<td>4.70 1.25</td>
<td>5  2  2-6</td>
</tr>
<tr>
<td>Total</td>
<td>53.9 5.91</td>
<td>53  7  43-68</td>
<td>55.2 6.58</td>
<td>54.5  9  43-66</td>
</tr>
</tbody>
</table>

Table 11.3: Parametric and non-parametric descriptive statistics for interview scores for short and long stay units. SD - standard deviation; IQR - interquartile range.
in high performing units have higher expectations and are more critical of local practice, resulting in lower scores based on reporting of practice in the interviews. Conversely, the total score from interviews with nurses was higher in short stay units than those with long stay, although the absolute difference was small.

<table>
<thead>
<tr>
<th>Professional group</th>
<th>Length of stay</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Short</td>
<td>Long</td>
</tr>
<tr>
<td>Doctor</td>
<td>54.5</td>
<td>58.2</td>
</tr>
<tr>
<td>Nurse</td>
<td>53.2</td>
<td>52.2</td>
</tr>
<tr>
<td>Overall</td>
<td>53.9</td>
<td>55.2</td>
</tr>
</tbody>
</table>

Table 11.4: Average total interview scores by professional group and unit performance status. Mean values provided.

The association between surgeons’ and nurses’ total scores at units stratified by performance was also examined. Pearson’s correlation coefficient was higher among units with short stay compared with long stay units (r = 0.311 and r = 0.092, respectively). However, this correlation was non-significant in both cases (p=0.381 and p=0.883, respectively). This finding may suggest that doctors and nurses at high performing units have more closely aligned views on the organisation and delivery of care within the unit, perhaps reflecting a more shared mental model of local clinical practice. However, given the small sample size, this must be interpreted as a preliminary finding, requiring further investigation in future work.
11.5 Discussion

The novel, semi-structured interview developed for this study yielded an overall score that was normally distributed and had an average of 54.3, with a standard deviation of 6.06. Total scores ranged from 43 to 68, out of a maximum possible score for the interview of 77. There was no significant difference in total interview scores between surgeons and nurses, or between units with short or long stay.

The highest individual question scores, indicative of better reported practice, were returned for questions on preoperative preparation of patients, routine postoperative management and organisation of care to detect postoperative problems. Responses to these questions also had some of the lowest variation observed. These results suggest that at the units studied, patients usually went through an effective preoperative assessment clinic and were well counselled about their surgery beforehand. Patients were normally seen regularly by consultants after their surgery, often within an enhanced recovery setting with goals for postoperative care. Complications were detected through regular review and early warning score systems, with good levels of supporting services such as interventional radiology and intensive care to manage complications when they arose.

By far the lowest score was returned for interviewees’ ratings of the level of medical and nursing staff, suggesting generally poor reported staffing levels, frequent turnover and regular use of agency or locum staff. However, variation in scores for this question was greater than for the questions discussed above.

At the individual unit level, there was only weak correlation between doctors’ and nurses’ total interview scores, and this did not reach statistical significance. An interesting subsidiary finding was that there appeared to be closer correlation
between surgeons’ and nurses’ total scores at units with short stay, than those with long stay, although neither correlation was significant. It is important to remember that the sample size for this study was small. This will have increased the chance that statistical tests have failed to show a significant difference when one genuinely exists within the true population (type II error).

The ability of this study to detect variation in the structures and processes of care may have been compromised not just by the small sample size, but also by the lack of participation among long stay units. Analysis of participation by length of stay in the previous chapter showed that long stay units were statistically significantly less likely to participate. Those that did agree to take part may not have been representative of all units with long stay. Indeed, 3 of the 5 long stay units included reported having tertiary colorectal practice. The complexity of cases in these units may have been an important confounding variable that significantly influenced results. Had it been possible to include a measure of this complexity in the risk-adjustment model, perhaps these units would not have been identified as outliers. Notwithstanding this observation, these 5 long stay units self-selected to participate in this study; therefore, they may have been more quality conscious and externally focused than their peers, perhaps making them more similar to the short stay units included. This could have reduced measurable differences in care organisation between short and long stay groups, compounding the lack of power due to the small sample size.

There are a number of important limitations that affected this study, many of which have been discussed in the previous chapter. Important limitations will be revisited here. Besides the small sample, only 2 staff members were interviewed at each site. While these participants were chosen to have good working knowledge of
the care delivered, as well as the wider functioning of the department and organisation, interviewees were selected by local members of staff, a process that could have introduced unknown bias. Interviewing a greater number of staff members may have resulted in a more accurate assessment of practice in included hospitals. The interview tool was novel and has not been validated. It is possible that it lacks sensitivity to measure genuine differences in care organisation. The study may also have been limited by unconscious bias introduced by the interviewer, which may have influenced coding of responses into numerical scores.

Beyond these limitations, this study has demonstrated the feasibility of using semi-structured telephone interviews to collect data on the structures and process of surgical care that are not available by other means. Members of the clinical team were willing to engage in discussion about the local organisation of care with a clinician researcher, building rapport quickly, enabling them to make critical comments about their unit during the relatively brief interviews conducted. The interviews were able to yield both qualitative and quantitative data. The scores for the interviews conducted appeared to show a reasonable degree of variation. The mean score was 54.3, with a standard deviation of 6.06. In the context of a flawed sample that may not have included units with the poorest organisation of care, this may suggest that the interview tool has the potential to measure variation in care organisation and delivery. However, much further work would be required with a larger sample of units to determine if the interview tool has reliability and validity as a measurement instrument.

Although not statistically significant, the average interview score for surgeons was higher than that of nurses that took part. This is consistent with other research examining teamwork in the operating theatre\textsuperscript{242,243} and on the ward\textsuperscript{244,245}, showing
that medical staff tend to report a more favourable assessment of teamwork and communication than nursing staff. It is also interesting that there appeared to be closer correlation of interview totals between surgeons and nurses at high performing units. While the correlation was non-significant, it should be considered in light of the qualitative findings reported in the previous chapter, documenting the excellent communication between surgeons and nurses at high performing units. This may suggest that a more shared understanding of local practice is important in achieving the best results. Clearly, further work is needed to explore this area in more detail.

While the quantitative analysis was limited by the sampling problems encountered, the semi-structured tool showed some degree of variation among the study units, suggesting that with appropriate refinement and a more rigorous recruitment strategy, it may represent a useful way of assessing the organisation of care within a colorectal unit. However, further work is required to identify and overcome barriers to participation among units with low performance if a deeper understanding is to be achieved.

11.6 Chapter summary

This chapter has described the exploratory quantitative analysis of a novel, semi-structured interview to examine care organisation in colorectal units. Difficulties engaging with poor performing units may have compromised the ability of the interview tool to measure significant variation in care. Much further work would be required to determine the potential of the interview as a reliable and valid measure of the organisation of care within a colorectal unit. Areas for further
research have been identified, such as exploring the barriers to participation in this type of study among units with long stay, and investigating the relationship between nurses’ and doctors’ scores at high and low performing units. Further work using a national questionnaire is currently ongoing, to try and build upon the qualitative and quantitative results of this study.
Chapter 12

Study 6: Post hoc analysis - comparison of participant and non-participant units using existing data

12.1 Chapter overview

This short chapter presents the results of a post hoc analysis conducted to compare the characteristics of units that did or did not participate in study 6.

12.2 Introduction

Previous research has suggested that surgical units and consultants that engage in voluntary disease registers have better outcomes than those that do not\textsuperscript{160,161}. When approaching Trusts to participate in the present interview study, the responses received from the selected Trusts varied greatly. Some were keen and enthusiastic, promptly approved the study and helped to coordinate the interviews. Others seemed disengaged and disorganised. After completing the inter-
views, comparison of participation by length of stay performance showed that those units that did not participate were more likely to have poor performance (see table 10.4). It was therefore decided to explore this difference between units stratified by participation status, using the publicly available data collected for study 5.

**Study question**

- Can existing data in the NHS be used to define measurable structure, process and outcome differences between colorectal units that did or did not participate in the present interview study?

**12.3 Methods**

The identification of units and collection of data for this analysis are described in study 5 (see chapter 9). The 2 additional units approached to participate in study 6 are also included. In contrast with study 5, units were stratified by whether or not they took part in study 6, rather than by their length of stay. Data were compared between these two groups using the independent samples Mann-Whitney U test. Due to the small sample size, $p<0.10$ was considered statistically significant.

**12.4 Results**

Data items were compared across all 22 units approached. 9 items were significantly different between participating and non-participating units (see table 12.1). In each case, the difference suggested better quality of care in participating units and Trusts, compared with non-participants. For example, participating units had
lower rates of cancellation of elective surgery per bed, and lower rates of staff reporting work-related stress in the preceding 12 months. Participating units also had higher ratings in various patient surveys, including the Friends and Family Test, Inpatient survey, and Cancer survey.

<table>
<thead>
<tr>
<th>Data item</th>
<th>Participant?</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Elective operations cancelled per bed</td>
<td>0.54</td>
<td>0.79</td>
</tr>
<tr>
<td>C difficile Trust apportioned rates</td>
<td>15.3</td>
<td>18.8</td>
</tr>
<tr>
<td>Staff survey: reported work related stress (%)</td>
<td>36.0</td>
<td>38.0</td>
</tr>
<tr>
<td>Friends and Family Test score (average)</td>
<td>76.7</td>
<td>70.3</td>
</tr>
<tr>
<td>Inpatient survey: doctors rating</td>
<td>8.6</td>
<td>8.3</td>
</tr>
<tr>
<td>Inpatient survey: nurses rating</td>
<td>8.4</td>
<td>8.1</td>
</tr>
<tr>
<td>Inpatient survey: operations and procedures rating</td>
<td>8.4</td>
<td>8.2</td>
</tr>
<tr>
<td>Cancer survey: involvement in decisions</td>
<td>73.0</td>
<td>70.0</td>
</tr>
<tr>
<td>Cancer survey: explanation of operation</td>
<td>78.0</td>
<td>71.0</td>
</tr>
</tbody>
</table>

Table 12.1: Significant differences between participating and non-participating units.
Median values provided.

12.5 Discussion

In marked contrast with the findings of study 5, when units were stratified by participation in the current interview study, there were 9 statistically significant differences between participating and non-participating colorectal units. With 56 comparisons at the 10% significance level, more differences were identified than would be expected to occur due to chance alone. In each case, the data suggested better performance or care quality in participating units. Therefore, willingness to engage in quality improvement research may be an important facet of high performing units; or, considered conversely, failure to engage in this type of work
may be associated with poorer quality of care and organisational performance.

This analysis highlights the challenge of trying to learn from units across the full range of performance, as those with the lowest level of performance may be particularly difficult to reach. Close examination of the care in such units may provide important insights into understanding surgical performance, and these clinical teams and organisations may stand to benefit the most from quality improvement research. Qualitative work with units that do and do not participate in this type of work may help identify important barriers to engagement, and help develop strategies to overcome any barriers identified.

12.6 Chapter summary

This chapter has presented a brief, post hoc analysis exploring differences between units that did and did not participate in this interview study. It has highlighted a challenge that has important implications for health services research. Strategies need to be developed to engage and learn from all units, not just those who are willing to participate in quality improvement work. A better understanding of the reasons for a lack of engagement may also help guide the implementation of research findings, to improve care for all patients treated within a health care system such as the NHS.
Part V

Discussion and conclusions
Section summary

This final section presents a broad discussion of the studies contained in this thesis, the limitations associated with the methodologies used, and the implications of the findings. It draws together all the preceding work and how it has helped further the present understanding of unit-level performance in colorectal surgery.
Chapter 13

Discussion

13.1 Chapter overview

This final chapter reviews the key findings and limitations of the original research presented in this thesis. The implications of this work are discussed, including directions for future research, before making some concluding remarks.

13.2 Key findings

This thesis has combined various research techniques, including questionnaires, interviews and large database work, to further the current understanding of department-level performance in colorectal surgery.

The first study, a patient survey, found that patients treated surgically for gastrointestinal tract cancer in the NHS experienced very little involvement in provider choice. However, the vast majority of these patients expressed that they would like to be involved in such decisions in some way, with over two-thirds indicating that they would have preferred more active involvement in choosing their provider than they had experienced. Participants attached the highest importance to provider-level information on wait times for cancer treatment, surgical errors, operative volume, and mortality rates, as well as hospital infection rates. While
further work is required to develop these findings, for example by exploring the information patients use to make decisions, they suggest that a number of the areas on which surgical quality improvement research has focused are also important for patients.

Studies 2 to 4 used routine administrative data from the HES database to describe and explore the mortality results after colorectal surgery over time and in various settings. Study 2 demonstrated that between 1998 and 2012, mortality after a colorectal resection for cancer fell significantly for nearly all patient groups examined, regardless of the urgency of the operation, surgical approach, screening eligibility or major complication status. The findings of this study may be interpreted to show that the observed falls in mortality probably relate to general improvements in care quality, rather than specific technical innovations such as the widespread adoption of laparoscopic surgery. The outcomes of care for surgical patients are likely to depend upon a wide range of aspects of medical care, in its broadest sense. These may include appropriate detection and management of comorbidities, high quality imaging and endoscopy, effective oncological treatments, and good quality surgical, anaesthetic and critical care. Incremental improvements across all of these areas may have combined to drive a reduction in postoperative mortality for nearly all patient groups undergoing colorectal surgery in this study.

Study 3 examined the impact of lengthening postoperative follow-up on the identification of institutional mortality outliers after colorectal surgery. In the cohort studied, the results showed that 90-day follow-up successfully identified early high mortality outliers at 30 days, while also identifying other units that developed high risk-adjusted mortality rates over a longer postoperative period. If it is accepted that high outlying mortality rates may be a marker of potential problems
with quality of care, it may be preferable to extend follow-up for examining surgical mortality to 90 days. Clearly there is a balance to be made between short-term outcomes, where the temporal link with surgical care may be strongest, and longer follow-up, which may allow better detection of delayed complications, but can also be influenced by other, non-surgical factors.

The fourth study explored whether individual colorectal units achieved similar relative performance for patients undergoing surgery on both an elective and non-elective basis. Outlier status was not consistent across these patient cohorts, and correlation of observed-to-expected mortality rates was moderate. In that study, examination of elective performance alone would have resulted in failure to detect units with high outlying non-elective mortality, and, by association, possible problems in quality of care for patients undergoing non-elective surgery. Examination of both cohorts was required to provide a broad assessment of the overall performance of a colorectal unit.

Studies 5 and 6 used different approaches to try and understand the key determinants of performance at units with the shortest and longest length of stay results after elective colonic surgery in the English NHS. Study 5 compared existing, publicly available data across a range of structures, processes and outcomes between units stratified by length of stay. One significant difference was found, and that was likely to have occurred by chance due to multiple comparisons. The lack of an association between certain variables, such as operative volume and staffing levels, and length of stay was surprising, and highlighted the need to collect novel data to detect more subtle differences between short and long stay units.

Study 6 involved development of a novel semi-structured interview protocol, to try to characterise and quantify the role of key structures and processes in the or-
ganisation of care among these high and low performing units. Problems enrolling units with long stay into the study precluded fair inclusion of interviews with both long and short stay units in the qualitative analysis performed. However, a good sample of short stay units was achieved, with all ten such units approached agreeing to take part. Qualitative analysis of the interview transcripts for these hospitals highlighted the importance of careful attention to the organisation and delivery of care. All units based their clinical care on the Enhanced Recovery Program. Consultants or trained ward nurses led the implementation of the ERP, day-to-day on the ward. Nurses were encouraged to exercise their clinical judgement to detect postoperative problems, and were empowered to act on their concerns. There was excellent communication between nursing staff and the consultants, with frequent senior medical input. These factors, above and beyond the ERP, seemed critically important to achieving the best length of stay results in the country.

This final study also provided other insights into unit-level performance. The high performing units that took part had similar operative volume and laparoscopy rates, when compared with the rest of the national cohort. There was also marked variation in analgesic approach across these ten units. These findings suggest that, while laparoscopy may be associated with a demonstrable reduction in length of stay within the controlled setting of a clinical trial, this effect may be overwhelmed by other factors in the care of colorectal patients, when situated in usual clinical practice. Analgesic strategy may also be less important than previously considered. It was also interesting to note that mortality and readmission rates among units with short length of stay were not significantly lower than other units nationally. Excellence in one area of care does not necessarily correlate with above average performance in other domains.
An exploratory quantitative analysis of the interview results was performed. This demonstrated the feasibility of conducting telephone interviews with clinical staff to gather novel, quantitative data on care organisation and delivery in surgical units. Previously highlighted problems recruiting long stay units compromised efforts to demonstrate a statistical association between interview scores and performance. The small sample size and self-selection of participating long stay units may have combined to hamper the ability of the interview tool to discriminate between different levels of performance. Further work is needed to develop these findings, and to explore the generalisation of the qualitative findings more widely.

The post hoc analysis reported, comparing the units that did and did not participate in the study, highlighted potential challenges in such work, and in health services research more generally. It is important to consider how to maximise engagement among units with poor performance. Research with this group of units may provide important insights into the determinants of outcomes for surgical patients, and improving care delivery in these units may provide direct benefits to improve the results for the patients they treat.

13.3 Limitations

There are a number of important limitations to the work presented in this thesis. These relate to the observational nature of the research, the extensive use of administrative data, and the focus on colorectal surgery.

All of the studies in this thesis used an observational design to try and better understand various aspects of surgical care. By its nature, observational research focuses on examining for associations to understand the phenomena observed.
Sampling and unmeasured confounding variables may have had an unknown impact upon the results, and this must always be remembered when making inferences about associations, regardless of how strong and plausible the association may be. For example, in the qualitative analysis of study 6 (see chapter 10), all the high performing units based their care on the ERP. Further work with a larger sample size should be conducted to examine the relationship between the ERP and surgical outcomes more deeply, across the full spectrum of surgical performance.

The key strengths and weaknesses of the HES administrative database were discussed in section 5.2.3. HES data has been used in this thesis with careful consideration of its limits. For example, operative and diagnostic codes have been used in the broadest possible ways, using groups rather than specific sub-codes. Nonetheless, data inaccuracies are inevitable, and even when data are correct, the absence of specific, clinically important variables, such as stage of disease for cancer, is an important limitation. While such limitations should always be remembered, HES data has been shown to perform similarly to voluntary disease registers for the large-scale study of health care\textsuperscript{164}. Therefore, HES data met the requirements of the studies presented in this thesis.

This thesis has focused almost exclusively on colorectal surgery. The colorectal speciality is a large and important area in the field of surgery, with over 30 000 major colorectal resections performed each year in England. However, technical and disease-specific considerations may limit the relevance of the findings of the present studies to other specialties, or in other health care systems internationally. Speciality-specific changes include the growth of laparoscopy and ERP care, and the introduction of bowel cancer screening. Health care policy may also have had a significant impact on colorectal practice. There has been a strong focus on cancer
care in England over recent decades, with targets attached to elective treatment pathways. More recently, the NHS has begun to publish consultant- and Trust-level mortality rates after colorectal surgery, amongst others\textsuperscript{18}. Whether this influences clinical practice has yet to be seen. However, all these factors must be considered before extrapolating the findings of this thesis to other areas of health care.

13.4 Understanding high performance

The results of the studies described above have helped contribute to a more sophisticated understanding of unit-level performance in colorectal surgery. In this section, the concept of the high performing unit is discussed and deconstructed, and a model for the achievement of high performance in a specific area of practice is presented.

Deconstructing high performance

The studies in this thesis help contribute to the conceptualisation of high performance in surgery. Building upon earlier research, the findings make it increasingly clear that there are no (or certainly very few) universally high, or low, performing units. Individual colorectal units may deliver good performance in specific areas of practice, but not in others. A unit may have high outlying mortality after elective colorectal surgery, yet the same team may deliver perfectly acceptable non-elective results. A unit with significantly lower than average mortality in the first 30 days after surgery may have average 90-day results. The lowest length of stay results in the country are not associated with significantly lower-than-average mortality or readmission rates.
Given the breadth of clinical practice, and the range of metrics available for its measurement, it is perhaps unsurprising that no units manage to achieve universally high performance across measures. When the full definition of quality of care is considered, as discussed in chapter 1, it may perhaps be considered that universal high performance is impossible. Naturally, units will have specific strengths and weaknesses. Some units will achieve low mortality for certain patients, others will have short length of stay, and some may be very effective at delivering patient centred care. Certain teams will be very efficiently organised, delivering patients through the diagnosis and treatment pathway more promptly than their peers, whereas others may be very good at minimising waste and delivering care for the minimum cost. Simultaneously achieving all of these may overwhelm even the most optimistic clinical or management team.

**How to achieve high performance**

The research in this thesis may help better understand how high performance, within a specific area of practice, may be achieved. To deliver the best results, as determined by a particular metric, all the necessary factors or ‘ingredients’ must appropriately align to achieve the desired outcome. The highest levels of performance may perhaps only be achieved in one, or perhaps a limited number of ways. Conversely, average levels of performance may be achieved in any of a great number of ways, due to the ability of clinical teams to compensate for problems in one area with corrective efforts elsewhere in the patient pathway.

The results of study 6 may be used to illustrate such a model of high performance. Achieving the shortest length of stay required that teams attended to
each of various critical areas of patient care. Patients’ expectations needed careful management from an early stage, before they underwent surgery. An appropriate ERP was a key requirement, but it was also vital to deliver this care efficiently, and this was best done by consultants or appropriately trained nurses. Excellent communication was required, both to allow rapid patient progression along the pathway, and to resolve problems promptly. In the units studied, all of these elements worked together in synergy to deliver the best length of stay results in the country.

13.5 Implications and future work

There are a number of implications that follow from the work presented, across various areas of research and health care more broadly. Several avenues for further work have also been identified. These will be discussed together in this section, grouped under the headings of: research; quality improvement; patient involvement; and policy. Several of the themes discussed in these subsections are relevant in more than one area, but they have been considered separately for convenience.

Research

The findings of this thesis have highlighted the importance of carefully defining the patient cohort and outcomes under consideration when examining unit-level performance. The strengths and weaknesses of individual units may result in rather different outcomes for elective or non-elective patients. Examining one group alone, or combining both, may result in a failure to identify key components of care that determine the outcomes for these different groups of patients. Similarly,
varying the time-point at which outcomes such as mortality are determined may have significant implications for unit performance. Considered positively, these differences represent an opportunity for further research to explore the reasons for differential performance in these areas. For example, a unit with low mortality after elective colorectal surgery and average non-elective results could be visited, and care assessed through interviews and observation, to understand the determinants of outcomes in these different areas.

It is also critical to consider the time-period under examination. If a particular study includes patients over a prolonged period of time, the magnitude of the background fall in mortality after surgery over time (certainly in colorectal surgery) may potentially overwhelm the effect of other influences on outcome. This finding also represents an opportunity for more research, specifically examining the relative performance of surgical units over time. If a particular unit is found to have changed its performance relative to its peers, from above average to below average, closer examination of the care at such a unit may yield novel insights into how to improve surgical performance.

Routine data has been shown to have important use in defining broad differences in care between units, but it was not adequate to identify how care differed across units with varying levels of performance, to truly understand how specific results may be achieved. Novel data must be collected to create new knowledge in this area. This thesis has shown that semi-structured telephone interviews are a feasible approach for collecting both qualitative and quantitative data on the organisation and delivery of care in a surgical unit. Ongoing work using a cross-sectional questionnaire design will hopefully build upon the findings of these interviews. Further exploration using similar techniques within colorectal surgery and
across other specialties is required to determine the generalisation of the findings presented in this thesis.

**Quality improvement**

The complex nature of colorectal performance highlighted in this thesis makes it clear that before embarking on specific quality improvement work, individual units or teams must clearly define and understand their own current level of performance, including their own strengths and weaknesses. This may require examination of patient subgroups and multiple different metrics. Doing so will help focus any improvement efforts on specific areas of need, and help to ensure that the correct measures are used to evaluate any changes.

It is also important to recognise the value of studying good performance, as well as examining errors and harm. It is clear that much has been learned through the examination of errors and adverse events over recent decades. However, broadening the focus of study to include the performance of successful units, as presented in this thesis, may help further advance our understanding of health care. The adoption of a more positive focus, studying the successes of a system as well as its failures, may also help improve engagement between researchers and the clinical community, enabling a more effective translation of research into practice.

The findings of this thesis also suggest that it is important to look beyond the immediate, clinical processes involved in surgical care, to try and understand surgical performance in more detail. The general decline in mortality documented in study 2, regardless of urgency, surgical approach, eligibility for screening or major complications, suggests that improvements may have arisen from broad
developments, rather than specific changes such as the introduction of laparoscopic surgical techniques. This is supported by the findings reported in study 6. For example, even among the ten units with the shortest length of stay in the country, laparoscopy rates varied widely. The organisation and delivery of surgical care, and non-surgical areas of medicine upon which surgical patients also depend (such as medical management of comorbidities, anaesthesia, radiology and intensive care) must also be given due consideration when working to further improve care quality. These areas have arguably been under-researched, and their importance under-appreciated, within the research and clinical community, perhaps because they are not immediately and uniquely surgical in nature.

The ultimate aim of all quality improvement work must be to improve care in clinical practice. Observational work is a necessary component in the development of interventions designed to improve key areas of practice. The work in this thesis, and the ongoing national colorectal questionnaire study, may form the foundation for an intervention to improve the outcomes of elective colorectal surgery. For example, the combined results of these studies may suggest that nurse-led care and strong nurse-consultant communication are significantly associated with reduced length of stay. Training for colorectal teams could be developed on this basis, helping nurses take on increasing responsibility, with clear and close support from consultant colleagues. The impact of such training could be evaluated within a carefully designed clinical trial. Clearly more work is required before such an intervention could be developed and implemented, but the current work lays a foundation upon which efforts towards this goal may be based. In addition, further research is needed to establish the relevance of these specific findings in other areas of surgical practice, and across medicine more broadly.
Patient involvement

The majority of gastrointestinal cancer patients that took part in study 1 expressed a clear desire to be more involved in decisions about provider choice than they had been during their previous investigation and surgical treatment. This finding has clear implications for patients and clinicians in the NHS, and should be interpreted carefully. It does not mean that patients want to make decisions for themselves. Rather, it is likely to represent patients’ wishes to have their opinions and views respected, and to have the opportunity to participate in decisions that have an impact upon them. Clinicians need to be reminded not to make assumptions about patient preferences, as there was no association between participant age or education, and their reported preferences. It is also important to note that patients’ preferences spanned the full range of responses. Clinicians should therefore explain to patients that they can choose their provider if they would like, and listen to their patients’ preferences if they want to explore this further. If a patient would like to consider the different providers available locally, the information rated as important in study 1 may provide a starting point for guiding choice. Clinicians should help their patients to understand and interpret such information, as certain data may not represent meaningful variation in quality of care. Previous research has shown that patients recognise the need for their doctors’ help to make sense of the decisions they face. More involvement in this type of decision may help patients develop a greater sense of control and reduce their anxiety, at a difficult and unsettling time for them.

However, greater involvement in decision-making and better understanding of surgical performance among patients may also have important unintended con-
sequences. In a study of provider choice, Boyce et al found that making people think about differences in quality between hospitals, arguably giving patients a more objective understanding of care quality, was associated with lower levels of satisfaction. Without such information, patients may prefer to believe that their local hospital is the best available, a belief they can sustain until presented with information or experience to the contrary. Therefore, well-intended efforts to improve patient involvement and provide more detailed information on the health care available may have negative consequences to increase anxiety and reduce satisfaction. As suggested above, the best approach may be for doctors to be guided by each patient’s response to the offer of choice, and to proceed accordingly.

**Policy**

The above findings concerning current patient involvement in provider choice have important implications for clinical leaders and policy makers. Patients are not currently adequately involved in provider choice for this to act as an effective lever to drive quality improvement. If provider choice is to be pursued as a mechanism for change, much needs to be done to increase patient involvement at each stage of the patient journey, and to better understand the way patients make decisions, so that they can be appropriately supported. It is unknown whether greater involvement would actually result in changes in referral patterns. Therefore, while greater involvement may be pursued to promote patient-centred care, even if this is achieved, it is not clear whether it will prove to be an effective strategy for improving quality of care.

The more nuanced and complex understanding of surgical performance pre-
sented in this thesis highlights the over-simplicity of current performance monitoring initiatives, such as the national publication of consultant and hospital mortality rates. NHS Choices began publishing department- and consultant-level mortality results after elective colorectal surgery for cancer during 2013\textsuperscript{224}. Representing the performance of a colorectal unit with a single metric for a narrowly selected patient group belies the true nature of colorectal practice. It neglects the importance of surgical resections for benign disease, and ignores the outcomes of patients undergoing surgery on a non-elective basis. If one of the objectives of publishing outcomes is to help drive quality improvement, it should be important to examine practice broadly, rather than restricting the focus to a subset of patients treated. To assess colorectal performance, or perhaps that of any surgical or medical speciality, it may be more accurate to consider a panel of metrics, including different outcome measures across key patient groups. While publication of the outcomes of surgical care is a great step forward for transparency in health care, restricting this within colorectal practice to mortality after elective cancer surgery will severely limit the ability of this initiative to detect care delivery problems.

13.6 Concluding remarks

This thesis aimed to develop the present understanding of unit-level performance in colorectal surgery, using a combination of research techniques to examine the structures, processes and outcomes of care.

It has successfully identified that gastrointestinal cancer patients are interested in being more involved in provider choice, and has shown the types of information patients consider important at the provider level.
By using large administrative datasets, it has highlighted the complexity of measuring and understanding colorectal performance. There have been significant improvements in post-operative mortality over time, apparently unrelated to some of the most salient advances in the care of colorectal cancer patients over recent years. When assessing performance, it is critical to consider what metrics are being used. To obtain an appropriately broad assessment, it may be necessary to consider multiple metrics across each of the key patient groups treated within a particular speciality.

Novel data on the structures and processes of care within colorectal units was collected using a semi-structured telephone interview. Data from high performing teams highlighted the importance of care organisation and implementation in achieving excellent results. These findings form the basis for ongoing further research, designed to determine their generalisation across the range of colorectal performance.

The work presented has helped me understand how I consider surgery should be practised to deliver the best possible care for patients, and how to apply this during my future career. Clinical care based upon the best available evidence may only realise its true potential if it is carefully delivered within a well-functioning team. This requires consideration of how each of the parts of a package of care should be implemented effectively. It also needs deliberate effort to nurture inter-professional relationships to ensure free and high quality communication between surgeons and nurses. Specific areas of practice should be selected as key priorities, together with appropriate measures by which to judge performance, in the understanding that it is unlikely to be possible to excel across all measures.

On a personal level, I have learned a great deal and developed a valuable
research skills base through undertaking the work presented. This PhD has been a highly enjoyable experience, and I am very keen to continue to pursue this research in combination with my clinical career as a surgeon. I will therefore be continuing my training on an academic pathway, in the hope of making further contributions in this important field.

13.7 Chapter summary

This final chapter has reviewed the key findings of the studies contained in this thesis. It has described its main limitations, and suggested possible future research strategies to develop an intervention to improve the quality of care that patients undergoing colorectal surgery receive.
References


[10] Department of Health. The NHS plan: a plan for investment, a plan for re-
pdf.

gov.uk/document/cm74/7432/7432.pdf.

uk/20130107105354/http://dh.gov.uk/prod{.}consum{.}dh/groups/dh{.}
digitalassets/documents/digitalasset/dh{.}093442.pdf.

[13] Bristol Royal Infirmary Inquiry. Learning from Bristol: the report of the
public inquiry into children’s heart surgery at the Bristol Royal Infirmary


[15] The Mid Staffordshire NHS Foundation Trust Inquiry. Independent In-
quiry into care provided by Mid Staffordshire NHS Foundation Trust Jan-
midstaffspublicinquiry.com/key-documents.

[16] The Mid Staffordshire NHS Foundation Trust Public Inquiry. Report of


[86] Robinson J and Wharrad H. Invisible nursing: exploring health outcomes at a global level. Relationships between infant and under-5 mortality rates and


[141] Schneider E C and Lieberman T. Publicly disclosed information about the


[199] Christou A, Katzenellenbogen J M, and Thompson S C. Australia’s national
bowel cancer screening program: does it work for indigenous Australians?


[259] Maggs F and Mallet M. Mortality in out-of-hours emergency medical admis-


[276] Judge T A, Piccolo R F, and Ilies R. The forgotten ones? The validity of


Part VI

Appendices
Section overview

This section contains additional material to supplement the work presented within the main body of this thesis.
Appendix A

Example literature search

A.1 Appendix overview

This appendix presents an example of the search strategy used during the literature review in chapter 3. This particular search was performed to identify evidence investigating whether there is an association between out of hours care provision and outcomes.

A.2 Search details

The search was performed on 23rd May 2013, using the OvidSP portal to search the MEDLINE database from 1946 to the present, as shown in table A.1.

After search number 13, the first hits were inspected. To focus the search and improve its specificity, some repeat searches were performed, restricted to title search only. After search number 18, titles and abstracts of the papers identified were screened, and relevant articles were retrieved. A list of the papers obtained is shown in table A.2.
<table>
<thead>
<tr>
<th>Number</th>
<th>Search term</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>mortalit$.ab,ti</td>
<td>431454</td>
</tr>
<tr>
<td>2</td>
<td>death$.ab,ti.</td>
<td>508438</td>
</tr>
<tr>
<td>3</td>
<td>morbidit$.ab,ti.</td>
<td>232384</td>
</tr>
<tr>
<td>4</td>
<td>complication$.ab,ti.</td>
<td>560636</td>
</tr>
<tr>
<td>5</td>
<td>(length adj3 stay).ab,ti.</td>
<td>34679</td>
</tr>
<tr>
<td>6</td>
<td>outcome$.ab,ti.</td>
<td>828277</td>
</tr>
<tr>
<td>7</td>
<td>(patient adj3 satisfaction).ab,ti.</td>
<td>20728</td>
</tr>
<tr>
<td>8</td>
<td>1 or 2 or 3 or 4 or 5 or 6 or 7</td>
<td>1994738</td>
</tr>
<tr>
<td>9</td>
<td>out of hours.ab,ti.</td>
<td>1131</td>
</tr>
<tr>
<td>10</td>
<td>time of day.ab,ti.</td>
<td>5780</td>
</tr>
<tr>
<td>11</td>
<td>(night adj5 day).ab,ti.</td>
<td>7977</td>
</tr>
<tr>
<td>12</td>
<td>9 or 10 or 11</td>
<td>14625</td>
</tr>
<tr>
<td>13</td>
<td>8 and 12</td>
<td>1917</td>
</tr>
<tr>
<td>14</td>
<td>out of hours.ti.</td>
<td>509</td>
</tr>
<tr>
<td>15</td>
<td>time of day.ti.</td>
<td>708</td>
</tr>
<tr>
<td>16</td>
<td>(night adj5 day).ti.</td>
<td>687</td>
</tr>
<tr>
<td>17</td>
<td>14 or 15 or 16</td>
<td>1899</td>
</tr>
<tr>
<td>18</td>
<td>8 and 17</td>
<td>199</td>
</tr>
</tbody>
</table>

Table A.1: Example literature search.
<table>
<thead>
<tr>
<th>First author</th>
<th>Title</th>
<th>Journal</th>
<th>Year, volume and pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asha, S E</td>
<td>No effect of time of day at presentation to the emergency department</td>
<td><em>Emergency Medicine Australia</em></td>
<td>2011;23(1):33-8</td>
</tr>
<tr>
<td>Beckett, D</td>
<td>Improvement in out-of-hours outcomes following the implementation of</td>
<td><em>Quarterly Journal of Medicine</em></td>
<td>2009;102(8):539-46</td>
</tr>
<tr>
<td>Berger, A</td>
<td>Out of hours percutaneous coronary interventions in acute coronary</td>
<td><em>Heart</em></td>
<td>2006;92(8):1157-8</td>
</tr>
<tr>
<td>Curtze, S</td>
<td>Does time of day or physician experience affect outcome of acute</td>
<td><em>International Journal of Stroke</em></td>
<td>2012;7(6):511-6</td>
</tr>
<tr>
<td>George, T J</td>
<td>Association of operative time of day with outcomes after thoracic</td>
<td><em>Journal of the American Medical Association</em></td>
<td>2011;305(21):2193-9</td>
</tr>
<tr>
<td>Kelz, R R</td>
<td>Time of day is associated with postoperative morbidity: an analysis</td>
<td><em>Annals of Surgery</em></td>
<td>2008;247(3):544-52</td>
</tr>
<tr>
<td>Kelz, R R</td>
<td>Time-of-day effects on surgical outcomes in the private sector: a</td>
<td><em>Journal of the American College of Surgeons</em></td>
<td>2009;209(4):434-45</td>
</tr>
<tr>
<td>Koike, S</td>
<td>Effect of time and day of admission on 1-month survival and</td>
<td><em>Resuscitation</em></td>
<td>2011;82(7):863-8</td>
</tr>
<tr>
<td>Lairez, O</td>
<td>Relationship between time of day, day of the week and in-hospital</td>
<td><em>Archives of Cardiovascular Diseases</em></td>
<td>2009;102(12):811-20</td>
</tr>
<tr>
<td>Maggs, F</td>
<td>Mortality in out-of-hours emergency medical admissions - more than</td>
<td><em>Journal of the Royal College of Physicians of Edinburgh</em></td>
<td>2010;40(2):115-8</td>
</tr>
</tbody>
</table>

Table A.2: Papers retrieved from example search.
<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Title</th>
<th>Journal</th>
<th>Year</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magid, D J</td>
<td>Relationship between time of day, day of the week, timeliness of reperfusion and in-hospital mortality for patients with acute ST-segment elevation myocardial infarction</td>
<td><em>Journal of the American Medical Association</em></td>
<td>2005</td>
<td>294(7):803-12</td>
</tr>
<tr>
<td>Modrall, G J</td>
<td>Time of day is associated with postoperative morbidity: an analysis of the National Surgical Quality Improvement Program data</td>
<td><em>Annals of Surgery</em></td>
<td>2008</td>
<td>248(3):500-1</td>
</tr>
<tr>
<td>Peled, Y</td>
<td>The effect of time of day on outcome of unscheduled cesarean deliveries</td>
<td><em>Journal of Maternal-Fetal and Neonatal Medicine</em></td>
<td>2011</td>
<td>24(8):1051-4</td>
</tr>
<tr>
<td>Seow, Y</td>
<td>Cold ischemia time, surgeon, time of day, and surgical complications</td>
<td><em>Transplantation</em></td>
<td>2004</td>
<td>77(9):1386-9</td>
</tr>
</tbody>
</table>

Table A.2 continued: Papers retrieved from example search.
Appendix B

Additional structure and process factors reviewed

B.1 Appendix overview

This appendix summarises the remaining literature identified and examined as part of the review described in chapter 3, that was not selected for presentation in that chapter. This additional material also contributed to the development of the interview protocol for study 6.

B.2 Structural factors

Technology

A number of studies have examined for an association between higher technology status and improved outcomes. Three studies scoring technology status according to predefined lists were identified. Shortell and coworkers found that higher equipment scores were associated with significantly lower risk-adjusted mortality in the intensive care unit. In a study of VA hospitals, combining questionnaires and site visits, Daley et al found units with low outlying mortality or morbidity had
higher technology scores than those with high outlying results. However, a later study in the same hospital system showed that a higher technology score was associated with higher risk-adjusted morbidity after general surgical procedures. Two studies were identified that used specific hospital services, such as open heart surgery or organ transplantation, as indicators of high technology. Silber et al reported lower postoperative mortality rates in hospitals with high technology, in a cohort of specific urology and general surgical patients. A later study of general surgical patients showed an association between high technology status and lower risk-adjusted mortality and failure-to-rescue rates.

The heterogeneity of measurement of technology limits comparison between the two types of study identified. Caution must be exercised before concluding there is a significant relationship between greater technology availability and improved outcomes. Technology may be a marker of other, more important factors in the local organisation and delivery of care, rather than a direct cause of improved outcomes.

Building layout and design

Ulrich et al undertook an extensive review of the literature regarding the impact of hospital design on the results of health care. The review included 459 papers, and concluded that specific design features (such as single-bed rooms, increased access to daylight, appropriate lighting, and views of nature) may improve patient and staff outcomes (including a reduction in hospital-acquired infections, medical errors and length of stay, with an increase in patient and staff satisfaction). The paper reported that while the quality of some studies was low, the consis-
tency of findings across the range of articles reviewed increased the strength of the conclusions reached.

To check the quality of the research, three papers reported as showing randomised evidence in favour of the impact of environmental factors on outcomes were retrieved and examined. This suggested that potential problems with the methodological rigour of the reviewed articles may have been under-reported. One study was described as examining length of stay and mortality in patients who had experienced a heart attack and ‘had been randomly assigned to sunny and dull rooms’. However, the original article did not describe any specific process of random allocation. Retrospective comparison of certain patient characteristics, such as age and sex, was performed and showed no evidence of systematic differences. This is clearly not equivalent to random, prospective allocation. Another study was cited as a prospective randomised trial providing evidence supporting a link between views of nature and reduced pain scores. The study was published as a three paragraph summary based on conference proceedings. 166 patients were included, allocated to 6 different visual stimuli. Therefore individual group size was small, restricted to a maximum of 28 patients. Methodological details regarding randomisation and statistical analysis were not provided and therefore could not be evaluated. In contrast to the quality of these studies, the third paper retrieved was a meta-analysis of 40 randomised clinical trials published in the *British Medical Journal*, examining interventions to prevent falls in elderly patients. The review included appropriate assessment of the methodological quality of each trial, including the Jadad score and details of concealment allocation.

It may cautiously be concluded that environmental factors may be a significant determinant of specific outcomes in certain patient populations, but high quality
evidence is limited.

Out-of-hours care

The term ‘out-of-hours’ may be variously interpreted. In this review, it included all care provided outside the usual working week, such as 08.00 to 17.00, Monday to Friday. Out-of-hours care has generally been studied from one of two stand-points: comparing weekdays against weekends, or day-time against overnight. While not a specific structural factor, out-of-hours work will be under the influence of structural factors such as staffing and resource levels. It is therefore considered in this section.

Significant observational evidence exists linking higher mortality rates with weekend care. Aylin et al compared mortality after elective surgery by day of the week, with Monday as the reference group\textsuperscript{251}. They found a consistently significantly higher postoperative mortality rate for elective surgery performed on weekends, with a tendency towards higher mortality for operations undertaken on Friday, dependent upon subgroup. Other studies of surgical care have also found higher death\textsuperscript{252} or complication\textsuperscript{253} rates for patients admitted on the weekend. Two studies of large cohorts of emergency admissions found an overall tendency towards higher death rates for patients admitted on the weekend, though this did not apply to all diagnosis or procedure groups\textsuperscript{254,255}.

A number of studies have found an association between overnight care and inferior outcomes. In two studies of emergent and non-emergent general and vascular surgical cohorts within both the VA and ACS NSQIP hospitals in the US, Kelz and colleagues found higher risk-adjusted mortality and morbidity rates in patients undergoing surgery overnight, although the findings were not entirely consistent
between the two studies\textsuperscript{256,257}. Egol et al found higher risk-adjusted mortality for trauma patients admitted overnight compared with during the day\textsuperscript{258}. Maggs and Mallet studied emergency medical admissions, and found higher mortality rates for patients admitted during the night, equivalent mortality comparing weekends and weekdays, and higher mortality for all out-of-hours versus in-hours admissions\textsuperscript{259}.

However, a number of studies have found equivalent outcomes between groups defined by time of day or day of the week. These include studies of thrombolysis for acute ischaemic stroke\textsuperscript{260}, admission to intensive care from the Accident and Emergency Department\textsuperscript{261}, and transplant surgery\textsuperscript{262,263}. The evidence for acute coronary syndrome is mixed\textsuperscript{264–267}.

While specific patient groups or procedures may have equivalent outcomes (perhaps due to clear diagnostic criteria, high professional awareness and well-evidenced treatments), a significant body of evidence based on large-scale databases suggests a broad association between inferior outcomes and care provided outside of routine office hours.

**Total work hours**

A number of papers have examined whether a reduction in total hours of work have been associated with any change in patient outcomes. The assessment of this relationship is complicated by the fact that most reductions have occurred as a result of large-scale reorganisations of care provision, precluding effective comparison between the intervention and a control group. In the US in 2003, the Accreditation Council for Graduate Medical Education introduced an 80-hour limit to the average working week for medical trainees. A number of researchers have
compared outcomes pre- and post-introduction of these regulations\textsuperscript{268,269}. A meta-analysis conducted by Jamal found a non-significantly raised risk of mortality in surgical patients associated with longer working hours\textsuperscript{270}. However, as all included studies adopted the pre- \textit{versus} post-regulation design, there is likely to have been confounding due to the background secular improvement in outcomes over time. Others have shown that reducing work hours is associated with greater hours of sleep, reduced attentional failures and reduced medical error rates, though no association with patient outcomes was shown\textsuperscript{271,272}.

Overall, the evidence does not show a clear relationship between total work hours and patient outcomes. It is likely that the interaction is complex, involving other changes in the organisation of working practice and supervision arrangements, as well as wider quality improvement initiatives. Potential negative results of reduced working hours, such as discontinuity of care and reduced physician experience, may well be offset by positive results, such as reduced fatigue and errors, or system-wide adaptations that may include increased senior supervision of trainees\textsuperscript{273}.

\textbf{B.3 Process factors}

\textbf{Leadership}

There is evidence of a link between different leadership styles and organisational outcomes. Most research on leadership has focused on nursing staff, with examination of related outcomes such as job satisfaction, retention and intention to leave. Cumminngs et al conducted a systematic review of leadership styles
within nursing\textsuperscript{274}. Leadership styles were dichotomised as relationship- or task-focused. Relationship-focused leadership included transformational and resonant styles, where individual and personal contributions are valued and recognised, directed towards achieving individual and organisational potential. Leadership that was task focused included management by exception and transactional styles, which tend to be reactive, addressing problems as they arise, and focus principally on delivery rather than potential. Relationship focused leadership was more favourably associated with positive outcomes, such as improved staff satisfaction, relationships and productivity, than leadership focused on task completion. Gilmartin and D’Aunno also found an association between transformational leadership in health care and improved outcomes such as individual and group satisfaction, retention and performance\textsuperscript{275}. Beyond the strictly medically focused literature, Judge and colleagues undertook a meta-analysis of studies examining the specific leadership concepts of consideration (the degree to which a leader shows concern and respect for followers) and initiating structure (clear definition of individuals’ roles and lines of communication)\textsuperscript{276}. They found a positive correlation between these factors and leadership outcomes, such as follower satisfaction and motivation. Overall, the literature suggests that leadership is important in determining staff, and potentially organisational, outcomes, but there was no evidence of a link between leadership and ‘hard’ outcomes such as mortality rates or length of stay.
Quality assessment and improvement

Two surgical outcomes reporting and quality improvement programs in the US have described significant improvement in outcomes associated with participation in their programs. The Northern New England Cardiovascular Disease Study Group developed a system of data collection to generate case-mix adjusted mortality rates\(^ {277}\). Their initial report presented the results of data collected over 6 years. Half-way through this period, they introduced feedback of outcomes to participating surgeons and institutions, training in continuous quality improvement methods, and site visits between participating units. Their analysis showed no change in mortality during the pre-intervention period. In the 3 years post-intervention, case-mix adjusted mortality fell by 24% (95% confidence interval 10-33%, \( p < 0.001 \)).

The VA NSQIP was first established in 1994 after completion of the National VA Surgical Risk Study\(^ {56}\). This ongoing program involves data collection by dedicated personnel in participating hospitals who prospectively record pre-, intra- and post-operative data on surgical patients. There is a system of checks to ensure data quality and accuracy. Data is submitted for central processing and feedback. Individual institutions receive reports comparing the care they provide and their risk-adjusted outcomes with peer group hospitals and the national average. The program is also designed to disseminate best practice from high performing institutions, and assist in the improvement of care in hospitals with poor performance. By 2005, 30-day mortality and morbidity within VA hospitals had fallen by 31% and 45% respectively\(^ {278}\). The success of the NSQIP resulted in its introduction to other private hospitals in America. Results within the private sector have similarly
shown improvement in outcomes over time.\textsuperscript{279}

While such longitudinal studies must be interpreted with some caution due to difficulties separating the causal role of particular programs from secular changes over time, these findings suggest that participation in specific initiatives that monitor performance and provide outcome reports may be associated with improvements in care over time. Work by Hibbard and colleagues suggests that feedback of performance (especially publicly) is significantly associated with improvements in care.\textsuperscript{140}

Within England, evidence suggests that participation in voluntary disease registries is associated with improved results. Typically, such registers facilitate central data collection, analysis and reporting, with feedback to participating units. Almoudaris et al used routine administrative data combined with data from the National Bowel Cancer Audit Program (NBOCAP) to show that, even at low rates of submission, participating English National Health Service Trusts had lower case-mix adjusted 30-day mortality rates and shorter length of stay than those that did not.\textsuperscript{160} Aylin and colleagues examined routinely collected data together with the National Vascular Database.\textsuperscript{161} When the mortality results for consultants that submitted to the voluntary register were compared with those of consultants who did not, there was a trend towards lower mortality for those participating in the National Vascular Database, though this was not statistically significant.

The published findings suggest that participation in voluntary registries or large-scale outcome monitoring and feedback programs is likely to be associated with better outcomes.
Management of adverse events and error

There is no clear evidence linking the management of errors or adverse incidents with improved outcomes. Mahajan described four key components to an effective adverse incident reporting system, including how data is collected, the specifics of the data that is recorded, how data is analysed, and the subsequent feedback and action based on the understanding gained. In the UK, the National Reporting and Learning System (NRLS), a national system for reporting and analysing adverse incidents, was established in 2003. Given the relative immaturity of the system, and the fundamental requirement of a change in culture for such a system to succeed, a lack of evidence supporting any impact on patient care is perhaps not surprising. Considering the lessons from other high-risk industries, it seems only logical that a hospital or health care system that encourages and facilitates open reporting of adverse events or errors will be better able to change practice to reduce the risk of similar events in the future.

Hutchison and colleagues have shown that a higher rate of reporting to the NRLS within a hospital is associated with higher staff-reported ratings of safety culture and a better risk-management rating from the NHS Litigation Authority. The authors did not show an association between reporting rate and ‘hard’ outcomes such as Hospital Standardised Mortality Rate (HSMR). However, accepting the four components of an effective reporting program proposed by Mahajan, reporting rates are only one facet of a complex system. Reporting rates do not indicate the quality of the data reported, nor the ability of the system to act upon it.

Measurement of the ability of an organisation to learn from error is likely to
be much more complex than an assessment of incident reporting rates, and a link
with outcomes has not yet been demonstrated.

B.4 Appendix summary

This appendix has summarised the additional evidence examined as part of the
literature review described in chapter 3. The evidence presented helped develop
a broad understanding of the determinants of department level performance in
surgery, and directly contributed to the development of the interview protocol for
study 6.
Appendix C

Study 1: specimen questionnaire

C.1 Appendix overview

This appendix contains the full patient questionnaire used for study 1, described in chapter 4.
Information needs and choice between providers of surgical care for gastrointestinal cancer

Choosing where to have tests and surgery

For the purpose of this questionnaire, please remember the time before you had surgery, when you had some tests to examine your insides and were then given the diagnosis of cancer. You may or may not have been given the option of choosing which hospital or clinic you went to for your test(s) and subsequent surgery.

For each of the following questions, please place a tick in the box under the option that best describes your opinion or experience. There is a comments box at the end of the survey for anything not covered by the questions or for you to share general comments and thoughts.

Being referred for tests

1. When you saw your General Practitioner (GP) initially, were you given the choice of which hospital or clinic to go to?

| Yes | No | I saw somebody other than my GP first e.g. doctor in A&E |

If you answered ‘No’, please skip the next question and move on to question 3.

2. What role did you have in the decision about which hospital or clinic you went to for your test(s)?

| A. I made the decision myself | B. I made the final decision after seriously considering my doctor’s opinion | C. My doctor and I shared responsibility for deciding | D. My doctor made the final decision after seriously considering my opinion | E. My doctor made the decision on his/her own |

3. Sometimes decisions may be made in ways that do not match a person’s preferences. What role would you like to have had in deciding about which hospital or clinic you went to for your test(s)?

| A. I would prefer that I made the decision myself | B. I would prefer that I made the final decision after seriously considering my doctor’s opinion | C. I would prefer that my doctor and I shared responsibility for deciding | D. I would prefer that my doctor made the final decision after seriously considering my opinion | E. I would prefer that my doctor made the decision on his/her own |
If you answered option E, please skip the next question and move on to question 5.

4. Who would you like to have discussed this decision with? Please tick the box(es) below that best matches your wishes. You may tick more than one option.

<table>
<thead>
<tr>
<th>Nobody else</th>
<th>My GP</th>
<th>A nurse at my GP practice</th>
<th>My spouse / family</th>
<th>Other (please specify below)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Other:

Having your operation

After tests you were given a diagnosis of cancer and likely went on to have further tests to assess your fitness for surgery.

5. Once you had these results and decided you would have an operation, did your specialist give you a choice of which hospital to go to for your surgery?

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
<th>Unsure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If you answered 'No', please skip the next question, and move on to question 7.

6. What role did you have in the decision about which hospital you went to for your operation?

<table>
<thead>
<tr>
<th>A. I made the decision myself</th>
<th>B. I made the final decision after seriously considering my doctor’s opinion</th>
<th>C. My doctor and I shared responsibility for deciding</th>
<th>D. My doctor made the final decision after seriously considering my opinion</th>
<th>E. My doctor made the decision on his/her own</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

7. At times decisions may be made in ways that do not match a person’s preferences. What role would you like to have had in deciding about which hospital you went to for your operation?

<table>
<thead>
<tr>
<th>A. I would prefer that I made the decision myself</th>
<th>B. I would prefer that I made the final decision after seriously considering my doctor’s opinion</th>
<th>C. I would prefer that my doctor and I shared responsibility for deciding</th>
<th>D. I would prefer that my doctor made the final decision after seriously considering my opinion</th>
<th>E. I would prefer that my doctor made the decision on his/her own</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
If you answered option E, please skip the next question and move on to question 9.

8. Who would you like to have discussed this decision with? Please tick the box(es) below that best matches your wishes. You may tick more than one option.

<table>
<thead>
<tr>
<th>Nobody else</th>
<th>My GP</th>
<th>A specialist hospital doctor</th>
<th>A specialist hospital nurse</th>
<th>My spouse / family</th>
<th>Other (please specify below)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Other:

CONTINUED ON THE NEXT PAGE
Information needs

We would like to better understand what information you would have considered important to know about the hospital and team treating you, once you knew you were going to have surgery.

The following list describes a number of aspects of a hospital in general and the surgical team in particular. Next to each is a scale of importance to you, ranging from 1 to 5, explained below.

<table>
<thead>
<tr>
<th></th>
<th>Not at all important to me</th>
<th>Slightly important to me</th>
<th>Important to me</th>
<th>Very important to me</th>
<th>Absolutely critical to me</th>
</tr>
</thead>
<tbody>
<tr>
<td>9. Rates of infection with ‘super-bugs’ in the hospital</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>10. How many beds there are in the hospital</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>11. How full the hospital usually is</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>12. How many consultants there are in the hospital</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>13. How many nurses, midwives and health visitors there are in the hospital</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>14. Whether the hospital successfully balances its income and spending</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>15. How satisfied staff are with the care they deliver across the whole hospital</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>16. Previous patients’ satisfaction with their care across the whole hospital</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>17. How long it takes me to get to the hospital from my home</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>18. Whether the whole hospital has an average number of deaths, or more or less than expected</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>19. How much it costs to park at the hospital per hour</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>20. How often patients are readmitted to hospital soon after receiving any form of care provided there</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>21. How often staff report things that may have gone wrong in the hospital</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>22. How often the hospital checks whether patients are at risk of clots on the legs or lungs while in the hospital</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>23. How many nurses, midwives and health visitors there are per bed in the hospital</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>24. How much of the hospital’s waste is sent for recycling</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Please circle the number beside each aspect of the hospital that best matches how important it would have been for you to know about each item when thinking about where you were going to have surgery. There is a comments box at the end of the survey for anything not covered by the questions or for you to share general comments and thoughts.
Information about the cancer and surgical treatment at the hospital

<table>
<thead>
<tr>
<th>Question</th>
<th>Not important</th>
<th>Slightly important</th>
<th>Important</th>
<th>Very important</th>
<th>Absolutely critical</th>
</tr>
</thead>
<tbody>
<tr>
<td>25. How long patients with your type of cancer wait for treatment</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>26. How many of this type of cancer operation the team performs each year</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>27. How often patients die soon after your type of cancer operation in this hospital</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>28. How well previous patients with your type of cancer rated the hospital</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>29. How often patients having any form of surgery stay in much longer than expected</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>30. How many operations of any type were cancelled at the last minute in the hospital over the previous year</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>31. The number of operations each year where the doctors and nurses accidentally left something behind inside the patient, such as a swab or instrument</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

32. Please use the box below for any comments or thoughts you wish to share. These may relate to anything you think is relevant, such as your experience and expectations of choosing between providers, the information you consider important, or something not included in the questionnaire.

Comments:
About you (optional)

It would be very helpful if you could fill in the following. This will help understand if different groups of people have different opinions or needs.

33. What type of cancer did you have surgery for?

| Upper GI cancer (oesophagus / gullet, stomach) | Lower GI cancer (colon, rectum) |

34. How long ago did you have surgery?

| Less than 1 year | 1 or 2 years | 3 to 5 years | 6 to 10 years | More than 10 years |

35. As far as you know, have you been cured by your surgery?

| Yes | No | Unsure |

36. Have you ever experienced any form of harm as a result of health care?

| Yes | No | Unsure |

37. What is your gender?

| Female | Male |

38. How old are you?

| Less than 40 | 40-49 | 50-59 | 60-69 | 70-79 | 80 or more |

39. What is your highest level of education?

| Secondary school | Vocational qualification | University | Postgraduate degree |

Thank you very much for giving your time to complete this questionnaire. If you have any questions or comments, please feel free to get in touch with Ben Byrne, lead researcher on this project, at benjamin.byrne@imperial.ac.uk.
Appendix D

Study 2: additional tables

D.1 Appendix overview

This appendix contains additional tables from study 2. These provide annual counts of operations and deaths for patient subgroups examined in that study. This data was used to create the figures presented in chapter 6 to illustrate trends over time. The data was also used to model for changes over time during the study period.

D.2 Tables

All of the following tables present annual counts of operations and deaths by urgency of operation and patient sex. Individual tables present subgroup data according to: surgical approach (laparoscopic or open); eligibility for bowel cancer screening (defined by age group); and presence of a major complication (defined as undergoing a surgical failure to rescue procedure or not).
<table>
<thead>
<tr>
<th>Year</th>
<th>Elective</th>
<th>Non-elective</th>
<th>D90 (%)</th>
<th>Cases</th>
<th>D90 (%)</th>
<th>Cases</th>
<th>D90 (%)</th>
<th>Cases</th>
<th>D90 (%)</th>
<th>Cases</th>
<th>D90 (%)</th>
<th>Cases</th>
<th>D90 (%)</th>
<th>Cases</th>
<th>D90 (%)</th>
<th>Cases</th>
<th>D90 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998-9</td>
<td>7755</td>
<td>554</td>
<td>6.1</td>
<td>12.5</td>
<td>7755</td>
<td>554</td>
<td>6.1</td>
<td>12.5</td>
<td>7755</td>
<td>554</td>
<td>6.1</td>
<td>12.5</td>
<td>7755</td>
<td>554</td>
<td>6.1</td>
<td>12.5</td>
<td>7755</td>
</tr>
<tr>
<td>1999-0</td>
<td>7907</td>
<td>537</td>
<td>6.8</td>
<td>12.6</td>
<td>7907</td>
<td>537</td>
<td>6.8</td>
<td>12.6</td>
<td>7907</td>
<td>537</td>
<td>6.8</td>
<td>12.6</td>
<td>7907</td>
<td>537</td>
<td>6.8</td>
<td>12.6</td>
<td>7907</td>
</tr>
<tr>
<td>2000-1</td>
<td>7838</td>
<td>535</td>
<td>6.8</td>
<td>12.5</td>
<td>7838</td>
<td>535</td>
<td>6.8</td>
<td>12.5</td>
<td>7838</td>
<td>535</td>
<td>6.8</td>
<td>12.5</td>
<td>7838</td>
<td>535</td>
<td>6.8</td>
<td>12.5</td>
<td>7838</td>
</tr>
<tr>
<td>2001-2</td>
<td>7278</td>
<td>537</td>
<td>7.4</td>
<td>12.9</td>
<td>7278</td>
<td>537</td>
<td>7.4</td>
<td>12.9</td>
<td>7278</td>
<td>537</td>
<td>7.4</td>
<td>12.9</td>
<td>7278</td>
<td>537</td>
<td>7.4</td>
<td>12.9</td>
<td>7278</td>
</tr>
<tr>
<td>2002-3</td>
<td>7512</td>
<td>523</td>
<td>7.0</td>
<td>12.5</td>
<td>7512</td>
<td>523</td>
<td>7.0</td>
<td>12.5</td>
<td>7512</td>
<td>523</td>
<td>7.0</td>
<td>12.5</td>
<td>7512</td>
<td>523</td>
<td>7.0</td>
<td>12.5</td>
<td>7512</td>
</tr>
<tr>
<td>2003-4</td>
<td>7468</td>
<td>470</td>
<td>6.3</td>
<td>10.8</td>
<td>7468</td>
<td>470</td>
<td>6.3</td>
<td>10.8</td>
<td>7468</td>
<td>470</td>
<td>6.3</td>
<td>10.8</td>
<td>7468</td>
<td>470</td>
<td>6.3</td>
<td>10.8</td>
<td>7468</td>
</tr>
<tr>
<td>2004-5</td>
<td>7453</td>
<td>463</td>
<td>6.2</td>
<td>10.9</td>
<td>7453</td>
<td>463</td>
<td>6.2</td>
<td>10.9</td>
<td>7453</td>
<td>463</td>
<td>6.2</td>
<td>10.9</td>
<td>7453</td>
<td>463</td>
<td>6.2</td>
<td>10.9</td>
<td>7453</td>
</tr>
<tr>
<td>2005-6</td>
<td>7810</td>
<td>454</td>
<td>5.8</td>
<td>9.3</td>
<td>7810</td>
<td>454</td>
<td>5.8</td>
<td>9.3</td>
<td>7810</td>
<td>454</td>
<td>5.8</td>
<td>9.3</td>
<td>7810</td>
<td>454</td>
<td>5.8</td>
<td>9.3</td>
<td>7810</td>
</tr>
<tr>
<td>2006-7</td>
<td>7411</td>
<td>405</td>
<td>5.5</td>
<td>9.0</td>
<td>7411</td>
<td>405</td>
<td>5.5</td>
<td>9.0</td>
<td>7411</td>
<td>405</td>
<td>5.5</td>
<td>9.0</td>
<td>7411</td>
<td>405</td>
<td>5.5</td>
<td>9.0</td>
<td>7411</td>
</tr>
<tr>
<td>2007-8</td>
<td>7061</td>
<td>404</td>
<td>5.7</td>
<td>8.5</td>
<td>7061</td>
<td>404</td>
<td>5.7</td>
<td>8.5</td>
<td>7061</td>
<td>404</td>
<td>5.7</td>
<td>8.5</td>
<td>7061</td>
<td>404</td>
<td>5.7</td>
<td>8.5</td>
<td>7061</td>
</tr>
<tr>
<td>2008-9</td>
<td>6551</td>
<td>329</td>
<td>5.0</td>
<td>6.0</td>
<td>6551</td>
<td>329</td>
<td>5.0</td>
<td>6.0</td>
<td>6551</td>
<td>329</td>
<td>5.0</td>
<td>6.0</td>
<td>6551</td>
<td>329</td>
<td>5.0</td>
<td>6.0</td>
<td>6551</td>
</tr>
<tr>
<td>2009-10</td>
<td>5912</td>
<td>305</td>
<td>5.2</td>
<td>5.5</td>
<td>5912</td>
<td>305</td>
<td>5.2</td>
<td>5.5</td>
<td>5912</td>
<td>305</td>
<td>5.2</td>
<td>5.5</td>
<td>5912</td>
<td>305</td>
<td>5.2</td>
<td>5.5</td>
<td>5912</td>
</tr>
<tr>
<td>2010-1</td>
<td>5634</td>
<td>258</td>
<td>4.6</td>
<td>4.8</td>
<td>5634</td>
<td>258</td>
<td>4.6</td>
<td>4.8</td>
<td>5634</td>
<td>258</td>
<td>4.6</td>
<td>4.8</td>
<td>5634</td>
<td>258</td>
<td>4.6</td>
<td>4.8</td>
<td>5634</td>
</tr>
<tr>
<td>2011-2</td>
<td>5008</td>
<td>201</td>
<td>4.0</td>
<td>4.0</td>
<td>5008</td>
<td>201</td>
<td>4.0</td>
<td>4.0</td>
<td>5008</td>
<td>201</td>
<td>4.0</td>
<td>4.0</td>
<td>5008</td>
<td>201</td>
<td>4.0</td>
<td>4.0</td>
<td>5008</td>
</tr>
<tr>
<td>Totals</td>
<td>98598</td>
<td>5975</td>
<td>6.1</td>
<td>6.1</td>
<td>98598</td>
<td>5975</td>
<td>6.1</td>
<td>6.1</td>
<td>98598</td>
<td>5975</td>
<td>6.1</td>
<td>6.1</td>
<td>98598</td>
<td>5975</td>
<td>6.1</td>
<td>6.1</td>
<td>98598</td>
</tr>
</tbody>
</table>

Table D.1: Number of resections and 90-day deaths, by year, urgency, surgical approach and sex of patient.

D90 - deaths within the first 90 postoperative days.
<table>
<thead>
<tr>
<th>Year</th>
<th>Male</th>
<th>Female</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
<td>D90 (%)</td>
<td>Cases</td>
<td>D90 (%)</td>
</tr>
<tr>
<td>1998-9</td>
<td>2335</td>
<td>105</td>
<td>1581</td>
<td>54</td>
</tr>
<tr>
<td>1999-0</td>
<td>2399</td>
<td>103</td>
<td>1515</td>
<td>55</td>
</tr>
<tr>
<td>2000-1</td>
<td>2330</td>
<td>104</td>
<td>1391</td>
<td>39</td>
</tr>
<tr>
<td>2001-2</td>
<td>2115</td>
<td>87</td>
<td>1304</td>
<td>37</td>
</tr>
<tr>
<td>2002-3</td>
<td>2179</td>
<td>88</td>
<td>1339</td>
<td>28</td>
</tr>
<tr>
<td>2003-4</td>
<td>2184</td>
<td>80</td>
<td>1343</td>
<td>24</td>
</tr>
<tr>
<td>2004-5</td>
<td>2185</td>
<td>85</td>
<td>1313</td>
<td>40</td>
</tr>
<tr>
<td>2005-6</td>
<td>2408</td>
<td>88</td>
<td>1475</td>
<td>36</td>
</tr>
<tr>
<td>2006-7</td>
<td>2404</td>
<td>67</td>
<td>1507</td>
<td>33</td>
</tr>
<tr>
<td>2007-8</td>
<td>2662</td>
<td>85</td>
<td>1670</td>
<td>34</td>
</tr>
<tr>
<td>2008-9</td>
<td>3159</td>
<td>77</td>
<td>1932</td>
<td>34</td>
</tr>
<tr>
<td>2009-10</td>
<td>3137</td>
<td>76</td>
<td>2012</td>
<td>30</td>
</tr>
<tr>
<td>2010-1</td>
<td>3338</td>
<td>59</td>
<td>2089</td>
<td>31</td>
</tr>
<tr>
<td>2011-2</td>
<td>3223</td>
<td>53</td>
<td>1930</td>
<td>23</td>
</tr>
<tr>
<td>Totals</td>
<td>36058</td>
<td>1137</td>
<td>22401</td>
<td>498</td>
</tr>
</tbody>
</table>

Table D.2: Number of resections and 90-day deaths among screening age patients (aged 60-69 years), by year, urgency and sex of patient.

D90 - deaths within the first 90 postoperative days.
### Table D.3: Number of resections and 90-day deaths among non-screening age patients, by year, urgency and sex of patient.

<table>
<thead>
<tr>
<th>Year</th>
<th>Male Cases</th>
<th>Male D90 (%)</th>
<th>Female Cases</th>
<th>Female D90 (%)</th>
<th>Male Cases</th>
<th>Male D90 (%)</th>
<th>Female Cases</th>
<th>Female D90 (%)</th>
<th>Male Cases</th>
<th>Male D90 (%)</th>
<th>Female Cases</th>
<th>Female D90 (%)</th>
<th>Male Cases</th>
<th>Male D90 (%)</th>
<th>Female Cases</th>
<th>Female D90 (%)</th>
<th>Male Cases</th>
<th>Male D90 (%)</th>
<th>Female Cases</th>
<th>Female D90 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998-9</td>
<td>1437</td>
<td>31</td>
<td>2.2</td>
<td>1051</td>
<td>36</td>
<td>3.4</td>
<td>1057</td>
<td>150</td>
<td>14.2</td>
<td>1236</td>
<td>160</td>
<td>12.9</td>
<td>390</td>
<td>391</td>
<td>10.0</td>
<td>360</td>
<td>26</td>
<td>7.2</td>
<td>500</td>
<td>170</td>
</tr>
<tr>
<td>1999-0</td>
<td>1499</td>
<td>29</td>
<td>1.9</td>
<td>1075</td>
<td>22</td>
<td>2.0</td>
<td>1044</td>
<td>156</td>
<td>14.9</td>
<td>1233</td>
<td>143</td>
<td>11.6</td>
<td>367</td>
<td>328</td>
<td>8.7</td>
<td>358</td>
<td>22</td>
<td>6.1</td>
<td>528</td>
<td>197</td>
</tr>
<tr>
<td>2000-1</td>
<td>1386</td>
<td>23</td>
<td>1.7</td>
<td>1050</td>
<td>19</td>
<td>1.8</td>
<td>1132</td>
<td>150</td>
<td>13.3</td>
<td>1193</td>
<td>124</td>
<td>10.4</td>
<td>372</td>
<td>287</td>
<td>7.5</td>
<td>308</td>
<td>21</td>
<td>6.8</td>
<td>491</td>
<td>180</td>
</tr>
<tr>
<td>2001-2</td>
<td>1300</td>
<td>27</td>
<td>2.1</td>
<td>954</td>
<td>18</td>
<td>1.9</td>
<td>1163</td>
<td>192</td>
<td>16.5</td>
<td>1224</td>
<td>124</td>
<td>10.1</td>
<td>373</td>
<td>30</td>
<td>8.0</td>
<td>336</td>
<td>28</td>
<td>8.8</td>
<td>516</td>
<td>171</td>
</tr>
<tr>
<td>2002-3</td>
<td>1441</td>
<td>31</td>
<td>2.2</td>
<td>1062</td>
<td>11</td>
<td>1.0</td>
<td>1201</td>
<td>191</td>
<td>15.9</td>
<td>1366</td>
<td>128</td>
<td>9.4</td>
<td>371</td>
<td>24</td>
<td>6.5</td>
<td>321</td>
<td>25</td>
<td>7.8</td>
<td>535</td>
<td>191</td>
</tr>
<tr>
<td>2003-4</td>
<td>1396</td>
<td>27</td>
<td>1.9</td>
<td>956</td>
<td>15</td>
<td>1.6</td>
<td>1239</td>
<td>157</td>
<td>12.7</td>
<td>1245</td>
<td>106</td>
<td>8.5</td>
<td>354</td>
<td>33</td>
<td>9.3</td>
<td>317</td>
<td>28</td>
<td>8.8</td>
<td>498</td>
<td>177</td>
</tr>
<tr>
<td>2004-5</td>
<td>1476</td>
<td>24</td>
<td>1.6</td>
<td>1016</td>
<td>11</td>
<td>1.1</td>
<td>1248</td>
<td>162</td>
<td>13.0</td>
<td>1302</td>
<td>131</td>
<td>10.1</td>
<td>339</td>
<td>329</td>
<td>8.4</td>
<td>340</td>
<td>26</td>
<td>7.6</td>
<td>510</td>
<td>185</td>
</tr>
<tr>
<td>2005-6</td>
<td>1509</td>
<td>30</td>
<td>2.0</td>
<td>1131</td>
<td>14</td>
<td>1.2</td>
<td>1381</td>
<td>188</td>
<td>13.6</td>
<td>1391</td>
<td>108</td>
<td>7.8</td>
<td>395</td>
<td>41</td>
<td>10.4</td>
<td>371</td>
<td>32</td>
<td>8.6</td>
<td>474</td>
<td>156</td>
</tr>
<tr>
<td>2006-7</td>
<td>1515</td>
<td>33</td>
<td>2.2</td>
<td>1146</td>
<td>14</td>
<td>1.2</td>
<td>1426</td>
<td>185</td>
<td>13.0</td>
<td>1482</td>
<td>117</td>
<td>7.9</td>
<td>359</td>
<td>23</td>
<td>6.4</td>
<td>312</td>
<td>29</td>
<td>9.3</td>
<td>465</td>
<td>158</td>
</tr>
<tr>
<td>2007-8</td>
<td>1499</td>
<td>20</td>
<td>1.3</td>
<td>1149</td>
<td>11</td>
<td>1.0</td>
<td>1449</td>
<td>166</td>
<td>11.5</td>
<td>1534</td>
<td>112</td>
<td>7.3</td>
<td>259</td>
<td>18</td>
<td>6.9</td>
<td>278</td>
<td>15</td>
<td>5.4</td>
<td>443</td>
<td>127</td>
</tr>
<tr>
<td>2008-9</td>
<td>1502</td>
<td>22</td>
<td>1.5</td>
<td>1165</td>
<td>16</td>
<td>1.4</td>
<td>1522</td>
<td>173</td>
<td>11.4</td>
<td>1552</td>
<td>119</td>
<td>7.7</td>
<td>295</td>
<td>26</td>
<td>8.8</td>
<td>274</td>
<td>22</td>
<td>8.0</td>
<td>426</td>
<td>133</td>
</tr>
<tr>
<td>2009-10</td>
<td>1452</td>
<td>12</td>
<td>0.8</td>
<td>1062</td>
<td>6</td>
<td>0.6</td>
<td>1574</td>
<td>165</td>
<td>10.5</td>
<td>1573</td>
<td>96</td>
<td>6.1</td>
<td>286</td>
<td>21</td>
<td>7.3</td>
<td>276</td>
<td>18</td>
<td>6.5</td>
<td>452</td>
<td>137</td>
</tr>
<tr>
<td>2010-1</td>
<td>1496</td>
<td>17</td>
<td>1.1</td>
<td>1209</td>
<td>11</td>
<td>0.9</td>
<td>1609</td>
<td>137</td>
<td>8.5</td>
<td>1527</td>
<td>75</td>
<td>4.9</td>
<td>314</td>
<td>19</td>
<td>6.1</td>
<td>318</td>
<td>20</td>
<td>6.3</td>
<td>435</td>
<td>121</td>
</tr>
<tr>
<td>2011-2</td>
<td>1618</td>
<td>14</td>
<td>0.9</td>
<td>1170</td>
<td>7</td>
<td>0.6</td>
<td>1557</td>
<td>124</td>
<td>8.0</td>
<td>1543</td>
<td>79</td>
<td>5.1</td>
<td>297</td>
<td>21</td>
<td>7.1</td>
<td>286</td>
<td>16</td>
<td>5.6</td>
<td>382</td>
<td>96</td>
</tr>
</tbody>
</table>

Totals: 20526 cases, 340 deaths (1.7%); 15196 cases, 211 deaths (1.4%); 18602 cases, 2296 deaths (12.3%); 19401 cases, 1622 deaths (8.4%).
<table>
<thead>
<tr>
<th>Year</th>
<th>Male</th>
<th>Female</th>
<th>Male</th>
<th>Female</th>
<th>Male</th>
<th>Female</th>
<th>Male</th>
<th>Female</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cases</td>
<td>D90 (%)</td>
<td>Cases</td>
<td>D90 (%)</td>
<td>Cases</td>
<td>D90 (%)</td>
<td>Cases</td>
<td>D90 (%)</td>
<td>Cases</td>
<td>D90 (%)</td>
</tr>
<tr>
<td>1998-9</td>
<td>7419</td>
<td>6.5</td>
<td>5927</td>
<td>6.0</td>
<td>379</td>
<td>20.1</td>
<td>163</td>
<td>25.8</td>
<td>2237</td>
<td>20.7</td>
</tr>
<tr>
<td>1999-0</td>
<td>7574</td>
<td>6.0</td>
<td>5991</td>
<td>5.2</td>
<td>379</td>
<td>22.2</td>
<td>164</td>
<td>19.5</td>
<td>2290</td>
<td>20.3</td>
</tr>
<tr>
<td>2000-1</td>
<td>7448</td>
<td>6.2</td>
<td>5593</td>
<td>4.7</td>
<td>377</td>
<td>19.4</td>
<td>174</td>
<td>20.1</td>
<td>2108</td>
<td>20.3</td>
</tr>
<tr>
<td>2001-2</td>
<td>6974</td>
<td>6.7</td>
<td>5251</td>
<td>4.5</td>
<td>347</td>
<td>21.0</td>
<td>168</td>
<td>23.2</td>
<td>2085</td>
<td>20.4</td>
</tr>
<tr>
<td>2002-3</td>
<td>7230</td>
<td>5.6</td>
<td>5429</td>
<td>4.2</td>
<td>372</td>
<td>18.8</td>
<td>159</td>
<td>10.7</td>
<td>2018</td>
<td>20.4</td>
</tr>
<tr>
<td>2003-4</td>
<td>7371</td>
<td>5.6</td>
<td>5443</td>
<td>4.8</td>
<td>369</td>
<td>16.0</td>
<td>181</td>
<td>16.6</td>
<td>2028</td>
<td>20.3</td>
</tr>
<tr>
<td>2004-5</td>
<td>7430</td>
<td>5.1</td>
<td>6027</td>
<td>3.7</td>
<td>409</td>
<td>18.3</td>
<td>192</td>
<td>17.2</td>
<td>1998</td>
<td>18.9</td>
</tr>
<tr>
<td>2005-6</td>
<td>8010</td>
<td>5.1</td>
<td>6027</td>
<td>3.7</td>
<td>422</td>
<td>15.4</td>
<td>189</td>
<td>19.0</td>
<td>1839</td>
<td>18.7</td>
</tr>
<tr>
<td>2006-7</td>
<td>7968</td>
<td>4.9</td>
<td>6126</td>
<td>3.7</td>
<td>422</td>
<td>15.4</td>
<td>189</td>
<td>19.0</td>
<td>1839</td>
<td>18.7</td>
</tr>
<tr>
<td>2007-8</td>
<td>8356</td>
<td>4.6</td>
<td>6355</td>
<td>3.3</td>
<td>438</td>
<td>18.7</td>
<td>209</td>
<td>14.8</td>
<td>1668</td>
<td>18.5</td>
</tr>
<tr>
<td>2008-9</td>
<td>8804</td>
<td>4.2</td>
<td>6606</td>
<td>2.8</td>
<td>531</td>
<td>13.9</td>
<td>258</td>
<td>19.0</td>
<td>1609</td>
<td>17.2</td>
</tr>
<tr>
<td>2009-10</td>
<td>8831</td>
<td>4.0</td>
<td>6651</td>
<td>2.8</td>
<td>563</td>
<td>12.6</td>
<td>250</td>
<td>11.6</td>
<td>1652</td>
<td>17.7</td>
</tr>
<tr>
<td>2010-11</td>
<td>9237</td>
<td>3.4</td>
<td>6883</td>
<td>2.4</td>
<td>584</td>
<td>11.5</td>
<td>275</td>
<td>9.5</td>
<td>1696</td>
<td>15.9</td>
</tr>
<tr>
<td>2011-12</td>
<td>9151</td>
<td>3.0</td>
<td>6684</td>
<td>1.8</td>
<td>527</td>
<td>8.9</td>
<td>248</td>
<td>7.3</td>
<td>1527</td>
<td>14.8</td>
</tr>
<tr>
<td>Totals</td>
<td>111648</td>
<td>5.0</td>
<td>84001</td>
<td>3.8</td>
<td>6058</td>
<td>16.2</td>
<td>2788</td>
<td>16.1</td>
<td>20822</td>
<td>19.0</td>
</tr>
</tbody>
</table>

Table D.4: Number of resections and 90-day deaths, by year, urgency, surgical failure to rescue and sex of patient. D90 - deaths within the first 90 postoperative days.
Appendix E

Studies 5 and 6: hospital-level length of stay results

E.1 Appendix overview

This appendix explains in detail how length of stay results were retrieved from the Dr Foster Real Time Monitoring version 8.0 (RTMv8) web tool for studies 5 and 6 (see chapters 9, 10, 11 and 12).

E.2 Retrieving length of stay results

The following steps were followed to create a spreadsheet of hospital-level risk-adjusted LLoS results for all NHS hospitals in England using the Dr Foster Intelligence RTMv8 web-based tool:

• Log in to the system at https://da.drfoster.co.uk/.

• Select ‘RTMv8’ from ‘My Tools’ menu.

• Select ‘Provider’ under ‘Unit Type:’, and select an NHS Trust under ‘Provider’.

• Select ‘Relative Risk’ from the ‘Reports’ menu.
• Select the following options on the displayed webpage:

  – **Outcome**: Length of stay (superspell).
  
  – **Basket**: Procedures - All.
  
  – **Chapter**: Lower Digestive Tract.
  
  – **Procedure Group**: Excision of colon and/or rectum.
  
  – **Admission Type**: Elective.
  
  – **Age Range**: 15+.
  
  – **View From**: January 2011.
  
  – **View To**: December 2012.
  
  – **Analyse By**: Subgroup.

• Check that ‘Length of stay (superspell)’ is still selected under ‘Outcome:’.

• Click on ‘Generate report’.

• Select ‘without rectum, with neoplasm’ and ‘without rectum, without neoplasm’.

• Select ‘Site’ under ‘Analyse By:’.

• Click on ‘Generate report’.

The RTMv8 tool provides the following information about the cohort of patients meeting the selected criteria:

• **Site** Hospital site within the Trust.
• **Spells** Number of patient spells within the cohort. As ‘superspells’ were selected earlier, single admissions during which care was transferred between consultants (considered as distinct ‘episodes’ of care in the HES database) or between hospitals (considered as separate ‘spells’) are considered together as one episode.

• **% of all** Proportion of cases meeting the selected criteria that are specific to each site. If only one site is included, this value will be 100.

• **Long LoS** The number of spells meeting criteria for long length of stay.

• **%** The proportion of spells for each specific site that meet criteria for long length of stay. This equates to \((\text{Long LoS} / \text{Spells}) \times 100\).

• **Expected** The number of spells that would be expected to have a long length of stay, calculated by the multivariable logistic regression.

• **%** The proportion of spells for each specific site that would be expected to have a long length of stay according to the regression model. This equates to \((\text{Expected} / \text{Spells}) \times 100\).

• **RR** The relative risk of experiencing a long length of stay for a specific site, after adjustment for case-mix by the logistic regression model, with a reference point of 100. This equates to \((\text{Long LoS} / \text{Expected}) \times 100\).

• **Low** The low limit of the 95% confidence intervals for the estimate of the RR presented above. This provides an indication of the degree of uncertainty in the sample estimate of the true population RR for the specific site under consideration, and is related to the sample size, i.e. number of spells.
• **High** The high limit of the 95% confidence intervals for the estimate of the RR presented above.

This process was repeated for all English NHS Trusts to obtain outcomes for all hospitals during the study period.

In some Trusts, not every patient record was identified with a specific hospital site, or multiple hospital site codes were used. The following rules were used to handle the results of the above analysis in these cases:

- **‘Unknown’ hospital site code.** If the search results were all associated with an ‘Unknown’ hospital site code, it was assumed there was only one hospital site performing colorectal surgery for the Trust under consideration, and all procedures were performed at this site. If there was one named hospital site as well as ‘Unknown’, it was assumed that all procedures were performed at the named hospital site, so the results were combined using the RTMv8 tool.

- **Multiple site codes.** Hospital site names were checked with individual NHS Trusts. Some site codes clearly represented new codes for the same site, and these results were therefore merged. In other instances, sites were renamed during the study period and new site codes assigned. These were also merged under one site name.

- **Exclusions.** If there were two named hospital sites as well as ‘Unknown’, those cases assigned the ‘Unknown’ code were discarded as they could not be reliably associated with a particular hospital site.
Appendix F

Study 6: interview piloting and refinement

F.1 Appendix overview

This appendix describes the piloting process for the interview protocol used in study 6, reported in chapters 10, 11 and 12.

F.2 Expert interviews

An initial interview protocol was developed as described in chapter 10. The proposed schedule was first piloted face-to-face with three expert interviewees (two senior academic colorectal surgeons and one senior colorectal nurse practitioner). Subsequently, the interview protocol was refined as follows:

- Questions reordered. The order of questioning was reorganised to improve the flow of the conversation. For example, the question about standardisation was revised to begin by asking if all consultants managed patients the same way, with more general questions about wider standardisation of care subsequent to this specific question. As another example, the question regarding the use of outcomes data within the department was moved to the
end. Within the question on communication and collaboration, the components of the question were re-ordered so that ward-level communication was discussed before communication at a higher level among leaders of the services.

- **Increased focus.** Some questions were rephrased to be more focused and specific. For example, the question, ‘Tell me about the level and quality of equipment in your hospital?’ was revised to a surgeon- or nurse-specific question regarding equipment in theatre or on the ward, respectively. The interviews revealed that the question, ‘At the ward level, is there clear leadership of the medical and nursing teams?’ tended to elicit affirmative responses. This was therefore changed to, ‘At the ward level, who defines the goals of the clinical team and what constitutes acceptable and unacceptable levels of care?’ The question, ‘Tell me about the supporting services available to help you manage an unwell patient with a complication such as an anastomotic leak?’ was changed to ask, ‘How do you consider the provision of supporting services, such as intensive care, radiology, or medical specialties such as cardiology, matches the need for such care for your patients?’ It was hoped that more focused questions would help generate more specific responses, allowing greater discrimination between units.

- **Areas to develop.** The expert interviewees suggested that more detailed questioning about the local approach to adverse events or clinical errors may generate useful data. It was hoped that this would provide some insight into the local culture and attitudes to error, as well as the ability of the team to learn from such events. This question was therefore separated from the
preceding question on clinical leadership. A review of the available evidence suggested an association between incident reporting and organisational culture, though no evidence was found providing a link with patient outcomes.

- **New areas to cover.** The expert interviewees considered it important to ask study participants for their thoughts regarding how length of stay may be reduced. This could help develop a greater understanding of length of stay after surgery by collecting together the insights of a number of experienced clinicians, and could potentially help direct future research.

- **Areas to discard.** Given the identification of areas considered important to explore further, it was necessary to reduce or eliminate other elements of the interview schedule to keep the overall interview to approximately 30 minutes. It was decided to remove the question on hospital structure and layout, and to rationalise the final fact-gathering questions. Questions that could be answered through other information sources if required were discarded.

### F.3 Peer interviews

A further 4 face-to-face and 5 telephone interviews were conducted with peers (surgical registrars, senior house officers and a research nurse) within the research department. During this pilot period, further refinements were made, including:

- **Introduction abridged.** As respondents were to be sent a detailed outline of the project before deciding to participate, the introductory preamble was shortened to remove redundant information.
• Questions anchored to clinical examples. To guide respondents, questions were based upon specific clinical scenarios or examples as far as possible. For example, the question on routine postoperative management was modified to begin with a specific clinical example.

• Revision of leadership question. The question regarding who sets specific goals for the ward team did not provide rich responses. Therefore, this was revised to ask about long- versus short-term goal setting. During piloting, this elicited detailed responses that provided some insight into the management of a unit’s priorities by its leaders.

• Areas to consider for future studies. Interviewees suggested two areas that may be appropriate for future study. Anecdotal experience suggested that the functioning of the Multi-Disciplinary Team meeting would be likely to vary significantly between Trusts. In addition, teamwork in theatre has been the subject of increasing study over recent years, with some evidence providing a link between the functioning of the clinical team at the time of surgery and clinical outcomes. While both of these areas were considered appropriate for assessment using the interview tool, such questions could only be posed to surgeons and not ward nurses or sisters, as they are not directly involved in these processes. Therefore, it would not be possible to triangulate the assessment of these processes through multiple interviewees in this study. However, they could be considered for inclusion in future studies if appropriate.
F.4 Appendix summary

This appendix has described the strategy adopted to pilot and refine the preliminary interview schedule for study 6. The final protocol is provided in full in appendix G.
Appendix G

Study 6: interview protocol

G.1 Appendix overview

This appendix contains the full schedule for the semi-structured interview developed for study 6. The development of this protocol is described in chapter 10.
HiPer study interview schedule
Length of stay after elective colonic surgery

Date

Time

Interviewer

Interviewee

Organisation

Opening the conversation

“Many thanks for agreeing to talk to me today. You have previously given consent for this interview (including recording if consent given) and I am very grateful for your time.

“Before we start, it is important that I point out that there are no right or wrong answers. I would like you to share both positive and negative opinions or observations you may have. Please simply answer the questions as honestly as you can, and I’d like to reassure you that everything you say will be kept confidential.”

“At this stage, do you have any questions?”
Introductory question

“I would like to start by asking you a very general question.”

1. “Could you please explain to me your role within the colorectal department?”
Structural factors (1 of 2)

Equipment

"I'd like to ask you some questions about the resources available within the hospital."

2. **SURGEONS:** “In theatres in your hospital, please could you tell me about the level and quality of equipment available to you?”

**NURSES:** “On the ward in your hospital, please could you tell me about the level and quality of equipment available to you?”

- Do you have all the equipment you need or would like?
- E.g. HD laparoscopes, staplers, energy devices / cardiac monitors, bladders scanners, BP machines.
- Is it up-to-date?
- Is it reliable or does it tend to need frequent repairs?
- Any problems with supply or stock control of consumables?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
</table>
Structural factors (2 of 2)

Human resources practices

3. “Tell me about the level of medical and nursing staff in your hospital. Do you have enough staff on the wards?”

- In hours and out of hours
- Use of locum / agency staff
- Turnover
- Hiring of replacements when leave or retire / often vacancies

"Are talented staff identified and developed, or do they tend look for opportunities elsewhere?"

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
</table>

381
Clinical process (1 of 4)
Preoperative
"I'd now like to move on to ask you some questions about the patient pathway."

4. “If a 63 year old woman has agreed in clinic to undergo a left hemicolecction for an early tumour, how would she be prepared and optimised for her operation?”

- Do you have a preoperative assessment clinic? Who leads this service?
- How do patients know what to expect? [Level of preop counselling.]
- Do patients receive no bowel prep, carbohydrate drinks, specific counselling about likely recovery, advice about discharge planning, etc?

"Sometimes specific medication needs to be stopped for surgery, such as clopidogrel, or patients need a valid Group and Save before theatre. In your experience, are local arrangements for such things highly reliable, or do patients often end up having surgery while still taking medication they should have stopped or having last minute blood tests?"

"If there is concern about a patient's fitness for surgery, how is this assessed and optimised?"

- Protocols / anaesthetist with specialist interest? Ad hoc referral? Who refers?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
</table>
Clinical process (2 of 4)
Operative technique

5. “Across the colorectal department, for elective colonic procedures, what proportion of patients do you estimate undergoes laparoscopic surgery?”

“On the whole, are most cases completed laparoscopically, or is there a high conversion rate?”

- Are there defined selection criteria for laparoscopic surgery?

SURGEONS: “Do you and your consultant colleagues tend to have similar set-ups in theatre, such as patient positioning, staplers and energy devices?”

SURGEONS: “Is the anaesthetic for colorectal patients standardised, or does it vary from anaesthetist to anaesthetist? E.g. agents used, TIVA, monitoring?”

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
</table>
Clinical process (3 of 4)
Postoperative (1 of 2)
Routine care

6. “After surgery, how would a fit 60 year old male who has undergone a straightforward right hemicolectomy be managed once returned to the ward?”

- Who sees them?
- How often?
- Are there specific goals of care? Such as?
- What elements of ERP? E.g. no NGT, no drains, patient-led early feeding, early structured mobilisation, regular laxatives.
- Involvement of MDT - nurses allowed to make specific management decisions, physiotherapists daily, etc.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
</table>
Clinical process (4 of 4)
Postoperative (2 of 2)
Complication detection and management

7. “Unfortunately some patients may experience complications after surgery. What systems are in place to detect if a patient becomes unwell?”

- Early warning scores & escalation protocols
- Any specific training for complication recognition for nursing / junior medical staff?

“How do you think the provision of supporting services, such as theatre time, intensive care and radiology, matches the need for such care for your patients?”

- All services you would like available?
- Interventional radiology available? 24 hours?
- Speed of access and input when required?
- In and out of hours?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Daily v/v</td>
<td>Daily v/v. Team easily bleeped. MEWS. Reasonable ITU support. No / patchy interventional radiology.</td>
<td>1-2x review per day. Team check in before leaving. MEWS with some protocols for escalation. Good ITU support. Good in-hours interventional radiology, some out of hours.</td>
<td></td>
<td></td>
<td></td>
<td>Twice daily review. MEWS / protocols for concern. Nurses empowered to bypass junior staff. 24hr interventional radiology. Excellent ITU support</td>
</tr>
</tbody>
</table>
**Institutional process (1 of 5)**
Standardisation / clinical pathways

"I'd now like to move on to ask some questions about the wider running of the department."

8. “Do you and your colleagues / the colorectal consultants all manage their patients the same way, or are there differences in your / their routine management plans?"

“Are any there any formalised or standardised policies and procedures for colorectal patients?”

- Just ERP, or others eg complication management, emergency care, etc.

“Do staff adhere to protocols or are there often significant deviations?”

- How do staff at all levels (FY/HCAs) know about details of ERP / etc?

“What is the attitude to standardisation of care in your unit?”

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
</table>
Institutional process (2 of 5)
Communication and collaboration between professional groups

9. “At times, the management plan for a patient is changed on the ward round, or overnight there may be significant changes in a patient’s condition. How do doctors and nurses coordinate patient care to ensure there is good information exchange on the ward?”

- Ward rounds for consultants & SpR / SHOs - with/without nurse?
- Handovers - joint or separate? Weekly joint meeting?

“How do nursing staff, and junior and middle-grade medical staff contact consultants to discuss patient care?”

- See each other frequently on ward / theatre / clinic
- Via switchboard or exchange mobile numbers

“Do the leaders of the surgical, anaesthetic and nursing services meet regularly to discuss problems with the service and how to improve patient care? How often?”

- Do leaders seek input from front line and feed back after meetings?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>On ward, read plans in notes.</td>
<td>Usually discuss patients, but non-consultant WR with nurse &lt;50%..</td>
<td>Usually non-consultant WR together &gt;75%..</td>
<td>All non-consultant WR together. Joined by physio / pharmacist. Consultants easily contacted via mobile. At least 3-monthly meetings. Excellent vertical communication.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Only nurse on consultant WR. Consultants often uncontactable. No routine meetings. Poor communication vertically.</td>
<td>Consultants available only through juniors / switchboard. 6-monthly or fewer meetings between leaders.</td>
<td>Consultants generally available via mobile / switchboard. 3-6 monthly meetings between leaders. Reasonable vertical comms.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Institutional process (3 of 5)
Clinical leadership and culture

10. “I’d like to ask about the goals and priorities of the unit. In your opinion, do the clinical and non-clinical leaders of the department have a clear long-term plan with a strong emphasis on excellence and high quality care, or is the primary focus on meeting day-to-day priorities and targets?”

- Do you feel one predominates? Differences between clinicians and managers?

“Within the unit, is there a sense of shared purpose towards common goals, or do individual members of staff primarily focus on their own priorities?”

- Do you get the impression that all members of the unit feel they belong to the team?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focused on short-term goals e.g. breach-times, finances. Lip-service to quality, secondary to targets. No clear shared goals.</td>
<td>Balance of short &amp; long term, quality vs targets, but tend to discuss / focus on targets more. Some common purpose.</td>
<td>Balance of quality and targets, with good emphasis on long term. Generally shared vision / goals.</td>
<td>Primary focus on quality and long-term, achieve targets through high standards. Shared vision on clear objectives.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Institutional process (4 of 5)
#### Attitudes to safety and adverse events

11. “I’d now like to ask about the local approach to adverse incidents. Suppose a patient having a hernia repair was marked on the wrong side for surgery, and this was only detected after the patient had been anaesthetised in theatre, but before surgery was started? What would be the attitude to such an incident?”

- Open and blame-free, or element of fear of being blamed / punished?
- Would such an error tend to be considered as arising because somebody got it wrong, or as a failure of the system?

“Sometimes patients may not be prescribed appropriate DVT prophylaxis but no harm occurs. On other occasions, a similar error may be associated with an adverse outcome for a patient. Is there any difference in the local handling of these kinds of errors?

“When incident report forms are completed, do they usually result in a response and some form of action? Or is there a feeling that nothing happens?”

- How are the lessons learned disseminated and fed back to staff at all levels?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Culture of blame / punishment, lip service to openness. Errors often not reported. Individuals at fault, minimal systems thinking. AIRs 'black hole'.</td>
<td>Primarily blame culture, some move away from. Individuals more than system failures. Errors not routinely learned from.</td>
<td>Open but sometimes blame / fears. More systems than individuals. Errors analysed, lessons not always disseminated.</td>
<td>Very open, no blame unless gross misconduct. Errors are human. Strong system approach. Lessons learned fed back.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Institutional process (5 of 5)
Outcomes assessment and quality improvement

12. “Across the department, what performance data is collected?”

- Who by? How often?

“How are data such as length of stay, mortality or complication rates used within the department?”

- Outcomes considered individually and privately
- Outcomes publicised internally / publicly?
- Compared against historical local practice / national standards / research standards

“Are performance data considered among consultants and managers, or are they fed back to all involved in patient care eg theatres, ward staff?”

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Only admin data, no registries. Individuals assumed to monitor own results. No outcomes review or feedback.</td>
<td>Some additional data collection, not department wide. Outcomes may be reviewed annually, only among leaders, not publicised.</td>
<td>Routinely submit to registries (no dedicated staff for this). Meet to discuss outcomes 6-12 monthly, not well publicised to all staff.</td>
<td>Thorough prospective data (e.g. 100% NBOCAP). At least 6-monthly review &amp; feedback, transparent outcomes reporting to all staff.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Interviewee opinion

"We are nearing the end of the interview now. I'd like to take this opportunity to ask for your insights into the management of colonic surgery patients."

13. “What do you think a colorectal department can do to reduce length of stay?”
Fact-gathering

14. “I would also like to ask a few specific factual questions about certain aspects of the hospital you work in.”

“Would you consider the mix of disease and patients operated on in your department to be typical of an average hospital, or do you have areas of specialist practice, such as recurrent cancer or IBD? **NB colon vs rectum.**”

<table>
<thead>
<tr>
<th>Question</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has the management of elective colonic surgery changed significantly since the beginning of 2011?</td>
<td>YES / NO</td>
</tr>
<tr>
<td>Overall, is reducing length of stay a priority within the colorectal department?</td>
<td>YES / NO</td>
</tr>
<tr>
<td>How do you think length of stay for elective colonic resections in your unit compares nationally?</td>
<td>LONGER / AVERAGE / SHORTER</td>
</tr>
<tr>
<td>How many colorectal consultants work in your department?</td>
<td></td>
</tr>
<tr>
<td>How many middle grades? (SpR, Staff Grades)</td>
<td></td>
</tr>
<tr>
<td>How many full-time trained nurses do you have on the colorectal ward during the day?</td>
<td></td>
</tr>
</tbody>
</table>

“That concludes our interview. Again, I am very grateful for your time and assistance with this study. We hope to gain a better understanding of surgical care as a result of this study and plan to share our early results with those who have participated. Would you like us to contact you with our findings? If so, how would it be best to get in touch? Thank you once again. Good bye.”

<table>
<thead>
<tr>
<th>Question</th>
<th>Options</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Would like to be sent findings?</td>
<td>YES / NO</td>
<td></td>
</tr>
<tr>
<td>Contact details</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix H

Study 6: quantitative analysis

supplement

H.1 Appendix overview

This appendix contains additional results of the quantitative analysis of interview data gathered in study 6, discussed in chapter 11.

H.2 Distribution of overall scores

Frequency histograms of individual (figures H.1 and H.2) and total question scores (figure H.3) are provided below.
Figure H.1: Frequency histograms of scores for questions 2 to 7.
Figure H.2: Frequency histograms of scores for questions 8 to 12.
Figure H.3: Frequency histogram of total interview scores.
Appendix I

Third party permissions

I.1 Appendix overview

This appendix contains proof of permission to reproduce published material.
Subject: Thank you for your RightsLink / Elsevier transaction
From: "Copyright Clearance Center" <rightslink@marketing.copyright.com>
Sent: 12/04/2013 10:02:40
To: benbyrne@doctors.org.uk

To view this email as a web page, go here.
Do Not Reply Directly to This Email
To ensure that you continue to receive our emails,
please add rightslink@marketing.copyright.com to your address book.

RightsLink

Thank You For Your Order!

Dear Mr. Ben Byrne,

Thank you for placing your order through Copyright Clearance Center's RightsLink service. Elsevier has partnered with RightsLink to license its content. This notice is a confirmation that your order was successful.

Your order details and publisher terms and conditions are available by clicking the link below:
http://e100.copyright.com/CustomerAdmin/PLF.jsp?ref=b8faad61-2956-4fbf-a05f-8d433b4a2260

Order Details
Licensee: Ben Byrne
License Date: Apr 12, 2013
License Number: 3126400817567
Publication: The Lancet
Title: Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma
Type Of Use: reuse in a thesis/dissertation
Total: 0.00 GBP

To access your account, please visit https://myaccount.copyright.com.

Please note: Online payments are charged immediately after order confirmation; invoices are issued daily and are payable immediately upon receipt.

To ensure we are continuously improving our services, please take a moment to complete our customer satisfaction survey.

B.1v4.2

+1-877-622-6543 / Tel: +1-978-646-2777
customercare@copyright.com
Subject: RE: Permission to reuse figure from 'Patient Safety' by Charles Vincent
From: "Wiley Global Permissions" <permissions@wiley.com>
Sent: 10/08/2015 10:22:40
To: "benbyrne@doctors.org.uk" <benbyrne@doctors.org.uk>

Dear Ben Byrne,

Thank you for your request.

Permission is granted for you to use the material requested for your thesis/dissertation subject to the usual acknowledgements and on the understanding that you will reapply for permission if you wish to distribute or publish your thesis/dissertation commercially. You must also duplicate the copyright notice that appears in the Wiley publication in your use of the Material.

Permission is granted solely for use in conjunction with the thesis, and the material may not be posted online separately.

Any third party material is expressly excluded from this permission. If any of the material you wish to use appears within our work with credit to another source, authorisation from that source must be obtained.

Yours Sincerely,

Rebecca Cook
Permissions Assistant
John Wiley & Sons Ltd
The Atrium
Southern Gate, Chichester
West Sussex, PO19 8SQ
UK

From: Ben Byrne <mailto:benbyrne@doctors.org.uk>
Sent: 30 July 2015 14:34
To: onlinelibrary sales
Subject: Permission to reuse figure from 'Patient Safety' by Charles Vincent

To whom it may concern,

I would like to include an image, based upon a figure in the book Patient Safety by Charles Vincent, in my PhD thesis at Imperial College London.
Do you grant your permission for me to do this?

Best wishes,

Ben Byrne
Clinical Research Fellow