Structural and Care Process Improvement of Ward-based Postoperative Care to Optimise Surgical Outcomes

Philip Han-Lung Pucher MD BSc MRCS

Thesis submitted to
Imperial College London
for the degree of
Doctor of Philosophy

Department of Surgery
Division of Surgery and Cancer
Imperial College London

July 2014
For my parents,
who have been ever loving, involved, and supportive, every step of the way.

And for Amy… obviously.
Acknowledgements

A thesis is no small undertaking. As with most things in life, however, the task is made easier through the help and support of friends, family, and colleagues along the way. Despite the individualistic nature of any academic award, there is no doubt that it is only with the help of many around me that I was able to accomplish this work. As such I can make no apologies for the lengthy acknowledgements to follow.

I would like to first give thanks to my supervisors: Professor Ara Darzi, for giving me the opportunity to pursue this work, regular feedback (always helpful, though often acerbic) throughout, and continued support within the department. These have been three great years of both personal and academic development upon which I will no doubt fondly reflect, on returning to clinical practice. Mr Rajesh Aggarwal, for taking a chance on me as a junior SHO five years ago, and convincing me to do a PhD in the first place. I am truly grateful for your unparalleled guidance, feedback, and mentorship these past years. I’m not sure one could expect any more from a supervisor, and I look forward to our productive friendship continuing in future.

To my many colleagues and fellow research fellows, your friendship has made the time in research not only bearable, but some of the most fun anyone has ever had whilst sat behind a desk! My second home, room 1029, home of jubs past (Chris Nicolay and Dan Cohen) and present – Mark Bagnall, George Bouras, Ann-Marie Howell, Jeremy Huddy, Max Johnston, and Laura Muirhead. Three weddings, an engagement, and three children in three years is some strong work. You will all be sorely missed! Pritam Singh, we will have to meet up regularly to make sure Bonne Bouche’s breakfast business doesn’t go bankrupt. As well as the many other research fellows and lecturers, too many to name.
Though not active researchers themselves, certainly no research within the department would be possible without Sejal Jiwan (a truly fine General Information Management Person), Karen Kerr, and Beth Janz. Thank you for everything – well, except the many VIP tours.

For the friends and family who have kept me sane and grounded throughout, even in the early days when I could not fathom I was ever going to finish a PhD. Specifically I would like to thank my parents, Martin and Corinna, and my brother Andrew, and Drs Alan Ward and Mark Almond.

Finally, for my beautiful wife Amy, who has incredibly put up with us living apart for almost the entire duration of this thesis, separated by train, car, and boat journeys. Despite being annoyed by my constant work-related stress, emails, and deadlines, my not infrequent travels abroad, and putting up with this all whilst cutting her own path through the thicket of surgical training knowing that I plan to shamelessly follow in her footsteps, we still managed to get married and treasure every minute we have together – this is for you.
Declaration of originality

I declare that the experiments described in this thesis were undertaken by me, under the supervision of Professor Lord Ara Darzi and Mr Rajesh Aggarwal. This thesis, and the work described herein, is my own except where explicitly referenced, and has not been previously submitted for a higher degree.

Philip Han-Lung Pucher

July 1, 2014
Copyright declaration

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.
Abstract

Much of the variation seen in surgical outcomes can be explained by differences in the quality of management of post-operative complications and ward-based care. The surgical ward round (WR) is critical to determining post-operative care and serves as the primary point of interaction between clinician and patient. Despite this, it is an area not subject to training or assessment at present.

This thesis demonstrates the high degree of variability which exists in the conduct of WRs. It establishes the link between suboptimal patient assessment and increased risk of preventable post-operative complications. These place patients not only at risk of short-term deterioration, but result in reduced long-term survival as well.

In order to quantify WR quality, a novel assessment tool has been developed and validated within a simulated environment. Ward simulation is a nascent branch of simulation which has been only preliminarily explored to date. A simulation environment was developed to take advantage of the known benefits of simulation such as controllability, reproducibility, and recordability, whilst maintaining a high level of fidelity and realism. An evidence-based curriculum for surgical WR training was designed and implemented in a simulation-based course. By focusing on structured generic processes of patient assessment and management, this resulted in significant improvement of trainee performance in routine WRs.

To ensure standardised and optimum management of specific conditions, checklists have proven themselves to be of great value in a number of surgical and medical disciplines. Surgical complications are common, yet their management often suboptimal. As part of this thesis, evidence-based protocols for the management of the six most common complications were designed and validated. The implementation of these in a simulation-
based randomised, controlled trial has resulted in greatly increased adherence to evidence-based standards of care, as well as improved communication and clinician performance.

This thesis explores the variance currently present in surgical ward rounds, and the potentially grave consequences of this for patient outcomes. To date, WRs have been one of the last areas of surgical care still dependent on the Halstedian principle of experiential learning alone. The tools have now been developed with which to assess, improve, and standardise critical structures and care processes in the assessment and management of the post-operative surgical patient. Future implementation of these and integration into surgical curricula will benefit clinician training, patient care, and surgical outcomes alike.
# Table of Contents

List of Tables and Figures ........................................................................................................ 15

Peer-Reviewed Publications Originating from this Thesis ............................................. 18

Presentations to Learned Societies ....................................................................................... 19

Aims of this Thesis .................................................................................................................. 21

Chapter 1: Literature Review .................................................................................................. 22

1.1 Introduction ......................................................................................................................... 23
   1.1.1 Patient safety and medical error ................................................................................. 23
   1.1.2 Error analysis and prevention ................................................................................. 26
   1.1.3 Beyond the operating theatre .................................................................................... 32

1.2 Measuring quality of care .................................................................................................... 34
   1.2.1 Volume and outcome ............................................................................................... 34
   1.2.2 Failure to rescue .................................................................................................... 37

1.3 Enhancing surgical performance outcomes through process-driven care ............ 42
   1.3.1 Introduction .............................................................................................................. 42
   1.3.1.1 Failure to rescue – variants and adaptations ....................................................... 43
   1.3.1.2 Structure, process, and outcome ....................................................................... 45
   1.3.1.2 Enhanced recovery protocols ............................................................................. 49
   1.3.2 Methods .................................................................................................................. 50
   1.3.2.1 Search strategy ................................................................................................... 50
   1.3.2.2 Selection criteria ............................................................................................... 51
   1.3.2.3 Quality analysis .................................................................................................. 51
   1.3.2.4 Statistical analysis .............................................................................................. 56
   1.3.3 Results ...................................................................................................................... 56
   1.3.3.1 Enhanced recovery protocols ............................................................................. 58
   1.3.3.2 Failure to rescue ............................................................................................... 59
   1.3.3.3 Statistical analysis .............................................................................................. 70
   1.3.4 Discussion .................................................................................................................. 72
   1.3.5 Conclusion ................................................................................................................ 76

1.4 Postoperative in-hospital morbidity impairs long-term patient survival: systematic
   review and meta-analysis .................................................................................................... 78
   1.4.1 Introduction .............................................................................................................. 78
   1.4.2 Methods .................................................................................................................. 80
   1.4.2.1 Search strategy ................................................................................................... 80
   1.4.2.2 Selection criteria ............................................................................................... 81
   1.4.2.3 Assessment of methodological quality .............................................................. 81
   1.4.2.4 Statistical analysis .............................................................................................. 81
Chapter 2: Identifying quality markers and improvement measures for ward-based surgical care

2.1 Introduction ........................................................................................................... 112
2.2 Methods ............................................................................................................... 114
  2.2.1 Subjects .......................................................................................................... 114
  2.2.2 Setting .............................................................................................................. 115
  2.2.3 Interview process ............................................................................................ 115
  2.2.4 Data analysis ................................................................................................... 116
2.3 Results .................................................................................................................. 116
  2.3.1 Current problems in ward-based care ............................................................. 119
  2.3.2 Defining the surgical ward round ................................................................. 121
  2.3.3 Required skill set ........................................................................................... 122
  2.3.4 Identifying quality markers for ward rounds ............................................... 122
  2.3.5 Improving surgical ward round quality ....................................................... 125
2.4 Discussion ............................................................................................................. 128
2.5 Conclusion .......................................................................................................... 131

Chapter 3: The simulated ward environment and assessment tools for ward-based surgical care

3.1 Development of a simulated environment for ward round assessment and training ........................................................................................................ 134
  3.1.1 Introduction ................................................................................................... 134
Chapter 4: Surgical ward round quality and impact on variable patient outcomes ................................................................. 177

4.1 Introduction ..................................................................... 178
4.2 Methods .......................................................................... 180
  4.2.1 Setting ......................................................................... 180
  4.2.2 Patients ........................................................................ 180
  4.2.3 Study protocol ............................................................... 181
  4.2.4 Patient outcomes .......................................................... 182
  4.2.5 Statistical analysis .......................................................... 182
4.3 Results ............................................................................. 183
  4.3.1 Patient demographics ..................................................... 183
  4.3.2 Ward round quality ......................................................... 185
  4.3.3 Complications and errors ............................................... 187
  4.3.4 Outcome analysis ........................................................... 189
4.4 Discussion ....................................................................... 191
4.5 Conclusion ....................................................................... 195
Chapter 5: Development and implementation of a ward round training curriculum

5.1 Development of an evidence-based curriculum for training of ward-based surgical care

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1.1</td>
<td>Introduction</td>
<td>197</td>
</tr>
<tr>
<td>5.1.2</td>
<td>Methods</td>
<td>199</td>
</tr>
<tr>
<td>5.1.2.1</td>
<td>Predevelopment analysis</td>
<td>199</td>
</tr>
<tr>
<td>5.1.2.2</td>
<td>Curriculum design</td>
<td>200</td>
</tr>
<tr>
<td>5.1.3</td>
<td>Results</td>
<td>200</td>
</tr>
<tr>
<td>5.1.3.1</td>
<td>Predevelopment analysis</td>
<td>200</td>
</tr>
<tr>
<td>5.1.3.2</td>
<td>Curriculum development</td>
<td>202</td>
</tr>
<tr>
<td>5.1.3.3</td>
<td>Knowledge and confidence</td>
<td>205</td>
</tr>
<tr>
<td>5.1.3.4</td>
<td>Didactic session</td>
<td>207</td>
</tr>
<tr>
<td>5.1.3.5</td>
<td>Simulated Ward Round</td>
<td>207</td>
</tr>
<tr>
<td>5.1.3.6</td>
<td>Assessment Tools</td>
<td>209</td>
</tr>
<tr>
<td>5.1.3.7</td>
<td>Debriefing feedback</td>
<td>209</td>
</tr>
<tr>
<td>5.1.3.8</td>
<td>Cost of curriculum</td>
<td>209</td>
</tr>
<tr>
<td>5.1.4</td>
<td>Discussion</td>
<td>210</td>
</tr>
<tr>
<td>5.1.5</td>
<td>Conclusion</td>
<td>211</td>
</tr>
</tbody>
</table>

5.2 Ward simulation to improve surgical ward round performance: a randomised controlled trial

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.1</td>
<td>Introduction</td>
<td>213</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Methods</td>
<td>215</td>
</tr>
<tr>
<td>5.2.2.1</td>
<td>Subjects</td>
<td>215</td>
</tr>
<tr>
<td>5.2.2.2</td>
<td>Study design</td>
<td>215</td>
</tr>
<tr>
<td>5.2.2.3</td>
<td>Simulation-based curriculum</td>
<td>217</td>
</tr>
<tr>
<td>5.2.2.4</td>
<td>Simulated ward environment</td>
<td>218</td>
</tr>
<tr>
<td>5.2.2.5</td>
<td>Assessment methods</td>
<td>219</td>
</tr>
<tr>
<td>5.2.2.6</td>
<td>Statistical analysis</td>
<td>221</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Results</td>
<td>221</td>
</tr>
<tr>
<td>5.2.3.1</td>
<td>Study participants</td>
<td>221</td>
</tr>
<tr>
<td>5.2.3.2</td>
<td>WR performance</td>
<td>221</td>
</tr>
<tr>
<td>5.2.3.3</td>
<td>Comparison to senior trainee cohort</td>
<td>226</td>
</tr>
<tr>
<td>5.2.3.4</td>
<td>Clinical errors</td>
<td>226</td>
</tr>
<tr>
<td>5.2.3.5</td>
<td>Confidence questionnaire</td>
<td>226</td>
</tr>
<tr>
<td>5.2.3.6</td>
<td>Course feedback</td>
<td>226</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Discussion</td>
<td>229</td>
</tr>
<tr>
<td>5.2.5</td>
<td>Conclusion</td>
<td>233</td>
</tr>
</tbody>
</table>

Chapter 6: Development and implementation of checklists for the management of postoperative complications

6.1 Development of surgical care checklists to optimise patient care following postoperative complications

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1.1</td>
<td>Introduction</td>
<td>235</td>
</tr>
</tbody>
</table>

12
Chapter 7: Conclusion ........................................................................................................... 271

7.1 Background synopsis ..................................................................................................... 272
7.2 Summary of findings ...................................................................................................... 275
7.3 Impact of this work ......................................................................................................... 280
7.4 Limitations ..................................................................................................................... 282
7.5 Future work .................................................................................................................... 284
7.6 Concluding remarks ....................................................................................................... 287

List of References ................................................................................................................. 288

Appendix ............................................................................................................................... 316

Appendix 1. Semi-structured interview protocol ................................................................. 317
Appendix 2. SWAT assessment tool................................................................. 319
Appendix 3. W-NOTECHS assessment tool...................................................... 320
Appendix 4. Final POSCH checklists for postoperative complications, optimised for use in Imperial College London, UK hospitals................................................. 321
List of Tables and Figures

Table 1.1  Structure, process and outcome variables .................................................. 46
Table 1.2  Jadad quality score for randomised trials ................................................... 53
Table 1.3  Demographics of ERP studies ................................................................. 61
Table 1.4  Results of ERP studies .............................................................................. 63
Table 1.5  Demographics of FTR studies ................................................................. 65
Table 1.6  Structure and process factors affecting failure to rescue .......................... 67
Table 1.7  Demographic details of included studies ............................................... 85
Table 1.8  Outcomes of included studies comparing cohorts with or without postoperative morbidity .............................................................. 88

Table 2.1  Participant demographics .......................................................................... 118
Table 2.2  Interviewees’ opinions on current ward round practice ............................ 120
Table 2.3  Identified markers for ward round quality ............................................... 124
Table 2.4  Recommendations for ward round quality improvement ........................ 127

Table 3.1  Example costs for ward furnishing for simulated ward ......................... 147
Table 3.2  Checklist of generic assessment and management processes .................. 159
Table 3.3  Subject demographics ............................................................................. 163
Table 3.4  Fidelity questionnaire responses .............................................................. 171

Table 4.1  Patient demographics .............................................................................. 184
Table 4.2  Assessed sources of clinical information during ward rounds ................. 186
Table 4.3  Summary of preventable adverse and risk events ................................... 188

Table 5.1  Factors critical to ward round quality ....................................................... 201
Table 5.2  Evidence-based module design ............................................................... 204
Table 5.3  Surgical Ward care Assessment Tool results .......................................... 222
Table 5.4  W-NOTECHS results with individual component scores ....................... 225
Table 5.5  Results of course questionnaire ............................................................... 228

Table 6.1  Type and incidence rate of postoperative complications ......................... 243
Table 6.2 Results of casenote review................................................................. 246
Table 6.3 Delays and adverse events in management of complications............. 247
Table 6.4 Subject demographics........................................................................ 261
Table 6.5 Simulated ward round performance metrics....................................... 266
Table 6.6 Baseline ward round performance by scenario.................................. 265
Table 6.7 Users’ responses to use of checklists.................................................. 266

Figure 1.1 Reason’s “Swiss cheese” model......................................................... 28
Figure 1.2 Newcastle-Ottawa quality assessment scale for cohort studies............ 55
Figure 1.3 Flow diagram of literature search results............................................. 57
Figure 1.4 Meta-analysis for variables affecting FTR rates.................................. 71
Figure 1.5 Flow chart of literature search results............................................... 83
Figure 1.6 Meta-analysis of overall survival after postoperative complications...... 89
Figure 1.7 Meta-analysis of disease-free survival after all postoperative complications................................................................. 91
Figure 1.8 Funnel plot assessment of potential bias.............................................. 93
Figure 1.9 Stages of assessment tool development.............................................. 107

Figure 3.1 The simulated ward at Imperial College London................................. 140
Figure 3.2 Example simulated patient scenarios............................................... 143
Figure 3.3 Setup of the simulated ward............................................................. 151
Figure 3.4 Sample vignettes for the three simulated patient scenarios ............... 155
Figure 3.5 Study protocol overview.................................................................... 157
Figure 3.6 Percentage of assessed patient assessment tasks completed............... 165
Figure 3.7 Percentage of assessed patient management tasks completed............. 166
Figure 3.8 W-NOTECHS scores ....................................................................... 168

Figure 4.1 Effect size of ward round quality for error and complication rates........ 190

Figure 5.1 Curriculum flow chart....................................................................... 203
Figure 5.2 Pre- and post-course confidence questionnaire.................................. 206
Figure 5.3 Systematic approach to conduct of surgical ward rounds............... 208
Figure 5.4 CONSORT diagram......................................................................... 216
Figure 5.5 Course feedback form...................................................................... 220
Figure 5.6  Boxplot of performance for control and intervention groups for SWAT and W-NOTECHS scores

Figure 6.1  Iterative checklist development process

Figure 6.2  CONSORT study flowchart

Figure 6.3  Failures to adhere to critical processes in the management of postoperative complications
<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title of Paper</th>
<th>Journal</th>
<th>Volume and Issue</th>
<th>Pages/Publication Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Pucher PH, Aggarwal R, Singh P, Tahir M, Darzi A.</td>
<td>Identifying quality markers and improvement measures for ward-based surgical care: a semi-structured interview study.</td>
<td></td>
<td>[under review]</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Pucher PH, Aggarwal R, Almond MH, Darzi A.</td>
<td>Surgical care checklists to optimise patient care following postoperative complications.</td>
<td></td>
<td>[under review]</td>
<td></td>
</tr>
</tbody>
</table>
Presentations to Learned Societies


Surgical care checklists to optimise patient care following postoperative complications. [Poster]


Aims of this Thesis

1. To identify potential quality markers for the surgical ward round and methods for improvement through stakeholder consensus.

2. To assess the variability of surgical ward rounds in current practice and its effect on patient outcomes.

3. To develop and validate tools to measure ward round quality.

4. To develop interventions to improve trainees’ assessment and management of the surgical patient during the ward round.
Chapter 1

Literature Review
1.1 Introduction

1.1.1 Patient safety and medical error

The historical uncertainty of life or death in disease has in the past resulted in an obscuring of differences between unavoidable deteriorations in patient condition, and avoidable injury caused through errors in care. In recent decades, however, there has been a growing awareness within the medical community and the greater public alike of the unnecessary risks and dangers patients are exposed to throughout the course of any given stay in hospital. The maturation of patient safety as an independent research discipline has contributed greatly to the broader understanding of medical error and adverse events.

*Medical error* may be defined as a process of care that may be an act of commission or omission, resulting from a failure to act, a planned action which fails to be completed as intended, or the use of an inappropriate plan implemented to achieve a given aim.¹ An *adverse event* is an injury or complication caused by medical management rather than by the primary disease process itself.² Whilst not all errors lead to adverse events, and not all adverse events are caused by medical error, the need to reduce both is obvious.

Adverse events and iatrogenic injury have been long recognised, but it is only through the pioneering patient safety studies of the 1990s, which sought to quantify their nature and incidence, that they became an acknowledged problem. Brennan and Leape’s seminal 1991 Harvard Medical Practice Study assessed 31,429 medical records of acute inpatients for all specialties other than psychiatry. These were screened by multiple physicians according to a defined search strategy to identify adverse events, defined as an injury caused through medical management rather than by the underlying disease, or negligence, defined as care falling below the standard expected of the physician community. The results of this study suggested several key findings – not only were rates of negligence and adverse events higher
than expected, but rates of adverse events for surgical patients were significantly higher still, when compared to other specialties. Whereas between 20 – 31% of non-surgical patients experienced negligent care, resulting in an adverse event incident rate of 3 – 4%, for surgical specialties (urology, neurosurgery, vascular, cardiothoracic, orthopaedic, and general surgery) these ranged between 19 – 39% and 4 – 16%, respectively. Overall, almost 70% of adverse events were found to have been directly caused through human error.

Catalysed by Brennan and Leape’s study, similar results were soon reported internationally. The 1995 Quality in Australian Health Care Study adopted a similar protocol to the Harvard study, reviewing 14,179 admissions across 28 Australian hospitals. Its authors reported a 16.6% adverse event rate, with half of these judged to have “high preventability.” Many of these adverse events were classed as serious events, with 14% of all events resulting in permanent disability, and 5% resulting in patient death. As with the Harvard study, authors of the Australian study reported significant differences in adverse event rates between specialties, with the highest proportion of serious events occurring in the surgical disciplines – with general surgery, in particular, found to have the highest overall rate of serious (resulting in permanent disability or death) adverse events (13.8%). Subsequent reports from Denmark and Canada confirmed similar findings.

In the UK, patient safety experts Charles Vincent and Graham Neale completed a similar study of English surgical patients in 2001, reporting that greater than 1 in 10 patients experienced an adverse event whilst hospitalised. Of these events, over half (53%) occurred in general ward care, as opposed to other more complex environments such as the operating theatre, or anaesthetic room.

Following on from the expanding base of academic evidence, it was arguably the United States’ Institute of Medicine’s landmark 1999 report, To Err is Human, which subsequently brought medical error to the forefront of public and political consciousness. In this report,
the non-governmental organisation estimated that between 44,000 and 98,000 deaths – figures extrapolated from the results of the Harvard Medical Practice Study and subsequent similar study carried out in the states of Utah and Colorado11 – in the United States annually were attributable to preventable medical error, thus establishing it as one of the top ten causes of death for the American population. Commenting on this, Lucian Leape, one of the authors of the Harvard study, famously compared it to “the equivalent of three jumbo-jet crashes every two days,”12 implying not only the scale of the problem, but also the indifference with which medical error had been previously considered – were an airplane disaster to occur daily, the airline industry would almost certainly no longer be in business.

Addressing the issue of patient safety in the UK, the Department of Health (led by Chief Medical Officer Sir Liam Donaldson) in 1999 published its strategy document, An Organisation With a Memory.13 It painted a bleak picture of the NHS, with an estimated adverse event rate of 10% of all admissions, resulting in over 400 serious injuries or deaths annually, costing the NHS an estimated £4.4 billion annually in additional hospital costs and potential liability claims alone, not including wider human costs or losses in economic productivity. It identified a lack of systematic methods to identify and respond to error, with existing systems of incident recording that were fragmented and incomplete. As a result, this publication set out a number of recommendations which would transform the landscape of patient safety in the UK. Prominent among these were the establishment of a clear mechanism for incident reporting, addressing safety culture within hospitals, and creating a national agenda for patient safety-focused research, resulting in the creation of the National Patient Safety Agency (subsequently disbanded in 2012, with responsibilities now largely incorporated into the NHS Commissioning Board Special Health Authority) and National Reporting and Learning System, a centralised national database and reporting system for adverse events.14
The Institute of Medicine’s *To Err is Human*, together with the Department of Health’s *An Organisation With a Memory* in the UK, have formed prominent rallying points of an ongoing greater movement in healthcare towards defining and improving the quality of care that healthcare systems, organisations and physicians deliver. Along with the huge resulting annual additional costs for the sequelae of adverse events (estimated at over £4 billion annually in the UK, and greater than USD $30 billion in the United States), these prompted calls to formally establish organisational bodies to study, reduce and prevent errors and adverse events in health care. The mandate to ensure quality provision of health care and suggest changes to improve care and reduce errors and adverse events has since been taken up by such national agencies as the Care Quality Commission (CQC) in the United Kingdom and the Institute for Healthcare Improvement (IHI) and the Agency for Healthcare Research and Quality (AHRQ) in the United States. The monitoring and selection of indicators of quality of care, once nebulous, has become a field of study unto itself, in an attempt to objectively quantify the performance and improve both clinical outcomes and cost-effectiveness, and leading to the design interventions to improve care. Despite this, it is also important to recognise that many of the most-cited figures, such as the Institute of Medicine’s attention-grabbing 98,000 annual deaths, are extrapolated from the relatively modest evidence of a small number of retrospective studies, rather than true prospective or actual population-level data, and further research remains required.

1.1.2 Error analysis and prevention

In order to reduce, or prevent, medical errors, it is first necessary to understand how they occur. In the case of a number of unfortunate, and catastrophic, medical errors which have occurred in the NHS over the past decades, resulting public inquiries have provided some insight. The Bristol Royal Infirmary Inquiry,15 was set up in 1998 to investigate a higher-than-
expected mortality rate in infants undergoing cardiac surgery in the preceding decade. Published in 2001, it was one of the most detailed investigations into NHS practice ever conducted, and contributed significantly to the wider reform of NHS services in conjunction with other major patient safety publications at the time. Rather than finding a single action, process, or staff member at fault, it placed blame upon a negative institutional culture, staff shortages, a lax approach to clinical safety, and a lack of external monitoring.

Similar findings have been reported by various subsequent inquiries, regardless of the nature of the event – whether it relate to a single patient, such as the fatal intrathecal injection of vincristine in a patient in Nottingham in 2001,\textsuperscript{16} or an entire Trust, such as the investigation into high mortality rates at Mid Staffordshire NHS Foundation Trust (resulting in the 2013 Francis Inquiry and the follow-up Berwick Report).\textsuperscript{17,18} Through these tragedies, spread across four decades of medical progress, runs a common theme: catastrophic error resulting from the failure of a system as a whole, rather than of a single procedure or person.

The concept of systems error in medicine has been long recognised. Perhaps its best-known exponent has been psychologist and human factors expert James Reason. Reason’s much-cited “Swiss cheese model” describes a system’s layers of defences against accidents or adverse events, each made porous by potential active or latent errors (figure 1.1).\textsuperscript{19} In a functioning system, these redundant layers work to prevent potential risks from turning into actual adverse events. When the system fails, however, successive barriers fall away, allowing “holes” within the “cheese” to align, allowing hazards to transform into losses.
Fig 1.1 Reason’s “Swiss cheese” model (from: Reason 1997)
Where failure results in an adverse event, it is imperative that a functioning system responds through analysis of the event, its cause, and institute preventative measures. This typically entails a cycle of investigation, data gathering, identification of care delivery problems and contributing factors, as described in Vincent’s London Protocol for Systems Analysis of Clinical Incidents. Specifically, it describes the assessment of three domains in the organisational accident causation model, adapted from the work of Reason. These relate to organisation and management culture (latent failures), contributory factors influencing practice (error and violation producing conditions), and care delivery problems (active failures). Such a systems-based approach to error and safety has been increasingly implemented in patient safety research.

Rather than focusing on a single intervention or process, such as the decision whether to prescribe or omit a given drug, a systems approach considers interactions between individual processes as the result of the above domains (organisation and culture, contributory factors, and care delivery). Once specific problems – failing layers within the “cheese” – are identified, interventions to reduce error, addressing the systems problem as a whole, may be considered.

Much has been made of the error prevention systems and safety cultures present in other high-technology industries such as aviation or the nuclear industry, and how they might be applied to medicine. In learning from past failures, such as air disasters or nuclear accidents, these professions have developed comprehensive cultures of safety and risk management, resulting in reliable service delivery despite their complex work environments. Some of the lessons learnt from these and other professions are now being applied in a surgical context. Team training and handover protocols derived from aviation or Formula One racing models, for example, have been shown to improve human factors performance of surgical teams. Catchpole et al, for example, reported the results of a multicenter study in which surgical teams completed a multi-session team training intervention, in which
training was provided by both human factors experts and team trainers from the aviation industry. Though this study only examined immediate post-intervention performance, rather than long-term skill retention, the authors reported that following the intervention, surgical teams displayed better communication and teamwork, and completed a significantly higher number of perioperative time-outs, briefings, and debriefings.

The generalisable nature of human factors across different professions has been highlighted by the work of Flin and colleagues. Parallels between the systems issues present in medicine and the oil drilling industry, for example, have been highlighted.\(^{28}\) Non-technical skills, complex technical tasks, production-focused organisation, and a male dominated workforce are equally present in both industries – as are the disasters which may results from systems failure. In her paper, Flin highlights the failings which led to the 2010 loss of the Deepwater Horizon oil drilling platform – weaknesses in the team, leadership, decision making, and safety culture of the rig\(^{28}\) – all issues which have been named as root causes of never events within the NHS, as well.\(^{29}\)

Further prominent examples of practice borrowed from other industries are the growing use of simulation and checklists, particularly in surgery. Surgical simulation and its applications to training will be discussed in greater detail in later chapters of this thesis. However, the growing adoption of surgical simulation has now largely supplanted the historical Halstedian method of training, in which surgical trainees were expected to acquire and practice operating skills through repeated procedures on patients (in what has colloquially been called the “see one, do one, teach one” approach). The use of simulation for surgeons’ technical skills training has provided an effective means for surgical trainees to practice basic operative skills in a safe environment whilst eliminating patient risk. Modern simulators may take many forms, depending on the task, skill, and level of complexity to be approximated, ranging from benchtop silicon models to box trainers, virtual reality simulations, and real tissue (animal or cadaver models). The addition of simulation-based
curricula continues to revolutionise the modern model of surgical training, with trainees now able to develop skills and demonstrate appropriate levels of competency prior to undertaking procedures on actual patients. With numerous studies demonstrating the efficacy of simulation training, more recent studies have demonstrated both its economic viability, and the effects of improved technical skill on patient outcomes as well.

Whereas simulation has played an important role in the improvement of surgical training, the aim of checklists has been to standardise the delivery of surgical care. Checklists are routinely used across all areas of the aviation industry ranging from manufacture, to pre-flight checks, to the management of in-flight crises. Noted surgical safety expert Atul Gawande comments on the universal utility of checklists in his bestselling 2010 book, The Checklist Manifesto. Beyond aviation, and the nuclear industries, he finds that checklists could be used to assure the standardised practice of any system, regardless of type. As an example, he cites the ability of a national restaurant chain (“The Cheesecake Factory”) to produce the same food, to the same quality standards, across the US regardless of restaurant location, size, or staffing, through the use of checklists to guide production. Adapted to the medical environment, the use of checklists to ensure completion of critical tasks, such as the administration of preoperative antibiotics, has seen significant improvements in clinical outcomes as a result.

One way in which the use of checklists may aid the standardisation of practice is through the limitation of particular types of error. In the analysis of the original Harvard Medical Practice study, Brennan, Leape, and colleagues also considered the types of adverse events which occurred, in addition to their incidence. Subsequent investigators, such as Vincent and colleagues in their analysis of errors within the NHS, applying a similar methodology. In both studies, authors noted the preponderance of “errors of omission,” errors which occurred due to failure to complete a task (e.g. failure to prescribe medication), compared
to “errors of commission,” from failed or incorrectly carried out processes (e.g. prescribing of incorrect medication). The prevention of omission errors through use of checklists has been specifically championed by Reason,37 and is discussed in greater detail in later chapters of this thesis.

1.1.3 Beyond the operating theatre

The growing use of team training, simulation, checklists, and other interventions throughout medicine and surgery have contributed to the great strides achieved in the improvement of patient safety in the past few decades. On a global level, the World Health Organisation (WHO) has specifically targeted patient safety as one of its primary objectives through the World Alliance for Patient Safety, launched in 2004.38 Through the Alliance, several notable successes have already been achieved, such as the global dissemination of the WHO Safe Surgery Checklist35 to improve intraoperative safety, and the recent launch of the Patient Safety Curriculum Guide39 which aims to integrate patient safety as a core theme the undergraduate training of future medical professionals.

Despite these achievements, major failures in healthcare provision continue to occur with worrying regularity. In the UK, the tragic failures at Mid Staffordshire NHS Trust remind us that much work remains to be done to improve care.17 A recent literature review, seeking to update the incidence of preventable adverse events based upon the methods cited in the Institute of Medicine’s To Err is Human (which was based on 1984 data), suggested that the incidence of failures has not decreased. In their study, James and colleagues reported that as many as 400,000 deaths were attributable to preventable adverse events in the US in 2011 alone.40

Research to increase patient safety in surgical patients has to date focused largely on the operative environment. Despite the improvement of technical training (e.g. incorporation of
simulation) and development of novel innovations for patient safety in theatre (e.g. WHO checklist, non-technical and team training), these have yet to be matched in the arenas of postoperative and ward-based care. Such inertia, however, belies the nature of medical error suggested by studies examining adverse events and medical error in surgery.

Examining again the results from the Harvard Medical Practice study, one of the sub-analyses also assessed the location in which adverse events took place. The majority of adverse events took place outside of the operating theatre (59% of all events), with the largest proportion taking place on the ward within the patient’s room (26.5%).

It was also within the patient’s room that the greatest proportion of negligent (i.e. preventable) adverse events took place (41.1%). These figures were later replicated by Neale and colleagues’ study of events within the NHS, wherein they found that 53% of all preventable adverse events occurred in general ward care. Unsurprisingly, the authors concluded that improvements in the quality of ward-based patient care were needed, and that further research to identify underlying causes of failures in patient care was necessary.

Subsequent research on adverse events and medical error has considered the types of problems associated with preventable deaths. Hogan et al reported the results of a retrospective case record review of 1000 deaths in NHS hospitals. Associated errors were variable, ranging from poor clinical monitoring (31%) to diagnostic errors (30%), and inadequate management (21%). This variability highlights the need for a systems approach to address failings in care – with such broad domains of error not addressable through single process interventions. Prior to approaching such interventions, however, there is a need to establish a reliable means to measure the quality of postoperative care.
1.2 Measuring quality of care

1.2.1 Volume and outcome

The selection and measurement of appropriate outcomes is a critical focus for the assessment of care quality and quality improvement. Variable morbidity and mortality rates have long highlighted the variable nature of care provided by different institutions.\textsuperscript{42} However, this has in the past represented too coarse a metric to appreciate potential underlying factors. In modern literature, hospital volume remains one of the most widely used surgical quality indicators. A growing body of literature suggests that patients treated by hospitals, and surgeons, with higher annual caseloads for a given procedure, fare better.\textsuperscript{43-45}

One of the first studies to examine the volume-outcome relationship was published by Luft and colleagues in 1979, in which the authors described a significantly reduced mortality rate after surgery in centres which performed 200 or more of a given procedure annually, compared to those with lesser surgical volume.\textsuperscript{46} Significantly, this relationship was found to hold only for complex procedures – such as major cardiac and vascular surgery, rather than for hip replacement or cholecystectomy – resulting in the authors’ call for the regionalization of complex surgical procedures. With other studies publishing similar findings,\textsuperscript{47, 48} centralisation was first formally addressed in the UK following the recommendations of the 1995 Calman-Hine report on improving NHS cancer services.\textsuperscript{49} Aiming to reduce the variability in cancer outcomes seen across different geographical regions of England and Wales, the Calman-Hine report, \textit{A Policy Framework for Commissioning Cancer Services}, resulted in the centralisation of cancer-related services including surgery into regional Cancer Units for common disease such as breast cancer, and larger Cancer Centres for rare or more specialised cancers, with specified minimum resources required for each to ensure
appropriate care could be provided, and ensure sufficient patient volume to ensure these were appropriately utilised.

Perhaps one of the influential studies on the relationship between operative volume and surgical outcome has been Birkmeyer and colleagues’ 2002 study published in the *New England Journal of Medicine*. This study utilised Medicare-derived data, thus covering the majority of patients undergoing surgery in the United States, for 1994 through 1999, assessing outcomes for patients aged 65 – 99 years for a selection of 14 surgical procedures selected for their relative complexity and significant risk of mortality. Approximately 2.5 million patient episodes, taking place at over 1000 hospitals across the United States, were included. Hospitals were divided by operative volume into quartiles, and rates of mortality for the various procedures compared across each quartile. With the data available from the Medicare database, authors were able to control for age, sex, race, year of the procedure, urgency of the procedure, patient comorbidities, and socioeconomic class in their regression model.

The results suggested clear differences in adjusted mortality between the hospital volume quintiles, as well as differences for the varying procedures. For carotid endarterectomy and nephrectomy, for example, the difference in mortality rates between high and low volume hospitals was limited (though still significant), with high volume hospitals reporting 0.3% and 1% lower morbidity rates, respectively. However, for more complex procedures such as pancreatectomy, for example, this difference was much greater – centres performing on average one or fewer pancreatectomies annually were associated with a 16.3% mortality rate, compared to 3.8% mortality in centres performing greater than 16 procedures a year, a more than four-fold difference.

Subsequent research has produced similar results, with a number of systematic reviews assessing the effect of surgical volume on outcomes for varied procedures including
oesophagectomy, pancreatectomy, and bariatric surgery, unanimous in their conclusion that hospitals in which a given procedure is performed (or disease managed) more frequently are likely to achieve better results than those with lower case volume.\textsuperscript{50-54}

Such findings continue to drive ongoing centralisation for complex procedures in both the UK as well as abroad, such as through the Leapfrog Group in the US, a consortium of over 150 health care organisations advocating common standards of practice with the goal of improving surgical outcomes.\textsuperscript{49,55} Despite this, some publications have suggested that the volume-outcome relationship may not be as strong as previously believed.\textsuperscript{56-58} Burns and colleagues reported two separate studies assessing results of colonic and rectal surgery, based upon UK data derived from the national Hospital Episodes Statistics (HES) database.\textsuperscript{57,58} Though it should be noted that in Birkmeyer’s study, colectomies were found to have only a modest difference in mortality between high- and low-volume centres (5.4 vs. 6.4%) and the difference in patient numbers (the studies by Burns et al included approximately 100,000 compared to the 2.4 million patients in the Birkmeyer study\textsuperscript{43}) raises the potential for type II error, Burns et al reported no significant differences in mortality between high- and low-volume surgeons. Negative results have also been reported by some authors investigating the volume-outcome relationship for more complex procedures such as oesophagectomy – Lin et al reported no significant association in their national study of Taiwanese patients,\textsuperscript{59} as have Rodgers et al in their study of oesophagectomy outcomes within the US National Inpatient Sample population.\textsuperscript{56}

Though these studies report findings which would appear to contradict earlier work on this topic, they have not negated the concept of association between volume and outcome, which has been largely accepted on the balance of available evidence, and has contributed significantly to current health policy.\textsuperscript{49,55} Rather, they highlight some of the limitations of the retrospective large-scale database analyses which constitute almost all current knowledge on the topic. The inherent pitfalls of utilising administrative or historical
databases mean that the number of available relevant confounding factors which may be controlled for vary according to the database used, often including only patient-related factors and excluding any hospital-related metrics. Furthermore, the quality of data is subject to the quality and completeness of data entry, oft performed by administrative staff rather than clinicians. Though more recent studies have sought to control for, and identify, additional significant hospital-related variables such as bed capacity, teaching status, or location, the underlying causes for this variation in mortality – be they failures in the pre-, intra-, or post-operative domain, or elsewhere – cannot be addressed through the consideration of volume and outcome alone.

1.2.2 Failure to rescue

The relationship between operative volume and outcome illustrates the impact that hospital-related factors may have upon patient outcomes. In their seminal 1992 study of 5,792 American Medicare patients, Jeffrey Silber and colleagues hypothesised that the factors underlying morbidity and mortality might be different from one another, whereas patient-related factors such as age and pre-existing disease are known play a significant role in a patient’s risk of developing a postoperative complication, Silber sought to investigate whether these or other factors also played a role in the likelihood of the same patient’s recovery or death following complications. Silber utilised a metric which he termed “failure to rescue” (FTR), defined as the number of patient deaths following an adverse event or complication, divided by the total number of patients who suffer an adverse event or complication – i.e. the “failure to rescue” of a patient who has developed a postoperative complication.

The authors examined two index procedures – elective cholecystectomy and transurethral prostatectomy – utilising data for 5,972 patients, admitted in 1985 and 1986, from the US
Medicare database, and performed multiple logistic regression analysis for both patient and hospital characteristics to measure association with FTR rates. This analysis revealed that whereas the likelihood of adverse events was strongly linked to nearly all patient variables, such as age, gender, disease severity, and any of five assessed comorbidities (such as diabetes), there was no significant relationship with any hospital variables, other than a small-effect association between percentage of surgeons with board certification (relative risk (95% confidence interval): 0.86 (0.8, 0.99). This was not the case, conversely, for FTR rates, wherein there was almost no association with patient factors (other than age, or presence of metastatic disease), but significant association with hospital-related factors including presence of surgical house staff, or percentage of board certified anaesthesists.

These findings demonstrated significant variation in hospitals’ abilities to “rescue” patients from complications or adverse events. Fully adjusted for factors such as patient variables and hospital case mix, Silber and colleagues concluded that the use of FTR rates might provide a more sensitive metric than mortality figures alone. In this manner, a hospital with a low rate of complications, for example, might still demonstrate a high FTR rate if failures in postoperative care were to occur. As suggested by Silber et al, such a situation might occur particularly in hospitals with an “easy” case mix, wherein a relatively low rate of complications might result in a low rate of adjusted mortality – with potential failings in care quality that might be detected through use of FTR thus going otherwise unmeasured.

Further studies have since expanded upon these findings, demonstrating generally poor or no correlation between hospital adjusted complication and mortality rates for other specialties and procedures. Arozullah et al reported the results of an analysis of 103,176 non-cardiac surgical patients derived from a prospectively maintained database of 123 US hospitals, assessing the relationship between pneumonia, as one of the most common postoperative complications, and mortality. Across all hospitals, authors found only a weak correlation between pneumonia rates and mortality (r = 0.21, p = 0.02), and no significant
correlation at all when hospitals were stratified by operative volume.\(^6^2\) Similarly, Shroyer et al reported results for 503,478 patients undergoing coronary artery bypass surgery, using data from the US Society of Thoracic Surgeons National Adult Cardiac Surgery Database.\(^6^4\)

Comparing risk-adjusted rates of mortality to the most common types of morbidity demonstrated weak correlations for respiratory failure, renal failure, and stroke (\(r = 0.26, 0.21, 0.20\), respectively, \(p < 0.001\) for all), and almost no correlation for deep sternal wound infections (\(r = 0.03, p = 0.479\)), with authors concluded that the use of measures of morbidity or mortality alone might be inadequate to assess care quality as a result.

The above-described work of Silber and others has major implications for surgical practice and surgical outcomes. Whereas the risk of postoperative morbidity is due to patient-related factors, rather than care- or centre-dependent factors related to actual processes of care, the reverse would appear to hold true for mortality. Thus, high-risk patients who undergo an operation can be expected to survive the procedure, and remain more likely to develop complications; however, it is failure to recover these patients from complications which results in variable mortality rates between hospitals. These studies’ findings suggest that the largest proportion of variability in care in current practice (and surgical outcomes, by extension) may be ascribed specifically to variable quality in the domain of postoperative care, measured by rates of failure to rescue.

As a result, there has been a surge of interest in the use of failure to rescue as an outcome by which to measure hospital care quality in the past decade.\(^6^5,^{66}\) Ghaferi and colleagues’ widely cited 2009 study, in particular, has suggested that whilst risk-adjusted morbidity rates between hospitals may be similar, analysis of FTR rates may reveal significant variations in care quality and outcome.\(^6^5\) In this study, the authors analysed data of 84,730 patients, who had undergone high risk vascular or general surgery from 2005-2007, from the United States’ National Surgical Quality Improvement Program (NSQIP) database. Rates of mortality and major morbidity (as defined by the authors, including infectious complications, renal failure,
pneumonia, bleeding, and others) were assessed. Performance of hospitals was compared by dividing them into quintiles of mortality.

After adjusting for patient factors, no differences were found in the risk of developing either any recorded complication, or major complications, across hospital quintiles. However, there was an almost twofold increase in mortality for patients following the development of a complication across quintiles, from 3.5% to 6.9%. The rate of death following major complications demonstrated a similar trend, with rates of failure (i.e. death) of 12.5% vs. 21.4%. Though this study was limited, again, by its retrospective, database-derived nature, specifically in this case as the NSQIP database records only a pre-defined set of complications rather than all complications in general, its results suggested underlying failings in post-operative patient care to be responsible for variability in outcomes. A further follow-up study conducted by the same group, but using different data (this time comparing outcomes for 269,911 patients from the US Medicare database) found similar results for patients undergoing major surgery (pancreatectomy, oesophagectomy, abdominal aortic aneurysm repair, coronary artery bypass grafting, aortic valve replacement, or mitral valve replacement). Whereas complication rates were similar across all hospital quintiles (ranging from 32.7% to 36.4%), FTR rates were significantly different (16.7% in the worst quintile for mortality vs. 6.8% in the best quintile).\textsuperscript{66} Though neither of these studies are able to provide insight into exactly which factors, or domains of care, are involved, Ghaferi and colleagues have suggested that factors affecting clinical care such as communication and effective patient management were probably among the key determinants of FTR, and that future research should concentrate on elucidating both organisational or structural factors and care processes which improved care.\textsuperscript{65}

FTR has been since been suggested as a more sensitive care quality indicator than other measures currently in use, including morbidity and mortality rates, or operative volume.\textsuperscript{67}

As the assessment of care quality seeks to define ever more detailed metrics, and with the
variability in FTR rates seemingly attributable to differences in ward care, using FTR rates as a measure of quality should allow a more accurate measurement of post-operative clinical care quality. These may be considered alongside established metrics such as morbidity and mortality, or indeed other newer metrics such as patient satisfaction and patient-related outcomes. Irrespective of this, however, the variations in FTR make clear that further investigation of factors which make clinical care in one hospital superior to that in another is required, if true inroads into the standardisation and improvement of surgical patient care are to be achieved.

The remainder of this chapter seeks to explore current concepts and evidence in the assessment of ward-based care and resulting surgical outcomes. The research of this thesis is placed in context, with the underlying factors relating to variable patient outcomes explored in a systematic review, as are the potential implications of failures in care, i.e. preventable postoperative complications, on long-term patient survival.
1.3 Enhancing surgical performance outcomes through process-driven care

1.3.1 Introduction

Though successful surgery is undoubtedly an important element of surgical care, the rarity of intraoperative mortality in modern practice, combined with the high rates of morbidity associated with major procedures\textsuperscript{67} mean that the prevention of post-operative complications, and treatment of them should they occur, is at least equally critical in the determination of patient outcomes.

Postoperative complications are commonly defined as “any deviation from the normal postoperative course,” not including sequelae inherent to the procedure itself,\textsuperscript{68} and include adverse events which may or may not be the result of medical error, as discussed earlier in this chapter. Complications may result in anything from no discernable difference to patient outcome, managed with standard medical treatment and full recovery, to permanent disability or even death, and are generally categorised in terms of severity accordingly. The methods used in the Harvard Medical Practice study, which graded complications on a 1-6 scale according to the severity of their sequelae (ranging from full recovery to death),\textsuperscript{3} have been continued, if refined, by the classification systems most widely in use currently. Though the more recently published Comprehensive Complication Index\textsuperscript{69} has been recommended to be used in its place, the most widely used Clavien-Dindo Classification defines 5 grades of complications according to the level of intervention required, wherein grades I and II require pharmacological treatment (antipyretics or antibiotics), whereas grade III entails operative intervention, grade IV organ failure requiring
intensive care, and grade V resulting in death.\textsuperscript{68} The Accordion Classification,\textsuperscript{70} recommends similar classification, but adds an additional level by differentiating between operative interventions requiring general anaesthetic and those which do not, i.e. endoscopic intervention (subdivided as grades IIIb and IIIa in the Clavien-Dindo Score).

The known variability of surgical outcomes serves to highlight the differences in care quality that exist between different health care providers.\textsuperscript{71} The associated differences in FTR rates across different institutions suggest, as previously described, that a large proportion of this variance may be attributed to a failure to identify, ameliorate or prevent post-operative complications.\textsuperscript{61}

1.3.1.1 Failure to rescue – variants and adaptations

Though FTR has since been increasingly adopted as a key outcome measure for care quality,\textsuperscript{67} data on underlying factors which influence failure rates remains sparse. Since its original description by Silber in 1992 to describe failures in postoperative care, several variant definitions of “failure to rescue” have been defined. Needleman and colleagues adapted the measurement of FTR to include “nurse sensitive complications,” (FTR-N) with outcomes measured for both surgical and non-surgical patients: pneumonia, shock, gastrointestinal bleeding, cardiac arrest, sepsis, and deep venous thrombosis.\textsuperscript{72} This cohort study of over 6 million patient episodes, derived from discharge data and nurse staffing data from 11 US states’ databases, examined nurse staffing-related metrics such as hours of care and proportion of all care hours provided by a registered nurse, controlling for hospital-related factors. Higher nurse staffing levels, and a higher proportion of daily patient care provided by registered nurses, was associated with decreased rates of these complications, and decreased FTR rates resulting from them.
Needleman’s definition has been since further revised by AHRQ to better suit a use of the term from a patient safety perspective, with the addition of renal failure to the previous set of complications (FTR-A), representing the common complications deemed most likely to be preventable. AHRQ has further revised Needleman’s definition to better suit a use of the term from a patient safety perspective, with the addition of renal failure to the previous set of complications (FTR-A), representing the common complications deemed most likely to be preventable. Almoudaris and colleagues have also proposed a further definition, “failure to rescue – surgical” (FTR-S), to measure patient deaths following unplanned reoperation. The authors reasoned that administrative databases, as seen in almost all other FTR-related studies, often fail to include in their coding an adequate proportion of actually occurring complications, or do not appropriately differentiate between new and pre-existing morbidities, making the use of such data for FTR analysis potentially inappropriate. Rather, they assessed patients undergoing reoperation, a condition much more reliably coded, as representing the group of patients with serious complications (i.e. by definition Clavien-Dindo IIIb or greater), analysing differences in FTR rates for 144 542 patients from the English HES database undergoing colorectal surgery. Here, too, significant differences were seen between hospital mortality quintiles, ranging from 11.1 to 16.8% (p = 0.002).

Despite this slowly growing nomenclature, however, Silber has argued that these more restrictive FTR measures, by definition, limit the inclusion of patient deaths such that up to 40% of deaths following postoperative complications are excluded. In a 2007 study, Silber and colleagues compared FTR-N, FTR-A, and FTR rates across 403,679 surgical patients derived from 1999-2000 US MEDPAR data. FTR-N and FTR-A were found to omit 49% and 42% of all recorded deaths, respectively. Furthermore, authors conducted a reliability analysis, randomly dividing patient groups for each hospital into two groups and comparing calculated FTR rates for each version of the metric – FTR was found to be significantly more reliable than FTR-A or FTR-N (rho = 0.32, 0.18, 0.18, respectively). Silber thus concluded that though FTR-N, FTR-A, and FTR-S are each useful in the analysis of their specific subgroup of interest, they are less useful in the measurement of surgical care quality overall, with the
original definition of FTR perhaps the best standard by which to judge postoperative care outcomes.

1.3.1.2 Structure, process, and outcome

The improvement of care quality, it must be highlighted, is not limited to the measurement of outcomes alone. In keeping with the paradigm originally described by the late Avedis Donabedian nearly 50 years ago, widely considered the pioneer of outcomes research, quality may be classified through an assessment of three interlinked domains: structure, process and outcome.\textsuperscript{76,77} Considering each of these three domains has its advantages and disadvantages, and impacts upon different levels of a health system (table 1.1).

_Structural measures_ reflect system factors at any level of care, such as staffing levels, hospital size or patient volume. However, the organisational factors which they represent are often difficult to act upon as they must be actioned on a broader institutional, regional or national scale to implement change.\textsuperscript{78} The previously discussed inverse relationship between hospital operative volume and mortality, for example,\textsuperscript{45,79} has contributed to the ongoing drive towards the centralisation of complex procedures in most Western countries to ensure that sufficient operative volume is available to maintain expertise. Whilst this may contribute to improvements in surgical outcomes, it has required significant health systems restructuring on a national scale, with restriction of specialist operations to high-volume tertiary centres.\textsuperscript{80,81}
### Table 1.1 Structure, process and outcome variables affect, and measure, the quality of care in health systems

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Advantage</th>
<th>Disadvantage</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>System and organisational factors</td>
<td>Easily measured</td>
<td>Requires broader scale policy change to implement</td>
<td>Hospital operative volume, Nurse staffing levels</td>
</tr>
<tr>
<td>Process</td>
<td>Actions carried out in giving or receiving care</td>
<td>Locally actionable, Can be implemented as interventional trial</td>
<td>Implementation affected by local practice culture</td>
<td>Thromboprophylaxis measures, Early post-operative feeding</td>
</tr>
<tr>
<td>Outcome</td>
<td>The effect of care on the health status of patients and populations</td>
<td>Easily measured, Broadly reported</td>
<td>Multiple confounding factors can affect validity</td>
<td>Morbidity, Mortality, Patient reported outcomes</td>
</tr>
</tbody>
</table>
*Process variables* describe actions carried out in giving or receiving care. In contrast to structural factors, these are simpler to modify on an individual basis, as they describe actions directly carried out by individuals or teams in contact with patients. As a result, process variables are also more amenable to direct observation, and as such are also well suited for randomised trials. Examples include the administration of beta blockers following myocardial infarction, or the advantages of early enteral feeding in general surgical patients.

*Outcomes* describe the effects of care on the health and well-being of patients and populations. Thanks to the widespread establishment of patient databases, outcomes are easily measured and readily calculated from available spreadsheets of figures, though as previously described, the use of such database-derived data is not without its pitfalls, being dependent, for example, upon the correctness and completeness of recorded data.

The demand for increased transparency and public accountability in healthcare, combined with the accessibility of data, has meant that outcomes such as morbidity, mortality, or indeed failure to rescue, represent the measures of quality most familiar to the modern physician (or layperson). However, in focusing on the defining of outcomes and ways to measure healthcare, one must not overlook the fact that Donabedian’s paradigm has always stressed the linkage of outcomes to the other two domains of process and outcome.

Measuring quality through outcome measures is but the first step in quality improvement; the next step is to develop interventions which may improve them. A health system cannot directly manipulate outcomes, but can modify health structures which may affect outcomes in turn. As individual clinicians, it is often impossible to change the larger structures within which we work, but we can attempt to modify the care processes which take place in the course of giving care to patients. It is within this domain that programmes of quality improvement must be contextualised.
Process-based interventions have in the past focused upon individual patient or diagnosis-specific problems, such as the above named examples. Particularly within the past decade, however, there has been a shift towards a more systems-based approach to care improvement, as described earlier in this chapter with the introduction of Reason’s “Swiss Cheese” model. One example of this has been the dissemination of bundled care processes to regulate entire treatment pathways rather than individual processes alone.

Such process-driven care aims to ensure high quality care through a focus on the implementation and measurement of care at a process level. A leading example of this has been the 2006 study described by Pronovost and colleagues, who reported the results of the implementation of five evidence-based procedures to reduce central venous catheter-associated sepsis.84 This was a pre-/post-intervention cohort study of 103 US intensive care units. The primary recorded outcome was the rate of catheter-related infections. Over an 18-month period (March 2004 – September 2005), enrolled centres introduced a care-bundle programme of hand washing, use of sterile barriers, cleaning of insertion sites with chlorhexidine, avoidance of femoral sites, and removal of unnecessary catheters. They did not specifically record compliance rates with the process measures, but gave as an example that prior to the study, only 19% of hospitals included chlorhexidine in their central venous catheter kits – compared to 84% stocking it in either kits or in the intensive care unit afterward. Following a 3-month implementation period, outcome data was collected and compared to pre-intervention data in a comprehensive model controlling for hospital factors, and geographic and data-related clustering. Authors found that over the course of the study, median infection rates were reduced from 2.7 per 1000 catheter-days to 0.84 A 3-year follow up study found sustained, significant improvement in outcomes.85 These findings have since led to international dissemination of the catheter care bundle, including to the UK, where the matching of the Michigan group’s results was specifically set out as one of the targets of Lord Darzi’s 2008 NHS Next Stage Review, High Quality Care for All.86 This has resulted in the
“Matching Michigan” initiative reporting significant reductions in infection rates in UK intensive care units, from 3.7 to 1.4 infections per 1000 catheter-days.87

1.3.1.2 Enhanced recovery protocols

Building upon evidence-based care bundles, enhanced recovery protocols (ERP), care pathways determined by collated best evidence peri-operative care processes, have led to a recent revolution in peri- and post-operative care. Colorectal surgery, in particular, has widely adopted the ERP principle,88 with evidence-based interventions such as avoiding peritoneal drains (which impair mobility but do not reduce the rate or severity of complications)89,90 or encouraging early post-operative oral feeding (which reduces risk of infection and length of stay).91 First described by Kehlet and colleagues in 1995, describing a case series of 9 patients undergoing laparoscopic colonic resection for cancer, the initial aim of this bundled care protocol was to reduce the intra- and postoperative stress response and its role in postoperative organ dysfunction, such as gut dysfunction in the form of postoperative ileus.92 A more extensive series of 60 patients was published in 2000.93 Through the use of epidural analgesia, laparoscopic approach, and early nutrition and mobilization, in both studies the authors were able to achieve a median length of stay of only 2 days, for a procedure which would ordinarily have been expected to incur a stay of at least 5-10 days.

The exact components of colorectal ERPs have been since further refined,94 with evidence from subsequent randomised trials strongly suggesting that by combining up to such evidence-based processes and thus setting daily targets for routine care within a care pathway, the aim of ERPs are to shorten the length of inpatient stays, and reduce morbidity and mortality.95 Increasingly, ERPs are no longer confined to colorectal surgery alone, with similar protocols now being introduced in other specialties as well.96
The introduction of process-driven care pathways has driven the standardisation of the practice of evidence-based medicine. Despite this, the growing body of evidence for failure to rescue makes clear that an unacceptable degree of variability in the quality of post-operative care persists.\(^{65,67,97}\) Though variability in FTR rates are now widely accepted, there remains a lack of understanding as to its underlying causes. In order for these to be addressed, it is necessary to consider the effects of structural and process factors on post-operative care quality and outcomes.

*The aim of this systematic review was to assess the currently available evidence for structure and process factors affecting post-operative surgical care and FTR rates.*

### 1.3.2 Methods

#### 1.3.2.1 Search strategy

A comprehensive search strategy was used to identify relevant articles assessing factors affecting the quality of postoperative care. The search therefore focused on underlying determinants of failure to rescue, as a measure of postoperative care quality following complications, and postoperative care protocols, as care processes dedicated specifically to standardisation of care in the perioperative phase.

The following search terms were used: failure to rescue, care protocol, care pathway, enhanced recovery, fast track, multimodal perioperative; these terms were combined using Boolean “and” operand with the term “surgery” or MeSH headings “general surgery” or “operative surgical procedures.” PubMed and EMBASE were searched for publications from inception to June 1, 2013.
Returned results were searched for titles and abstracts for candidate articles according to pre-defined selection criteria (detailed below). These were retrieved for full-text review, with publications meeting the search criteria included in final analysis. Relevant full-text article bibliographies were also searched by hand to ensure all relevant articles were included.

The literature search and data extraction were both performed by two independent researchers. Any disagreements were discussed within the research group and resolved by consensus.

1.3.2.2 Selection criteria

Studies included were those which assessed either structural variables or bundled postoperative care processes, protocols or pathways, and their effects on postoperative surgical outcomes. Studies that addressed only pre- or intra-operative variables (without addressing postoperative care) were excluded. Only articles published in English (or with English translation available) were included in final data synthesis.

Owing to the large number of publications on ERPs, inclusion of publications on the effectiveness of ERP programs was limited to randomised controlled trials of acceptable quality. Based upon previously published definition and methodology, ERP was defined as a peri-operative care pathway which included modification of no less than seven specific care processes, compared to conventional care, as part of the protocol. For other structural or process factors, all study types were included, subject to quality analysis.

1.3.2.3 Quality analysis
Quality of randomised trials was evaluated using a previously validated score developed by Jadad et al. The Jadad score assigns or deducts points across seven domains of assessment, including appropriate blinding, randomisation, and reporting, for a final score of 0 – 5 (table 1.2). A score of 3 or more was considered acceptable quality and suitable for inclusion.
<table>
<thead>
<tr>
<th>Item</th>
<th>Points allocated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Was the study described as randomised?</td>
<td>1</td>
</tr>
<tr>
<td>Was the method used to generate the sequence of randomisation described, and was it appropriate (e.g. computer-generated)?</td>
<td>1</td>
</tr>
<tr>
<td>Was the study described as double-blinded?</td>
<td>1</td>
</tr>
<tr>
<td>Was the method of blinding described and was it appropriate?</td>
<td>1</td>
</tr>
<tr>
<td>Was there a description of dropouts from the study?</td>
<td>1</td>
</tr>
<tr>
<td>Deduct 1 point if the method used for randomisation was described but was inappropriate (e.g. patients allocated alternately, or according to date of birth)</td>
<td>-1</td>
</tr>
<tr>
<td>Deduct 1 point of the study was described as double-blinded but the method of blinding was inappropriate</td>
<td>-1</td>
</tr>
<tr>
<td><strong>Total score</strong></td>
<td><strong>/5</strong></td>
</tr>
</tbody>
</table>
Cohort trials were assessed using the Newcastle-Ottawa scale for cohort studies (fig 1.2). This scale assigns points on the basis of valid group selection, comparability, and outcome assessment for a score of 0 – 9. Moderate and high quality studies, with a score of 7 or higher, were included.
**Fig 1.2** Newcastle-Ottawa quality assessment scale for cohort studies

**Note:** A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

### Selection

1) **Representativeness of the exposed cohort**
   - a) truly representative of the average ______________ (describe) in the community ✷
   - b) somewhat representative of the average ______________ in the community ✷
   - c) selected group of users eg nurses, volunteers
   - d) no description of the derivation of the cohort

2) **Selection of the non exposed cohort**
   - a) drawn from the same community as the exposed cohort ✷
   - b) drawn from a different source
   - c) no description of the derivation of the non exposed cohort

3) **Ascertainment of exposure**
   - a) secure record (eg surgical records) ✷
   - b) structured interview ✷
   - c) written self report
   - d) no description

4) **Demonstration that outcome of interest was not present at start of study**
   - a) yes ✷
   - b) no

### Comparability

1) **Comparability of cohorts on the basis of the design or analysis**
   - a) study controls for ______________ (select the most important factor) ✷
   - b) study controls for any additional factor ✷ (This criteria could be modified to indicate specific control for a second important factor.)

### Outcome

1) **Assessment of outcome**
   - a) independent blind assessment ✷
   - b) record linkage ✷
   - c) self report
   - d) no description

2) **Was follow-up long enough for outcomes to occur**
   - a) yes (select an adequate follow up period for outcome of interest) ✷
   - b) no

3) **Adequacy of follow up of cohorts**
   - a) complete follow up - all subjects accounted for ✷
   - b) subjects lost to follow up unlikely to introduce bias - small number lost - > ____ % (select an adequate %) follow up, or description provided of those lost ✷
   - c) follow up rate < ____% (select an adequate %) and no description of those lost
   - d) no statement
1.3.2.4 Statistical analysis

Data was extracted for analysis and performed in consultation with a statistician in Stata 12 for Windows (StataCorp, College Station, TX, USA). Multiple regression analysis was performed to identify the component processes with the greatest effect on outcome for care pathways with common components. Where sufficient comparable data was available, pooled meta-analysis was performed for other structural and process factors identified.

1.3.3 Results

The initial database search identified 1945 non-duplicate articles. After applying selection criteria, 50 full-text articles were retrieved for further analysis. Nine studies were excluded on the basis of study quality. These were mostly cohort studies in which compared groups had significantly different baseline demographics (i.e. inappropriate comparison of different populations), or minor significant differences between groups that were not appropriately controlled for. Two further studies were excluded as they used a different definition of failure to rescue, reporting only deaths following re-operation (“failure to rescue – surgical,” FTR-S) but not taking into account complications in general. Finally, two studies were excluded which used a subset of previously published data to report on identical outcomes measures. For studies that investigated structural or process factors based on a subgroup analysis of a previously published dataset, only results from the larger dataset analysis were extracted for analysis. This resulted in the inclusion of 37 relevant articles for final analysis (see figure 1.3).
Figure 1.3. Flow diagram of literature search results.

FTR-S: failure to rescue – surgical, AAA: abdominal aortic aneurysm
Studies were divided into two categories – those which prospectively assessed ERPs and their effect on outcomes, and those which retrospectively analysed failure to rescue rates to identify related structures or processes.

Care process interventions in the form of ERPs were identified in twenty-three studies, applying to procedures in urology and general, bariatric, hepatobiliary, and vascular surgery. Sixteen studies concerning structure and process factors in FTR were identified and analysed.

1.3.3.1 Enhanced recovery protocols

Colorectal ERP studies included a total of 2,570 patients. Demographics of the included studies, as well as the details of the protocol used in each study are shown in Table 1.3. All studies scored 3 (moderate quality) on the Jadad scale, with appropriately randomisation and reporting, but lacking blinding.

There was a high degree of variation in the components of various ERP programmes. Most commonly included in ERPs were early mobilisation from the day of surgery (100%), feeding on the morning of surgery (91%), avoidance of post-operative nasogastric tubes and early post-operative feeding (both 87%). Least common components were the use of goal-directed post-operative intravenous fluid therapy, routine chest physiotherapy, and daily telephone follow-up post-discharge (all 4%).

Outcomes analysis (table 1.4) showed a significantly reduced length of stay across 21 of 23 studies for ERP treatment versus conventional care. Ren et al\textsuperscript{102} reported a non-significant reduction in length of stay (5.7±1.6 days for ERP vs. 6.6±2.4 days for conventional care) in their randomised trial of 597 patients undergoing colectomy. Similarly, van Bree et al\textsuperscript{103} reported a mean stay of 5.9 days with ERP vs. 6 days with conventional care in their series of 35 patients following laparoscopic colectomy.
There was no difference in readmission rates, with all studies which reported total length of stay, including readmissions, found a significantly reduced length of stay for ERP patients. Six of 23 studies reported a significant reduction in morbidity with ERP, though no significant differences were found by the remaining 17 studies. None reported any difference in mortality.

### 1.3.3.2 Failure to rescue

A total of 16 large cohort studies were identified, comprising some 20 million patient episodes and over 40,000 nursing staff in primarily American centres (see table 1.5). Categorising the associated factors assessed by each study according to Donabedian’s definition, these studies examined 18 structural and 2 process factors. Structural factors related to nursing staff, hospital size, resources, and accreditation, as well as the percentage of higher trained medical staff (table 1.6).

Most thoroughly assessed was the association between nurse staffing levels and FTR rates. Seven of nine studies reported significant increases in FTR rates with an increase in nursing workload, as measured by a reduction in the number of nurses or nursing hours available per patient. Aiken and colleagues also demonstrated the reduction in FTR rates shown with better working environments and higher levels of nurse education. Nursing education was reported in a further study by Aiken and colleagues to be separate from experience alone, with no association between nursing experience and FTR rates found. The proportion of care provided by registered nurses (RNs) or agency staff was shown to have a weak but significant relationship in a study by Blegen and colleagues (regression coefficient 0.008, p < 0.001), but no relationship in two other studies.

Several studies examined differences in hospital types and resources and differences in FTR rates. Whereas teaching hospitals, centres meeting benchmark resource, clinical and
research levels to be awarded National Cancer Institute (NCI) accreditation,\textsuperscript{107} and recognised centres of nursing excellence (Magnet hospitals),\textsuperscript{108} demonstrated strong reductions in FTR rates, safety net status did not confer any beneficial effects.\textsuperscript{109} The greater technological resources of centres with transplant, cardiac, burns or renal units were reflected in lower FTR rates in a study by Ghaferi et al,\textsuperscript{97} though when measured by the Saiden Index, a measure of institutional technological resource availability, in a further study by Blegen et al, this difference was not seen.\textsuperscript{109}

Conflicting reports were found for hospital size and operative volume. Whereas studies by Silber et al\textsuperscript{110} and Ghaferi et al\textsuperscript{97} showed remarkable decreases in FTR (odds ratios 0.87 and 0.65, respectively) for high capacity hospitals (greater than 200 beds), Friese et al\textsuperscript{107} could not show the same. For operative volume, in a second study by Ghaferi et al,\textsuperscript{67} an almost threefold increase in FTR rate was seen in low volume hospitals. However, this association was not seen in two other studies, which found no significant relationship between FTR and volume.\textsuperscript{107, 111} Higher hospital occupancy\textsuperscript{97} and ICU capacity\textsuperscript{111} were associated with decreases in FTR in single studies.

Process factors identified were whether surgery was performed by a subspecialist surgeon,\textsuperscript{112} and the direction of anaesthesia by a clinician anaesthetist, rather than an anaesthetic practitioner,\textsuperscript{110} both of which resulted in decreased FTR rates.
<table>
<thead>
<tr>
<th>Study</th>
<th>n (ERP)</th>
<th>n (CC)</th>
<th>Procedure</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>i</th>
<th>j</th>
<th>k</th>
<th>l</th>
<th>m</th>
<th>n</th>
<th>o</th>
<th>p</th>
<th>q</th>
<th>r</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jones, 2013 (UK)</td>
<td>46</td>
<td>45</td>
<td>Partial hepatectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lemanu, 2013 (New Zealand)</td>
<td>40</td>
<td>38</td>
<td>Laparoscopic sleeve gastrectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ni, 2013 (China)</td>
<td>80</td>
<td>80</td>
<td>Partial hepatectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim, 2012 (Korea)</td>
<td>22</td>
<td>22</td>
<td>Laparoscopic distal gastrectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ren, 2012 (China)</td>
<td>299</td>
<td>298</td>
<td>All colonic resections</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang, 2012 (China)</td>
<td>49</td>
<td>50</td>
<td>Laparoscopic colectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang, 2012 (China)</td>
<td>40</td>
<td>38</td>
<td>Laparoscopic colectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yang, 2012 (China)</td>
<td>32</td>
<td>30</td>
<td>Open colectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garcia-Botello, 2011 (Spain)</td>
<td>61</td>
<td>58</td>
<td>All colonic resections</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Magheli, 2011 (Germany)</td>
<td>25</td>
<td>24</td>
<td>Laparoscopic prostatectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Bree, 2011 (Belgium)</td>
<td>Lap</td>
<td>18</td>
<td>Laparoscopic colectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Open</td>
<td>18</td>
<td>Open colectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vlug, 2011 (Netherlands)</td>
<td>Lap</td>
<td>100</td>
<td>Laparoscopic colectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Open</td>
<td>93</td>
<td>Open colectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang D, 2010 (China)</td>
<td>45</td>
<td>47</td>
<td>Subtotal or total gastrectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ionescu, 2009 (Romania)</td>
<td>48</td>
<td>48</td>
<td>Open colectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muehling, 2009 (Germany)</td>
<td>49</td>
<td>50</td>
<td>Open AAA repair</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muller, 2009 (Switzerland)</td>
<td>76</td>
<td>75</td>
<td>Open colectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serclova, 2009 (Czech Rep.)</td>
<td>51</td>
<td>52</td>
<td>Open bowel resection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Khoo, 2007 (UK)</td>
<td>35</td>
<td>35</td>
<td>Open colectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gralla, 2007 (Germany)</td>
<td>25</td>
<td>25</td>
<td>Laparoscopic prostatectomy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------</td>
<td>---------</td>
<td>----------------</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>----</td>
<td>---</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Gatt, 2005 (UK)</td>
<td>2005</td>
<td></td>
<td>Open colectomy</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderson, 2003 (UK)</td>
<td>2003</td>
<td></td>
<td>Hemicolecotomy</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td>☑️</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ERP: enhanced recovery protocol; CC: conventional care; AAA: abdominal aortic aneurysm; lap: laparoscopic

- a: pre-op counseling
- b: pre-op nutrition
- c: no bowel preparation
- d: post-op ileus prevention
- e: epidural analgesia
- f: no premedication
- g: transverse incision in right hemicolecotomy
- h: avoidance of intra-abdominal drains
- i: no post-op nasogastric tube
- j: early urinary catheter removal
- k: limited intravenous fluids
- l: early oral intake
- m: structured mobilisation program
- n: standardised anaesthetic
- o: regular post-op antiemetic
- p: standardised analgesia
- q: goal-directed fluid therapy
- r: routine chest physiotherapy
- s: daily telephone follow-up in community
<table>
<thead>
<tr>
<th>Study</th>
<th>Initial length of stay</th>
<th>Readmission rate</th>
<th>Total length of stay</th>
<th>Complication rate</th>
<th>Mortality rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ERP CC</td>
<td>ERP CC</td>
<td>ERP CC</td>
<td>ERP CC</td>
<td>ERP CC</td>
</tr>
<tr>
<td>Jones, 2013</td>
<td>113</td>
<td>4.2% 0%</td>
<td>4[3-5] 7[6-8]</td>
<td>17% 31.3%</td>
<td>2% 2%</td>
</tr>
<tr>
<td>Lemanu, 2013</td>
<td>114</td>
<td>2 [no IQR] ***</td>
<td>20% 21.1%</td>
<td>25% 21.1%</td>
<td>0% 0</td>
</tr>
<tr>
<td>Ni, 2013</td>
<td>115</td>
<td>6.9±2.8 8.0±3.7  *</td>
<td>30% 46.3%*</td>
<td>0% 0</td>
<td></td>
</tr>
<tr>
<td>Kim, 2012</td>
<td>116</td>
<td>5.36±1.46 7.95±1.98 ***</td>
<td>4.5% 0%</td>
<td>13.6% 18.2%</td>
<td></td>
</tr>
<tr>
<td>Ren, 2012</td>
<td>102</td>
<td>5.7±1.6 6.6±2.4</td>
<td>9.7% 9.4%</td>
<td>0% 0</td>
<td></td>
</tr>
<tr>
<td>Wang G, 2012</td>
<td>117</td>
<td>4(2-12) 5(3-48) **</td>
<td>12.2% 20%</td>
<td>2% 0</td>
<td></td>
</tr>
<tr>
<td>Wang Q, 2012</td>
<td>118</td>
<td>5.5[5-6] 7[6-8] ***</td>
<td>5% 21.1%*</td>
<td>0% 0</td>
<td></td>
</tr>
<tr>
<td>Yang, 2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(China)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Garcia-Botello, 2011</td>
<td>120</td>
<td>4.15±2.18 9.23±6.97 ***</td>
<td>3% 5%</td>
<td>0% 0</td>
<td></td>
</tr>
<tr>
<td>Magheli, 2011</td>
<td>121</td>
<td>3.6±1.2 6.7±0.9 ***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Bree, 2011</td>
<td>103</td>
<td>Lap 3.9 n/a 5.6% 11.8%</td>
<td>11.2% 29.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Open 5.9 6 5.6% 11.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vlug, 2011</td>
<td>112</td>
<td>Lap 5[4-8] 6[4-8.5] *</td>
<td>6% 6.4%</td>
<td>54% 55%</td>
<td>2% 2%</td>
</tr>
<tr>
<td></td>
<td>Open 6[4.5-10] 7[6-10.5] 8.0% 7.1% 7[5-11] 7[6-13] *</td>
<td>76.0% 73% 4% 2%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang D, 2010</td>
<td>122</td>
<td>6[6-7] 8[7-8] **</td>
<td>2% 2%</td>
<td>20% 14.9%</td>
<td>0% 0</td>
</tr>
<tr>
<td>Ionescu, 2009</td>
<td>124</td>
<td>6.43±3.41 9.16±2.7 ***</td>
<td>0 0</td>
<td>6.43±3.41 9.16±2.7 ***</td>
<td>12.2% 22.4%</td>
</tr>
<tr>
<td>Muehling, 2009</td>
<td>125</td>
<td>10[6-49] 11[8-45] *</td>
<td>36% 16%* a</td>
<td>0% 0</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Median (Range)</td>
<td>Median [Inter-quartile Range]</td>
<td>Percentages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>----------------</td>
<td>------------------------------</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muller, 2009</td>
<td>5(2-30)</td>
<td>9(6-30) ***</td>
<td>2.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>17%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>37% **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serclova, 2009</td>
<td>7(5-11)</td>
<td>9(7-22) ***</td>
<td>21.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>48.1% **</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Khoo, 2007</td>
<td>5(3-37)</td>
<td>7(4-63) ***</td>
<td>9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5(3-37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7(4-63) ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gralla, 2007</td>
<td>3.60±1.22</td>
<td>6.72±0.94 ***</td>
<td>4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>24%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>56% *</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gatt, 2005</td>
<td>5[4-9]</td>
<td>7.5[6-10] *</td>
<td>5.30%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>47%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>75%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderson, 2003</td>
<td>3(2-7)</td>
<td>7(4-10) **</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3(2-7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7(4-10) **</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>29%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>45%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>9%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Results reported as median [range], median [inter-quartile range], or mean ± SD.

ERP: enhanced recovery protocol; lap: laparoscopic; ^medical morbidity
<p>| Study                  | Study population                                                                 | Specialty                                           | Quality | a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t |
| Aiken, 2013 (US)      | 1,295,068 patients across 665 centres, convenience sample                       | General, orthopedic, and vascular surgery           | 8       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| McHugh, 2013 (US)     | 641,187 patients across 564 hospitals, convenience sample                      | General, orthopedic, and vascular surgery           | 7       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Aiken, 2011 (US)      | Survey data from 39,038 nurses across 665 hospitals, AHA data                  | General, orthopedic, and vascular surgery           | 7       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Blegen, 2011 (US)     | &gt;1.1m patient episodes, 54 hospitals                                            | All except psychiatry and obstetrics               | 8       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Ghaferi, 2011 (US)    | 37,865 patients, Medicare database 2005-2007                                    | Gastrectic, pancreatic, esophageal resection        | 7       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Mell, 2011 (US)       | 2,616 patients                                                                  | Elective abdominal aortic aneurysm repair           | 9       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Friese, 2010 (US)     | 4,618 patient episodes, 164 hospitals (subset data)                            | Any cancer surgery                                  | 7       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Ghaferi, 2010 (US)    | 8,862 patient episodes, 672 hospitals, AHRQ data                                | Pancreatectomy                                      | 8       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Silber, 2009 (US)     | 4,658,954 patient episodes, 3,270 hospitals                                     | General, orthopaedic, vascular surgery              | 7       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Sochalski, 2008 (US)  | 457,786 patients, 343 hospitals                                                  | Cardiology (n=348,720), surgery (n=109,066)        | 8       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| Rafferty, 2007 (UK)   | 118,752 patients, 3,984 nurses, 30                                              | General, orthopaedic                                | 7       |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |</p>
<table>
<thead>
<tr>
<th>Study</th>
<th>Hospitals and Vascular Surgery</th>
<th>Procedure</th>
<th>Number of Patients/Episodes</th>
<th>Year</th>
<th>Country</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith, 2007 (US)</td>
<td>1,864 patients, 214 hospitals</td>
<td>Elective gastrectomy for primary malignancy</td>
<td>111</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aiken, 2003 (US)</td>
<td>10,184 nurses, 232,342 patient episodes in 168 hospitals</td>
<td>General, orthopedic, and vascular surgery</td>
<td>105</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aiken, 2002 (US)</td>
<td>10,184 nurses, 232,342 patient episodes in 168 hospitals</td>
<td>General, orthopedic, and vascular surgery</td>
<td>134</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Needleman, 2002 (US)</td>
<td>5,075,969 medical, 1,104,659 surgical patients, 799 hospitals</td>
<td></td>
<td>72</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Silber, 2000 (US)</td>
<td>217,440 patient episodes, 245 hospitals</td>
<td>Surgery</td>
<td>110</td>
<td>7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Structural factors** - a: nurse staffing levels, b: nurse education levels, c: nursing work environment, d: proportion of care provided by RN, e: “safety net” hospital status, f: teaching intensity, g: hospital size, h: NCI centre accreditation, i: hospital volume, j: high technology, k: hospital occupancy, l: board certification of anaesthetists, m: board certification of surgeons, n: surgical house staff, o: ICU capacity, p: nurse experience, q: magnet hospital status, r: % of agency nurse staff

**Process factors** - s: surgery performed by specialist surgeon, t: anaesthetic directed by anaesthetist
### Table 1.6  Study outcomes, structure and process factors affecting failure to rescue

<table>
<thead>
<tr>
<th>Factor</th>
<th>Study</th>
<th>Modification</th>
<th>Effect on FTR rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nurse staffing levels</td>
<td>134</td>
<td>Increase patient: nurse ratio by 1</td>
<td>OR (95% CI): 1.07 (1.02, 1.11)**</td>
</tr>
<tr>
<td></td>
<td>109</td>
<td>Increase number of nursing hours / patient</td>
<td>OR (95% CI): -0.023; p &lt; 0.1</td>
</tr>
<tr>
<td></td>
<td>97</td>
<td>Increase patient: nurse ratio by 1</td>
<td>OR (95% CI): 1.06 (1.01, 1.11)****</td>
</tr>
<tr>
<td></td>
<td>72</td>
<td>Increasing nursing hours: patient ratio from 25th to 75th percentile</td>
<td>% change (95% CI): -5.9% (1.5, 10.2)**</td>
</tr>
<tr>
<td></td>
<td>133</td>
<td>1st (6.9 – 8.3) vs 4th (12.4 – 14.3) quartiles for patient: nurse ratio</td>
<td>OR (95% CI): 1.29 ***</td>
</tr>
<tr>
<td></td>
<td>110</td>
<td>25% increase in nurse: bed ratio</td>
<td>OR (95% CI): 0.95 (0.93, 0.98)***</td>
</tr>
<tr>
<td></td>
<td>111</td>
<td>High RN/patient ratio (&gt; dataset median)</td>
<td>OR (95% CI): 0.68 (0.40, 1.17), p 0.17</td>
</tr>
<tr>
<td></td>
<td>132</td>
<td>Increase in 1 hour / day of care per patient</td>
<td>% change (95% CI): ** -0.47% (-0.59%, -0.35%)</td>
</tr>
<tr>
<td></td>
<td>104</td>
<td>Increase patient: nurse ratio by 1</td>
<td>OR (95% CI): 1.04 (1.02, 1.06) *</td>
</tr>
<tr>
<td>Nurse work environment</td>
<td>104</td>
<td>Increase in PES-NWI by 1 standard deviation</td>
<td>OR (95% CI): 0.93 (0.90, 0.95) **</td>
</tr>
<tr>
<td>Nurse education</td>
<td>104</td>
<td>10% increase in BSN certified staff</td>
<td>OR (95% CI): 0.96 (0.94, 0.98)***</td>
</tr>
<tr>
<td>Nurse experience</td>
<td>105</td>
<td>1 year increase in mean experience</td>
<td>OR (95% CI): 1.01 (0.98, 1.03)</td>
</tr>
<tr>
<td>Type of nursing staff</td>
<td>109</td>
<td>For increasing proportion of nursing care provided RN</td>
<td>Regression coefficient: 0.008 ***</td>
</tr>
<tr>
<td></td>
<td>72</td>
<td>25th vs 75th percentiles for proportion RN nursing hours / total</td>
<td>% change (95% CI): 3.9 (-1.1, 8.8), p 0.12</td>
</tr>
<tr>
<td></td>
<td>106</td>
<td>Agency nurses &gt;15% vs &lt;5%</td>
<td>OR (95% CI): 1.02 (0.98, 1.06)</td>
</tr>
<tr>
<td>Hospital accreditation</td>
<td>Safety net status</td>
<td>Regression coefficient:</td>
<td>0.094</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------------------</td>
<td>--------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>NCI centre accreditation</td>
<td>No % FTR rate (SD)</td>
<td>Yes</td>
<td>4.86 (3.1)</td>
</tr>
<tr>
<td>Magnet hospital status</td>
<td>OR (95% CI)</td>
<td>0.81 (0.72, 0.91) ****</td>
<td></td>
</tr>
<tr>
<td>Teaching intensity</td>
<td>No trainees</td>
<td>% FTR rate (SD)</td>
<td>4.65 (3.4)</td>
</tr>
<tr>
<td>&gt; 4 beds / resident</td>
<td>5.14 (2.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 4 beds / resident</td>
<td>4.37 (1.8); p = 0.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Teaching hospital status</td>
<td>OR (95% CI):</td>
<td>0.66 (0.53, 0.82)</td>
<td></td>
</tr>
<tr>
<td>Member, Council of Teaching Hospitals</td>
<td>OR (95% CI):</td>
<td>0.84 (0.79, 0.89)</td>
<td></td>
</tr>
<tr>
<td>&gt; 0.6 residents / bed vs non-teaching</td>
<td>OR (95% CI):</td>
<td>0.88 (0.85, 0.92)</td>
<td></td>
</tr>
<tr>
<td>Residency program</td>
<td>Presence of surgical house staff</td>
<td>OR (95% CI):</td>
<td>0.99 (0.93, 1.06)</td>
</tr>
<tr>
<td>Hospital size</td>
<td>&lt; 100 beds</td>
<td>% FTR rate (SD)</td>
<td>4.46 (4.6)</td>
</tr>
<tr>
<td>100 – 250 beds</td>
<td>5.21 (3.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;250 beds</td>
<td>4.2 (1.9); p = 0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bedsize &gt; 200</td>
<td>OR (95% CI):</td>
<td>0.65 (0.48, 0.87) ****</td>
<td></td>
</tr>
<tr>
<td>≥ 200 vs &lt; 200 beds</td>
<td>OR (95% CI):</td>
<td>0.87 (0.81, 0.94) ***</td>
<td></td>
</tr>
<tr>
<td>Hospital volume</td>
<td>Lowest quartile</td>
<td>% FTR rate (SD)</td>
<td>5.01 (4.5)</td>
</tr>
<tr>
<td>Highest quartile</td>
<td>4.43 (1.8); p = 0.50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High volume hospital (&gt;15 gastrectomies / year)</td>
<td>OR (95% CI):</td>
<td>0.22 (0.04, 1.13), p 0.07</td>
<td></td>
</tr>
<tr>
<td>High vs low volume hospitals</td>
<td>OR (95% CI):</td>
<td>2.89 (2.40, 3.48) ****</td>
<td></td>
</tr>
<tr>
<td>High technology</td>
<td>Transplant, cardiac, burns, renal centre</td>
<td>OR (95% CI):</td>
<td>0.65 (0.52, 0.91) ****</td>
</tr>
<tr>
<td>As measured by Saiden Index</td>
<td>No significant relationship</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Occupancy level</td>
<td>Daily census &gt;50% capacity</td>
<td>OR (95% CI):</td>
<td>0.56 (0.32, 0.98) ****</td>
</tr>
<tr>
<td>ICU capacity</td>
<td>High ICU capacity (&gt; dataset median)</td>
<td>OR (95% CI):</td>
<td>0.53 (0.29, 0.97) *</td>
</tr>
<tr>
<td></td>
<td>FTR (%)</td>
<td>OR (95% CI)</td>
<td>p-value</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>---------</td>
<td>----------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Board certification (anaesthetist)</td>
<td></td>
<td>25% increase of certified staff</td>
<td>0.97 (0.95, 0.99) **</td>
</tr>
<tr>
<td>Board certification (surgeon)</td>
<td></td>
<td>Certified surgeon performs operation</td>
<td>0.80 (0.68, 0.94) **</td>
</tr>
<tr>
<td></td>
<td></td>
<td>25% increase of certified staff</td>
<td>0.98 (0.95, 1.02)</td>
</tr>
<tr>
<td>Treatment by specialist</td>
<td></td>
<td>Vascular surgeon</td>
<td>16.2%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-vascular surgeon</td>
<td>32.3% *</td>
</tr>
<tr>
<td>Treatment by anaesthetist</td>
<td></td>
<td>Effect of non-anaesthetist direction of anaesthetic care</td>
<td>1.10 (1.01, 1.18) *</td>
</tr>
</tbody>
</table>

FTR: failure to rescue, OR: odds ratio, CI: confidence interval, SD: standard deviation; PES-NWI: practice environment scale of the nursing work index; BSN: bachelor of science in nursing; ICU: intensive care unit

* p < 0.05; ** p < 0.01; *** p < 0.001; **** significant with no p value reported
1.3.3.3 Statistical analysis

Regression analysis was performed for the effects of individual ERP processes on length of stay and morbidity rates, but failed to identify a significant relationship for any individual factor.

Availability of data to allow pooled meta-analysis was possible for three structural factors. All results reported as odds ratio (95% CI). Increasing patient:nurse ratios (i.e. reduction in nurse staffing levels) uniformly increased FTR rates, with an overall effect size of 1.08 (1.05, 1.11). Both teaching hospital status, 0.83 (0.75, 0.90) and larger hospitals with greater than 200 beds, 0.78 (0.57, 0.99), resulted in decreased rates of FTR, though both demonstrated significant data heterogeneity as measured by I-square test (I-square 78.2%, p = 0.010, and I-square 77.3%, p = 0.036, respectively) (figure 1.4).
Figure 1.4 Meta-analysis for impact of increasing of patient:nurse ratio by 1, teaching hospital status, hospital size > 200 bed capacity. Odds ratio < 1 favours decrease in FTR rate (decreased mortality), > 1 favours increase in FTR (increased mortality).

Dot: odds ratio, whiskers: 95% confidence interval, diamond: pooled odds ratio with 95% confidence interval.
1.3.4 Discussion

This review has captured the impact of structural factors and care processes on the quality of post-operative care. The identification of these factors associated with differences in outcomes such as failure to rescue rates is an important step in the improvement of ward-based surgical care. By understanding significant structures and processes, and the means by which each can be implemented, interventions to improve care quality may be considered.

Considering structural factors, the strongest evidence was identified for nurse staffing rates, with a change in nurse staffing ratios equivalent to an additional patient per nurse (i.e. increased nursing workload) associated with a small, but significant increase in FTR (OR 1.08 (1.05, 1.11)). In addition, there were significant trends for lower FTR rates in teaching hospitals and larger centres (i.e. increased bed capacity).

Conflicting findings were noted for effects of high technology capabilities of institutions; though it is plausible that this may be explained by the variety of surgical procedures that were compared. Blegen et al analysed a cohort of “all-comers,” with the inclusion of all specialties (both surgical and non-surgical) other than psychiatry, finding no significant relationship between hospital resources and FTR rates for this “routine” patient group. The “Saidin index” used in their study represents a weighted sum of technologies and services available to an institution, with weighting determined by the percentage of hospitals within a dataset which lack the same resources. The study reported by Ghaferi et al, conversely, included only patients who had undergone pancreatectomy – a major, and complex, procedure – and found a significant effect of technological resources for FTR rates (OR 0.65 (0.52, 0.91)). Patients undergoing major surgery are most likely to develop complications, and therefore benefit from access to additional high technology resources. Findings of a recent study by Funk et al would appear to support this hypothesis, wherein
mortality rates following oesophagectomy, in hospitals with three or more of high nursing ratio, complex medical, bariatric, or lung transplantation services or positron emission tomography (PET) demonstrated an adjusted mortality rate less than half of those without, 5.0% vs 12.5%.

There was strong evidence for the positive effect of the implementation of ERP process bundles. Despite ranging across seven different sub-specialties, and an even greater number of different surgical procedures, almost all reviewed studies achieved significant reductions in length of stay with the implementation of ERP. Considering the two exceptions, the lack of significant findings may be explained by statistical error. Ren and colleagues’ sample consisted of patients undergoing colorectal procedures of varying complexity, ranging from the relatively routine right hemicolecotomy, to the more major abdominoperineal resection. These rather different procedures (though not individually reported) will have resulted in a broad range of hospital stay durations, affecting statistical analysis and increasingly the likelihood of underpowering or type 2 error. Van Bree and colleagues reported length of stay for only part of their sample, with a relatively small sample size (as few as 17 patients in the laparoscopic colectomy / conventional care group), and should therefore be discounted. In addition to reduced hospitalisations, many studies also reported a decrease in post-operative morbidity, demonstrating the potential advantages of standardised, process-led care – despite unstandardised patient groups and care protocols.

From this review, it is clear that structures and processes both have a role to play in determining outcomes. As discussed in the introduction, however, the ease of implementation, and clinical impact, of each may vary greatly. To effect change by changing structural factors, e.g. through the hiring of additional nursing staff or acquisition of additional resources, can require major reorganisation. This is often difficult, costly, and may be unfeasible altogether, for example, in smaller hospitals lacking prerequisite catchment populations or clinical expertise. ERP, by comparison, exemplifies the potential
impact a process-driven approach may have on care without incurring a significant additional resource burden.

Recent research by Birkmeyer et al\textsuperscript{136} has shown that the cost of potentially avoidable complications may be more than $5000 per patient, a strong financial incentive for centres to improve practice, especially if faced with bundled lump sum payments for services. The findings of this review demonstrate the potential of a process-driven approach to care and its effect on outcomes, in agreement with recent meta-analyses of colorectal ERP,\textsuperscript{95,137} which have found ERP to significantly reduce lengths of stay with no difference in readmission and mortality rates.

Despite this, it is worth noting that the implementation of new care processes, whilst not necessitating change on an organisational scale, can be no less challenging to implement.\textsuperscript{138} Studies have shown that in centres which implement ERP, adherence to self-imposed protocols is highly variable, with published audits of practice showing 60 – 70\% compliance rates for ERP processes.\textsuperscript{139-141} With these figures being published by larger academic centres only, and the high likelihood of publication bias, actual average compliance rates are likely to be lower still.

This failure to fully assess the implementation of protocols is too often overlooked. In one cohort study of colorectal ERP reporting compliance rates,\textsuperscript{142} an average implementation of only 7.4 of 13 ERP components was reported, the worst being 13\% compliance on the intake of 2 carbohydrate drinks pre-operatively. Furthermore, analysis of papers in this review reveals a lack of consensus on ERP processes, with the components initially described by Kehlet’s group\textsuperscript{88} implemented to variable degrees. Only a single component process was targeted for inclusion in all trials – structured early mobilisation.

Given the past success of protocols in the operating room, it is not surprising that further OR protocols, both generic and specific, have already been trialed. Borrowing from the multiple
checklists of the aviation industry, individual “crisis pathways” have been evaluated in the operating theatre for specific intra-operative complications, an approach which could be expanded into the post-operative setting to address specific postoperative complications. By addressing key areas of morbidity, and potential mortality, directly through such protocols, a greater effect on these metrics might also be expected compared to ERP, whose effects on morbidity and mortality remain unclear.

Other attempts at monitoring and modifying generic processes surrounding routine patient post-operative care have shown some promise – basic ward round checklists and daily “goal setting” for patient management have improved communication and documentation on the ward. The need to expand these attempts further and to improve and standardise care is clear.

This systematic review has limitations which must be considered. Compliance rates in the implementation of ERP were infrequently reported and highly variable, so that the degree of implementation of all purported ERP sub-processes in all studies cannot be certain. In searching for failure-related structure and process factors, our search relied on the keyword “failure to rescue”. Though FTR is a widely accepted term, it is possible that relevant studies may have been missed which used different terminology (e.g. mortality rates adjusted for patient and hospital factors), though hand-searching of references attempted to account in part for this. FTR-based studies were exclusively cohort studies drawn from large multi-centre databases, which are vulnerable to the known limitations of risk-adjustment, quality and type of data entry. Finally, the definition of FTR used to define outcomes and patient groups in the studies analysed was variable, with several studies utilising the AHRQ definition in favour of Silber’s original definition.
1.3.5 Conclusion

This systematic review highlights the positive impact of process-driven care to improve ward based surgical care, and the logistic and financial advantages in their implementation compared to structural factors. Current post-operative ward based care is inconsistent, as illustrated by FTR rates which demonstrate variability in recovery rates from post-operative complications. Whilst there is strong evidence for the impact of nurse-to-patient ratios on outcomes, the effect of other structural factors on failure rates is less clear. Process-driven care protocols can be implemented on a local basis to improve and standardise care with positive impact on outcomes, but to date have been implemented in a limited fashion, most notably in the form of ERP.

Given the evidence of this review, and the comparable implementability of process versus structural interventions themselves, it seems logical, and necessary, to consider a process-based approach to improving postoperative care. Considering ERPs as a successful example for implementing evidence-based peri-operative care, however, it is important to recognise also that they represent a model of care for the “best-case” patient, with a reduced length of stay and without suffering major morbidity in the postoperative phase. ERPs fail to address patients who do develop complications, thus necessitating “off protocol” care – wherein the risk of suboptimal and unstandardised care may place them at risk of failure to rescue.

This raises two questions. First, how to address process-based factors which may improve the detection and management of postoperative complications (i.e. reduction or prevention of FTR), as well as reducing their incidence. The exploration of this question forms the bulk of the remainder of this thesis.

Second, however, it also raises the question of the longer-term impact of postoperative complications themselves. Beyond the certainly valid, but short-term, arguments of patient
discomfort\textsuperscript{147} and economic burden,\textsuperscript{136} when the ultimate outcome of interest remains patient survival, just how important is the absolute rate of complications compared to the rate of failure to rescue from them? Is it potentially sufficient to focus on the management of complications, such as through the development of protocols, to ensure full recovery? A literature review was therefore performed to investigate the role of morbidity prevention on long-term outcomes beyond in-hospital failure or rescue.
1.4 Postoperative in-hospital morbidity impairs long-term patient survival: systematic review and meta-analysis

1.4.1 Introduction

The previous review (chapter 1.3) discussed the potential improvement in management of complications, i.e. reduction in FTR rates, which may be achieved with the implementation of appropriate evidence-based processes and structural measures such as the now widespread use of enhanced recovery protocols in colorectal surgery, and centralized treatment in high-volume centers. However, whilst these changes have contributed to significant reductions in mortality and failure to rescue rates, the absolute incidence of postoperative morbidity in major surgery remains high.

Major surgery is routinely associated with high rates of postoperative morbidity of up to 50%. These may be further categorized as surgical/technical or non-surgical/non-technical complications. Surgical complications relate to the procedure itself, and may be attributed to intraoperative factors, such as failure of a poorly fashioned anastomosis. In contrast, non-surgical complications such as postoperative pneumonia or deep vein thrombosis relate to generic processes in the postoperative phase of care and can occur in any patient.

It has become increasingly accepted that the majority of generic complications, though common, are preventable through the implementation of simple care processes, and by ensuring the provision of quality care. The previously described Keystone-Michigan project for the prevention of central venous catheter-associated sepsis is one such example. Other process-improvement projects have continued to be developed in a similar vein, such as the
US Surgical Care Improvement Project (SCIP). This national quality partnership project mandates the implementation of 20 generic process measures, such as the administration of preoperative antibiotics, adequate blood glucose level control, and prevention of hypothermia, and has resulted in a reduction of postoperative infectious complications as a result.\textsuperscript{152}

The reduction of preventable complications has numerous important implications for both patients and hospitals. The immediate consequence of postoperative morbidity is to prolong patients’ hospital stay, and to exert a negative effect on patient wellbeing and quality of life, from which patients may take months to recover.\textsuperscript{153, 154} A recent systematic review by Bouras et al reported that adverse events resulted in an up to 14\% decrease in subjective quality of life, as measured by validated questionnaire instruments, even at follow-up of 15 to 64 months postoperatively.\textsuperscript{147} For health systems, complications result in an increased cost of care and use of resources.\textsuperscript{155, 156, 157} Eappen and colleagues assessed financial and administrative data for 12 US hospitals, considering all surgical discharges over the course of 2010.\textsuperscript{157} In their analysis of 34,256 patient episodes, they found that the incidence of one or more surgical complications resulted in an up to USD$ 39,017 (95\% CI $20,069, $50,394) greater cost to the hospital, depending on the insurance/payer status of the patient.

Traditionally, focus has been placed upon these short-term effects of postoperative morbidity. However, it has been suggested that complications have a detrimental effect on long-term survival as well, even following appropriate treatment and full recovery. Khuri and colleagues’ reported the results of an analysis of 105,000 surgical patients, derived from US NSQIP-linked data.\textsuperscript{158} With a mean follow-up of 8 years, long-term survivals following surgery with, or without, post-operative complications were compared using a stepwise logistic regression model, well adjusted for multiple patient factors. Overall, for all procedures, authors found that 5-year survival was reduced by almost a third for patients
with post-operative morbidity compared to those without (58% vs. 40%). This trend persisted throughout subanalyses for different complication types such as wound infection or pneumonia, as well as individual procedures. Whilst the mechanism for this is unclear, if such a relationship were to be definitively demonstrated, the need to take further measures to not only treat, but prevent postoperative morbidity, would be compelling. However, the current state of evidence remains unclear, as other authors have failed to find an association between complications and long-term survival.\textsuperscript{149, 159}

\textit{A systematic review was undertaken, with the aim to review best published evidence to assess the impact of post-operative morbidity on long-term survival after surgery.}

\section*{1.4.2 Methods}

\subsection*{1.4.2.1 Search strategy}

A systematic review was conducted in accordance with PRISMA\textsuperscript{160} guidelines. Literature search and data extraction were both conducted by two independent researchers. Any discrepancies were agreed upon by consensus. MEDLINE and Web of Science databases were searched from inception up to July 1, 2013. The search used the following terms: “surgery” and “long-term survival” and “postoperative” and (“complications” or “morbidity”), and included the MeSH terms “morbidity” and “postoperative period.” Following de-duplication of search results, an initial review of titles and abstracts was conducted to identify articles of potential interest, which were then retrieved for full-text
analysis and independent data extraction. In addition, reference lists of retrieved articles were hand-searched for additional relevant references.

### 1.4.2.2 Selection criteria

All surgical disciplines and procedures were considered for inclusion in this review. Studies comparing long-term survival in adult patient cohorts with and without post-operative morbidity as a primary outcome were considered for inclusion in final analysis. Studies assessing only the effect of procedure-specific complications, or of “surgical” complications secondary to technical error (as defined by the authors in each case), such as anastomotic leak following colonic resection, or bronchopleural fistula following thoracic surgery, were excluded. Studies including only paediatric populations were also excluded.

### 1.4.2.3 Assessment of methodological quality

Studies considered for inclusion in final analysis were required to meet a minimum level of objectively assessed methodological quality. Study quality was assessed using the Newcastle-Ottawa Scale (NOS) for cohort studies. This scale, discussed in greater detail previously (Chapter 1.3.2.3), assesses cohort sample selection, data collection and analysis, and assigns a final score of 0-9. Based on previous studies’ methodologies, we defined a NOS score of 7 or greater as acceptable.

### 1.4.2.4 Statistical analysis

Data analysis was performed in Stata 12 (StatCorp, College Station, TX). Meta-analysis of overall survival and disease-free survival (DFS) was performed using a random effects model.
Z-test was used to assess the overall effect. Data heterogeneity was assessed using the I-square test. In all cases, a p-value of less than 0.05 considered statistically significant.

To assess the potential impact of poor quality studies, additional sensitivity analysis was performed with previously excluded studies with NOS ≤ 6.

The risk of publication bias was assessed with funnel plots and Egger’s test.164

1.4.3 Results

1.4.3.1 Search results and study selection

The pre-defined search strategy returned 5,395 non-duplicate references (figure 1.5). Following screening of titles and abstracts, 25 papers were retrieved in full-text form and assessed for eligibility. All 25 articles reported the effects of post-operative complications on long-term survival, however eight of these were of poor methodological quality as judged by an NOS score of 6 or less and were therefore excluded. Poor NOS scores were due to inadequate or unreported follow-up, lack of patient demographic information, and lack of appropriate controls for confounding factors (i.e. use of unadjusted univariate regression models). A hand search of reference lists of retrieved studies resulted in inclusion of one additional study missed in the original search.
Figure 1.5 Flow chart of literature search results.

Initial search (n = 5925) → De-duplication (n = 530)

Records screened (n = 5395) → Exclusion of irrelevant results (n = 5370)

Full-text assessed for eligibility (n = 25) → Exclusion of poor quality studies (n = 8)

Studies included in final data synthesis (n = 18) → Additional results from hand-searching of references (n=1)
1.4.3.2 Study demographics

A total of 18 cohort studies were included in final data synthesis, with reported results for 134,785 patients with an overall complication rate of 22.6% (see table 1.7). The largest cohort study, by Khuri et al., analysed results of 105,951 patients following a range of procedures for both malignant and benign disease: abdominal aortic aneurysm repair, lower limb arterial bypass, carotid endarterectomy, colectomy, cholecystectomy, pneumonectomy, or total hip replacement. All remaining studies were studies of single procedures for primary malignancies or resection of colorectal liver metastases.

Whereas most studies considered complications in general, four studies reported the effects of infectious or septic complications separately. For two of these studies, these were defined as systemic sepsis based upon clinical findings, such as direct evidence of infection (i.e. pus) or pyrexia, combined with microbiological or radiological evidence of infection. The other two studies did not specifically report their criteria for defining infectious complications. Three of the four studies analyzing infectious complications assessed operative procedures involving intra-abdominal anastomoses and included anastomotic leaks in their analysis. However these made up a small proportion of total complications, 11%, 22%, and 26%, respectively.

Reported median postoperative follow-up of the included 18 studies ranged from 29 – 68.5 months, with an overall median follow-up of 43 months. Study quality as measured by NOS score was good, with a mean score of 8.2 (range 7-9). All studies included in final analysis utilized fully adjusted regression models, controlling for factors such as age, comorbidities (local data or database-derived Charlson score), procedure, and tumour type, where appropriate.
Table 1.7 Demographic details of included studies comparing patient cohorts without or with post-operative morbidity.

<table>
<thead>
<tr>
<th>Study</th>
<th>Data source</th>
<th>Procedure</th>
<th>Diagnosis</th>
<th>n</th>
<th>Type of comp.</th>
<th>Follow-up (months)</th>
<th>NOS score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andalib, 2013, Canada</td>
<td>Provincial database</td>
<td>Any lung surgery</td>
<td>Lung cancer</td>
<td>4,033</td>
<td>All</td>
<td>60</td>
<td>7</td>
</tr>
<tr>
<td>Mrak, 2013, Austria</td>
<td>Retrospective single-centre</td>
<td>Rectal resection</td>
<td>Rectal cancer</td>
<td>811</td>
<td>All</td>
<td>61.2 ± 3.7</td>
<td>7</td>
</tr>
<tr>
<td>Hii, 2013, Australia</td>
<td>Retrospective single-centre</td>
<td>Esophag-ectomy</td>
<td>SCC or adeno-carcinoma</td>
<td>311</td>
<td>All</td>
<td>51 (2-151)</td>
<td>9</td>
</tr>
<tr>
<td>Mavros, 2013, USA</td>
<td>Retrospective single-centre</td>
<td>Partial hepatectomy</td>
<td>CLM</td>
<td>251</td>
<td>All</td>
<td>33.6 [16-63]</td>
<td>9</td>
</tr>
<tr>
<td>Tokunaga, 2013, Japan</td>
<td>Retrospective single-centre</td>
<td>Gastrectomy</td>
<td>Gastric adeno-carcinoma</td>
<td>765</td>
<td>Infectious</td>
<td>63</td>
<td>9</td>
</tr>
<tr>
<td>Tan, 2012, USA</td>
<td>SEER-Medicare database</td>
<td>Nephrectomy</td>
<td>Renal cell carcinoma</td>
<td>12,618</td>
<td>All</td>
<td>32 (1-132)</td>
<td>8</td>
</tr>
<tr>
<td>Rueth, 2011, USA</td>
<td>SEER-Medicare database</td>
<td>Any lung surgery</td>
<td>Stage I NSCLC</td>
<td>3996</td>
<td>All</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Chauhan, 2010, USA</td>
<td>Retrospective single-centre</td>
<td>Partial hepatectomy</td>
<td>Cholangio-carcinoma</td>
<td>51</td>
<td>All</td>
<td>28 (19-106)</td>
<td>8</td>
</tr>
<tr>
<td>Schepers, 2010, UK</td>
<td>Retrospective single-centre</td>
<td>Partial hepatectomy</td>
<td>CLM</td>
<td>121</td>
<td>All</td>
<td>38</td>
<td>8</td>
</tr>
<tr>
<td>Farid, 2010, UK</td>
<td>Retrospective single-centre</td>
<td>Partial hepatectomy</td>
<td>CLM</td>
<td>705</td>
<td>All</td>
<td>36</td>
<td>9</td>
</tr>
<tr>
<td>Tanaka, 2010, Japan</td>
<td>Retrospective single-centre</td>
<td>Partial hepatectomy</td>
<td>CLM</td>
<td>312</td>
<td>All</td>
<td>39</td>
<td>7</td>
</tr>
<tr>
<td>Tsujimoto, 2008, Japan</td>
<td>Retrospective single-centre</td>
<td>Gastrectomy</td>
<td>Gastric adeno-carcinoma</td>
<td>1,332</td>
<td>Infectious</td>
<td>68.5 (3-211)</td>
<td>9</td>
</tr>
<tr>
<td>Ito, 2008, USA</td>
<td>Retrospective single-centre</td>
<td>Partial hepatectomy</td>
<td>CLM</td>
<td>1,067</td>
<td>All</td>
<td>41</td>
<td>9</td>
</tr>
<tr>
<td>Country</td>
<td>Study Year</td>
<td>Study Design</td>
<td>Procedure</td>
<td>Procedure Details</td>
<td>Procedure Type</td>
<td>Complication Type</td>
<td>Complications Recorded</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
<td>--------------</td>
<td>-------------------------</td>
<td>--------------------------------------------</td>
<td>----------------</td>
<td>-------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Australia</td>
<td>2007</td>
<td>Retrospective</td>
<td>Partial hepatectomy</td>
<td>CLM</td>
<td>197</td>
<td>All</td>
<td>54</td>
</tr>
<tr>
<td>China</td>
<td>2007</td>
<td>Retrospective</td>
<td>Colorectal resection</td>
<td>Colorectal adenocarcinoma</td>
<td>1,657</td>
<td>All</td>
<td>45.3</td>
</tr>
<tr>
<td>USA</td>
<td>2005</td>
<td>NSQIP database</td>
<td>Multiple(^b)</td>
<td>Various</td>
<td>105,951</td>
<td>All</td>
<td>96</td>
</tr>
<tr>
<td>Italy</td>
<td>2004</td>
<td>Retrospective</td>
<td>Left colonic or rectal resection</td>
<td>Colorectal adenocarcinoma</td>
<td>192</td>
<td>Infectious ≥ 60</td>
<td>9</td>
</tr>
<tr>
<td>France</td>
<td>2003</td>
<td>Retrospective</td>
<td>Partial hepatectomy</td>
<td>CLM</td>
<td>311</td>
<td>All</td>
<td>29</td>
</tr>
</tbody>
</table>

Results reported as median (range), median [IQR], or mean ± SD.

Type of comp.: type of complications recorded, NOS: Newcastle-Ottawa Scale, CLM: colorectal liver metastases; SCC: squamous cell carcinoma, NSCLC: non-small cell lung cancer.

\(^a\) Separate results for patients undergoing partial hepatectomy or isolated liver perfusion reported

\(^b\) 6 procedures included: AAA, lower limb bypass, carotid endarterectomy, colectomy, cholecystectomy, lobectomy/pneumonectomy, total hip replacement.
1.4.3.3 Impact of morbidity on long-term survival

Post-operative complication rates ranged between 10.6%\textsuperscript{179} – 58.2%\textsuperscript{166} (see table 1.8). All but one study reported 5-year overall survival rates (Ito and colleague reported DFS only).\textsuperscript{173} Amongst these 17 studies, all but one reported significant reductions in 5-year survival following complications, with the exception of Mrak and colleagues’ report of 811 rectal resection patients, in which no significant difference in survival rates was found between groups.\textsuperscript{181}

Results sufficient for meta-analysis were reported by ten studies, representing 20,755 patients. Hazard ratios for survival (HR (95% CI)) ranged between 0.86 (0.58, 1.28)\textsuperscript{181} and 3.58 (1.46, 8.27)\textsuperscript{169} following adjusted multivariate regression modeling. Meta-analysis demonstrated an overall hazard ratio with significantly reduced overall survival following any complication, HR (95% CI) 1.29 (1.21, 1.34), p < 0.001, though there was significant data heterogeneity, I-square 52.9%, p 0.024 (figure 1.6).

Analysis of four studies (2,994 patients) reporting overall survival following infectious complications demonstrated similar results, with reduced overall survival, HR 1.92 (1.50, 2.35), p < 0.001, with no significant heterogeneity (figure 1.6).
Table 1.8 Outcomes of included studies comparing patient cohorts without or with post-operative morbidity.

<table>
<thead>
<tr>
<th>Study</th>
<th>Comp. rate</th>
<th>Severe comp. rate</th>
<th>5-year DFS</th>
<th>p-value</th>
<th>5-year overall survival</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Without</td>
<td>With</td>
<td>Without</td>
<td>With</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andalib</td>
<td>58.2%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>73.8%</td>
<td>0.001</td>
</tr>
<tr>
<td>Mrak</td>
<td>34.5%</td>
<td>-</td>
<td>65.7%</td>
<td>66.8%</td>
<td>n/r</td>
<td>72.4%</td>
</tr>
<tr>
<td>Hij</td>
<td>64.6%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>74mos&lt;sup&gt;a&lt;/sup&gt;</td>
<td>36mos&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Mavros</td>
<td>21.9%</td>
<td>5.6%</td>
<td>22.7%</td>
<td>6.0%</td>
<td>n/r</td>
<td>46.0%</td>
</tr>
<tr>
<td>Tokunaga</td>
<td>-</td>
<td>7.2%</td>
<td>84.5%</td>
<td>64.90%</td>
<td>&lt; 0.001</td>
<td>86.8%</td>
</tr>
<tr>
<td>Tan</td>
<td>37%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>69.5%</td>
<td>59.9%</td>
</tr>
<tr>
<td>Rueth</td>
<td>54.2%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>67.4%</td>
<td>55.6%</td>
</tr>
<tr>
<td>Chauhan</td>
<td>69%</td>
<td>47.0%</td>
<td>40mos&lt;sup&gt;a&lt;/sup&gt;</td>
<td>14mos&lt;sup&gt;a&lt;/sup&gt;</td>
<td>n/r</td>
<td>36mos&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Schepers</td>
<td>29.8%</td>
<td>8.0%</td>
<td>-</td>
<td>-</td>
<td>52.0%</td>
<td>41.0%</td>
</tr>
<tr>
<td>Farid</td>
<td>28%</td>
<td>17%</td>
<td>26%</td>
<td>13%</td>
<td>0.001</td>
<td>28%</td>
</tr>
<tr>
<td>Tanaka</td>
<td>26%</td>
<td>9.8%</td>
<td>31.3%</td>
<td>20.2%</td>
<td>0.03</td>
<td>55.4%</td>
</tr>
<tr>
<td>Tsujimoto</td>
<td>10.6%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>78.1%</td>
<td>57.9%</td>
</tr>
<tr>
<td>Ito</td>
<td>42%</td>
<td>25.0%</td>
<td>33.0%</td>
<td>25.0%</td>
<td>0.005</td>
<td>-</td>
</tr>
<tr>
<td>Schiesser</td>
<td>30%</td>
<td>13.7%</td>
<td>-</td>
<td>-</td>
<td>42.0%</td>
<td>25%</td>
</tr>
<tr>
<td>Law</td>
<td>27.3%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>70.0%</td>
<td>57.0%</td>
</tr>
<tr>
<td>Khuri</td>
<td>18.3%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>57.6%</td>
<td>39.5%</td>
</tr>
<tr>
<td>Nespoli</td>
<td>23%</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>57.0%</td>
<td>18.0%</td>
</tr>
<tr>
<td>Laurent</td>
<td>30%</td>
<td>28.0%</td>
<td>12.0%</td>
<td>0.001</td>
<td>42.0%</td>
<td>21.0%</td>
</tr>
</tbody>
</table>

Comp. rate: complication rate, severe comp. rate: rate of complications with Clavien-Dindo grade ≥ III, DFS: disease-free survival, n/r: not reported, sig: reported as statistically significant but precise p-value not given

<sup>a</sup>median, <sup>b</sup>Separate results for patients undergoing partial hepatectomy or isolated liver perfusion reported
Figure 1.6 Meta-analysis of overall survival after all post-operative complications (n=20,755, p < 0.001) and infectious complications (n=2,994, p < 0.001). Solid dots indicate hazard ratios, whiskers represent 95% CI. Diamond represents overall effect.
1.4.3.4 Impact of morbidity on disease-free survival

Of eight studies reporting 5-year DFS rates, five reported significantly lower rates for the cohort with complications. Two of the remaining studies found lower median survival (14 vs. 40 months)\(^{169}\) or 5-year DFS rates (6% vs. 22.7%),\(^ {165}\) but did not give results of statistical analyses. Mrak and colleagues reported a slightly higher (66.8% vs. 65.7%) DFS in their patient cohort with complications, however this was not significant (fig 1.7).

Meta-analysis of six studies (4,294 patients) indicated significantly reduced DFS after any complication (figure 1.6), HR 1.26 (1.10, 1.42), \(p < 0.001\), as well after infectious complications, HR 1.55 (1.12, 1.99), \(p < 0.001\) (2 studies, 1,470 patients). Heterogeneity as measured by I-square test was non-significant in both cases.
Figure 1.7 Meta-analysis of disease-free survival after all post-operative complications (n=4,294, p < 0.001) and infectious complications (n=1,470, p < 0.001). Solid dots indicate hazard ratios, whiskers represent 95% CI. Diamond represents overall effect.
1.4.3.5 Impact of severity of morbidity on survival

The effects of complication severity were investigated by three studies, with conflicting results. Results published by Ito et al\textsuperscript{173} and by Mrak et al\textsuperscript{181} found no difference in survival when stratifying complications by grade as measured by Clavien-Dindo classification.\textsuperscript{68} In their relatively small series of 197 patients undergoing hepatic resection for metastatic colorectal disease, however, Schiesser et al reported a significant difference in survival when patients were grouped as severe (Clavien-Dindo grade III/IV) or none/mild (grade 0/I/II), median survival 2.1 vs. 4.1 years, \( p = 0.012 \), respectively, DFS HR 1.8 (1.2, 3.8), \( p = 0.005 \).\textsuperscript{172}

1.4.3.6 Sensitivity analysis

The impact of excluded poor-quality studies was assessed through a further sensitivity analysis. Of eight excluded studies, four reported hazard ratios suitable for analysis.\textsuperscript{182-184} Inclusion of these in meta-analysis had no effect on findings for overall survival, with similar resulting hazard ratios for all complications (1.29 (1.22, 1.36), \( p < 0.001 \)), or infectious complications (1.52 (1.42, 1.62), \( p < 0.001 \)).

1.4.3.7 Assessment of publication bias

Assessment of risk of bias via funnel plots suggested possible publication bias with positive skew of poorer-quality studies (i.e. overrepresentation of poor quality studies demonstrating reduced long-term survival after complications) (figure 1.8), Egger’s bias coefficient 1.69, \( p = 0.001 \).
Figure 1.8 Funnel plot assessment of potential bias for overall survival after any postoperative complications
1.4.4 Discussion

This review presents a summary and meta-analysis of current evidence on the effects of postoperative complications on long-term survival. These results show a significant association between the incidence of complications following major surgery and earlier all-cause mortality, independent of potentially confounding variables such as premorbid conditions, age, or underlying diagnosis.

Previous studies\textsuperscript{185,186} have reported the detrimental effects of specific surgical complications, i.e. anastomotic leak, on tumour recurrence in colorectal cancer. Although the precise mechanism underlying the association remains unclear, a recent meta-analysis reported a near 3-fold increase in risk of local recurrence, as well as significant increases in distal recurrence and decreased long-term survival after anastomotic leak (OR 2.9).\textsuperscript{187}

The findings of this thesis chapter go further still, suggesting that the incidence of any complication reduces overall patient survival. Thus, the importance of post-operative care is highlighted as being not only key to the management of complications and short-term outcomes, but also the avoidance of preventable morbidity in the determination of long-term survival. Furthermore, though there was insufficient evidence to include this in a pooled meta-analysis, it is possible that this effect is not limited to cancer surgery alone, as exhibited by the large cohort analysis of patients with predominantly benign disease by Khuri and colleagues.\textsuperscript{158}

In their analysis, Khuri et al also reported the most common complications across the six different procedures assessed by their study. With the exception of abdominal aneurysm repair, in which “failure to wean” – based upon their database classification, though this is arguably not a complication in itself – was listed as the primary complication (followed by pneumonia), the most frequent complications were infectious in nature and included pneumonia, urinary tract infection, and wound infection.\textsuperscript{158} The meta-analysis presented in
this thesis reveals that the relationship between complications and reduced survival is stronger still when considering only infectious complications, with a higher hazard ratio, compared to outcomes after all complications: HR 1.92 (1.50, 2.35) for infectious complications (with no significant data heterogeneity) vs. 1.28 (1.22, 1.34) for all complications (heterogenous data, I-square 52.9%, p 0.024). Both hazard ratios were highly statistically significant with p-values of less than 0.001, and significantly different from one another with no overlap of 95% confidence intervals. Despite this statistically significant association, the mechanism by which complications, septic or otherwise, affect survival remains uncertain. Postoperative morbidity often leads to prolonged inpatient stay, which may delay adjuvant chemo- or radiotherapy. However, this small delay, in a relatively small proportion of patients, cannot account for the large reduction in long-term survival (nearly half in the case of infectious complications). Other authors have investigated the role of elevated levels of pro-inflammatory cytokines, as are seen in the context of infection and systemic sepsis. Increased serum concentrations of interleukins IL-1, IL-6, and IL-10, and subsequent down-regulation of antigen-presenting cells in response, may play roles in promoting tumor recurrence.\textsuperscript{179, 188-190} Mokart and colleagues reported the result of a prospective study in which 30 patients undergoing major surgery (defined as such by the authors, including oesophagectomy or pelvic exenteration) underwent serial daily blood sampling with immunoassay of proinflammatory cytokines. IL-6, in particular, was found to undergo a significant rise in the immediate (day 0 onward) postoperative period in response to surgical trauma. In a subset of patients which went on to develop systemic sepsis, this rise was two to six times higher than in controls who did not develop complications. Other cytokines related to inflammation were assessed by did not demonstrate any significant differences between groups, leading authors to theorise a proinflammatory and immunoregulatory role for IL-6.\textsuperscript{188} Laboratory studies have reported significant attenuation of acute lung injury\textsuperscript{191} and chemotoxicity\textsuperscript{192} with the activation of IL-1.
receptor antagonists. Though further research is required, it is therefore conceivable that such cytokine-mediated mechanisms may play roles in overall immune function and survival, beyond tumor recurrence alone.

The findings presented in this review, suggesting that the effects of largely preventable postoperative complications extend beyond immediate surgical outcomes to long-term survival, have several important implications. First, they highlight the importance of morbidity as key measure of care quality. Metrics such as mortality, length of stay\textsuperscript{95} or failure to rescue\textsuperscript{65} are limited in that they measure only outcomes in the immediate perioperative phase of care, whereas survival and recurrence rates measure only long-term effects. The use of morbidity rate bridges these two domains. It provides a measure of in-hospital, perioperative care quality and associated economic costs,\textsuperscript{193} as well as being significantly linked to longer-term patient outcomes and survival, as demonstrated by this review.

Second, continuing systems improvement and interventions to reduce complications must be sought. The ongoing SCIP project has sought to standardize generic postoperative care processes to reduce mortality, with some success, as part of a systemic approach to care.\textsuperscript{152} Enhanced recovery protocols (ERP) are another example of this, with protocolled processes for perioperative patient management to reduce length of stay, and potentially also morbidity and mortality.\textsuperscript{95} Additionally, measures to improve training of junior clinicians, beyond the traditional surgical focus on operative ability, must be considered to address the growing recognition of the importance of non-technical skills.\textsuperscript{194}

Finally, these findings give further impetus to the need to investigate underlying causes of preventable complications. The significant variability seen in post-operative care has been well documented in the literature surrounding failure-to-rescue.\textsuperscript{61,65} Critical care processes in the postoperative phase include the transfer of information (handovers) and the surgical
ward round (WR). Investigation into the handover process has identified regularly recurring pitfalls and irregularities which may affect clinical care, though direct links to outcomes have been more difficult to establish. Less still is known about the role of WRs in patient care. The conduct of WRs is a complex process, and is dependent on the diagnostic, management, teamworking and communication skills of the responsible clinical team. Failures to appropriately assess, diagnose, and manage potential complications appropriately will, unsurprisingly, result in poorer patient outcomes. However, further research into these and other processes at the point-of-care is required, if we are to fully understand the root causes of preventable complications and adapt practice to improve results.

By limiting the included studies in this review to those with transparent and thorough methodology, we have attempted to maintain a high standard of evidence. However, there remain limitations to this review which must be considered with the interpretation of its results. The broad scope of searching such generic terms as “morbidity,” “complications,” or “survival” meant that additional search terms were required to narrow the results, which may have resulted in relevant results being missed, though to mitigate this, varying search strategies were attempted before deciding on the final criteria, and hand-searching of retrieved reference lists used to identify additional articles. Included studies were database-based, and largely retrospective, cohort studies. As such, the quality of data, especially as relating to complications rather than primary diagnoses, cannot be assured. Whilst randomising patients to the avoidance or experiencing of complications is clearly impossible, prospective collection of large patient cohort data, at least, would be of potentially greater quality and reduce the risk of bias. Finally, Egger’s test (bias coefficient 1.69, p = 0.001) suggested significant positive skewing through publication bias. Unfortunately, current common practice in published literature tends toward the reporting of statistically significant hazard ratios resulting from multi-variate regression models only (i.e. non-
publication of negative findings). This is reflected by the under-representation of publications with negative (or less positive) findings on the lower end of the quality scale as might be expected on a normally distributed funnel plot. However, by limiting the search to high-quality studies we attempted to limit this impact. Additional sensitivity analysis with inclusion of the excluded low-quality studies in the regression analysis demonstrated no difference in outcome.

### 1.4.5 Conclusion

In conclusion, this review and meta-analysis provides evidence of a reduction of long-term and disease-free survival associated with the incidence of postoperative complications. Whilst the underlying mechanism is unclear, the implications of these findings are not: the effects of complications are not limited to the immediate postoperative period, with their sequelae felt by patients long after hospital discharge.

In the context of the preventability of many of the most common complications, there is a clear need for further action to reduce morbidity rates. Through the identification of root causes of error, intervention development, and appropriate staff training, the prevention of complications may improve patient outcomes and survival, as well as reducing the overall cost of care and burden to health systems in future.
1.5 The surgical ward round

The literature reviews presented in this chapter have sought to examine the current known evidence regarding the variable nature of surgical outcomes. These reviews have identified factors underlying this variability, with consideration of the long-term effects of the preventable complications which may result from substandard care. These highlight the postoperative, ward-based phase of care as the critical part of the surgical patient’s pathway in which adverse events occur, and failures to rescue result. Though structural factors play a role, the results presented in this thesis suggest that it is through improvement of process that the greatest improvements in clinical care may be achieved.

The primary care process of postoperative care, serving as the focal point of interaction between patient and surgeon on the ward, is the surgical ward round. In the space of a few minutes, a multidisciplinary group made up of surgeons, trainees, nursing staff and any number of allied health professions (physiotherapists, dietitians, pharmacists) must contend with a myriad of processes. These include, but are not limited to, patient assessment, diagnosis, prognosis and treatment planning, medication prescription, communication with team members, other disciplines and the patient and relatives, discharge arrangements and cost analyses, and the education of trainees. These processes must each be considered, executed, communicated and recorded, at which point the round moves on to repeat the process again with the next patient, often all whilst under pressure from competing clinical commitments, such as completing the round in time for the start of the day’s operating list. The great potential for error in such a situation should not come as a surprise.

Certain aspects of ward care present known entities of error. The well studied incidence of errors and potential for harm in drug prescribing and dispensing, for example, is illustrated in a cohort study reported by Dean et al, which observed hospital prescriptions over the course of four weeks. The authors estimated that over 36,000 prescriptions had
been written over the course of the observation period, or one prescription every 20 seconds. This study reported an overall error rate of 1.5%, more than a quarter of which were deemed potentially serious. Whilst this rate may appear relatively small, it nonetheless equates to 135 serious, and preventable, prescribing errors per month within a single academic hospital. In a further study of ward round prescribing, Looi et al observed a convenience sample of medical ward rounds over 6 months in a single centre, reporting the highly variable proportion of patient drug charts – an otherwise prominent component of daily patient care – which were being regularly checked by the responsible physician, ranging from as low as 45%.

Despite this propensity for error, and the central position of ward rounds in the delivery postoperative care, there is a paucity of research dedicated to its assessment or improvement. Thus, the scientific examination of ward rounds, to establish sources of failure and variability, and develop interventions for care improvement, forms the remainder of this thesis.

### 1.5.1 Current conduct and training for ward rounds

In an era of increasing simulation, credentialing and standardised practice, the conduct and teaching of rounds remains largely rooted in tradition. Though Halsted’s apprenticeship model of surgical training has been gradually consigned to history in many other domains, replaced by structured curricula, didactics, and proficiency-based advancement, it lives on in the case of the surgical ward round (WR). Currently, no tools for assessment of ward round quality, or teaching instruments for its instruction, exist. Rather, surgical trainees must rely on an experiential model of learning, emulating the WR conduct of their seniors, and adapting a varied model of practice through trial and error. Such a model offers few
guarantees of effectiveness. Highlighting this problem, Nikendei et al reported an observational study of 45 final year German medical students taking part in a simulated ward round.201 42% of students had already undergone specific training for ward rounds. All conducted a simulated ward round of three standardised patient scenarios, which were subsequently assessed using video review and checklists mapped to local predefined learning goals. Authors reported extremely poor performance, with failure by the majority of students to meet levels of performance expected for their level in all five learning domains, which included information gathering, communication, physical examination, chart review, and team communication. Appropriate notes were recorded in charts in only 7% of cases overall.

The body of currently published scientific evidence examining ward rounds is sparse. Previous work has examined ethnographic and organisational aspects of the WR, such as Creamer et al’s brief audit of time management within the WR. They observed four days of WRs, comparing time spent at the bedside and in transit between wards, with over an hour of travel time over the course of study period, resulting in authors calling for a minimisation of outlier patients to minimise travel times between wards.202 Other authors have reported patient surveys or first-person perspectives on the patient’s perception of the ward round and its potential to cause feelings of anxiety or intimidation among patients and junior trainees alike.203-205

Previous efforts to address deficiencies in WR conduct have largely taken the form of proformas,206 hand-over sheets and daily goals forms.144,145 In a study by Pronovost and colleagues, staff were surveyed on their understanding of daily treatment goals for patients on an intensive care unit. Prior to the study, less than 10% of staff were able to demonstrate this successfully. The introduction of templates (“daily goals forms”) to facilitate the transfer of information, specifically outlining goals for mobilisation, medication and treatment changes, and other aspects of patient care, resulted in improved volume and
quality of clinical communication, with 95% of staff able to show understanding of patient care goals, which may also have contributed to reductions in length of ICU stays in two studies.\textsuperscript{144,145} Data from the Joint Commission, the US organisation responsible for quality assessment and accreditation of healthcare organisations, have repeatedly singled out communication error as one of the primary root causes of sentinel adverse events, illustrating the importance of such an approach.\textsuperscript{207} Despite these moderate successes, however, implementation of goals forms and similar templates have been confined to small trials within medical wards and intensive care units.

1.5.1.1 The Productive Ward

A more self-directed initiative was introduced by the NHS Institute for Innovation and Improvement (now no longer in existence, largely incorporated into NHS Improving Quality) in 2007 – The Productive Ward. This series of thirteen modules – ranging from ward meals to patient assessment and communication – aims to provide a guide for nursing-led, self-directed local process improvement projects to improve patient care, communication, and ward efficiency. Though scientific assessment of the results of implementation has in part been precluded by its self-directed nature, a number of narrative publications, summarised in a recent review by Wright and McSherry, suggests that it has had a positive effect in both cost effectiveness and communication.\textsuperscript{208}

1.5.1.2 Ward rounds in medicine: Principles for best practice

In October 2012, the Royal College of Physicians (RCP) and Royal College of Nursing (RCN) issued a joint publication titled Ward rounds in medicine: Principles for best practice.\textsuperscript{209} In this document, these two national bodies highlighted the problem of WRs, stating:
“Despite being a key component of daily hospital activity, ward rounds remain a much neglected part of the planning and organisation of inpatient care. There remains considerable variability in both the purposes and conduct of ward rounds.”


Resulting from multiprofessional workshops and a review of available evidence, the RCP and RCN set out a series of recommendations for safe and effective ward rounds. Examples of expected roles for the multidisciplinary members of the WR are set out. Doctors are expected to lead the round, review charts, and update status, progress, and further management plan for the patient. They may be supported by a nurse, providing an update on clinical status and performing quality and safety checks such as pressure ulcer prevention and venous thromboembolism prophylaxis, as well as other allied health professionals (AHPs) including pharmacists and other carers.

Specific recommendations included the prioritisation of the WR as a critical part of patient care, with planned dedicated time set aside for this. An emphasis is placed upon the presence of nursing staff on the round and multidisciplinary engagement, with communication of WR results to all members involved in patient care. Doctors are advised to prepare in advance for WRs, and consider de-briefing afterward.

Despite these valuable guidelines, the document does not, in fact address how the WR itself should be conducted at the bedside. Though it recommends the use of checklists for patient safety issues, and that doctors should review drug charts for each patient during the round, it does not seek to advise the multidisciplinary team on how best to accomplish the stated aims of the round itself – that is, how to best assess the patient, diagnose any deviations from routine treatment and patient progress, and plan continued care.
There is a clear need to expand efforts to improve and assess ward rounds beyond these preliminary efforts. These must address more generalised system errors, as well as specific medical errors occurring in the course of patient care assessment and management, to improve clinician performance and patient outcomes.

1.5.2 Ward round assessment

In order to improve process, one must be able to first measure it. Quantification of performance is critical to ensuring progression of competent and safe trainees. With the introduction of revalidation by the General Medical Council in the UK, all doctors – whether trainees or not – will be required to demonstrate evidence of their competency to practice on an annual basis.

The field of medical and surgical education research has long devoted significant efforts to the development and validation of a plethora of assessment tools, resulting in a diverse selection of tools and scores with which to assess clinical practice. A 2009 systematic review published by Kogan et al summarised the then-current state of evidence for tools for assessment of clinical skills of medical trainees. Reporting the results of a search of over 10,000 articles, they found some 55 separate tools deemed suitable for the task. Of note was the generic nature of the vast majority of these tools. Some 58% of tools were intended for formative, rather than summative, assessment alone. Of the tools which claimed supporting validity evidence, all utilised semi-subjective numeric scales to describe various aspects of performance (e.g. communication, history taking, etc.), anchored by behavioural adjectives.
The most investigated and best validated tool, authors concluded, was the mini-Clinical Examination eXercise (mini-CEX). Originally developed by the American Board of Internal Medicine, the mini-CEX allows for rating of behaviours including history taking, physical examination skills, and clinical judgement, on a 1-4 scale anchored by descriptors ranging from “below expectation” to “meets expectations for completion” and “above expectation.”

Further tools continue to be developed in this area, such as the Clinic Assessment and Management Examination – Outpatient (CAMEO), which expands upon mini-CEX with the use of specific descriptive behaviours for each domain, and the potential to be used by patients to rate their attending clinician as well.

Despite the large number of assessment tools to choose from, their generic nature precludes a specific application to ward rounds. Generic assessment tools are well suited for overall assessment of trainee ability and progression, but lack the necessary detail to provide informative, summative feedback on performance for individual skills. Additionally, considering the tools assessed in Kogan et al’s systematic review, the authors noted that of the 55 scoring systems, none demonstrated any validity evidence relating to clinical performance or outcomes – instead using differences in performances of trainees of different experience levels, or questionnaire feedback. If WR performance is to be assessed and improved, it would appear necessary to consider the design of a novel assessment tool with which to specifically quantify WR performance, considering processes relevant to WR quality.

1.5.3 Development of assessment tools
The development of a novel assessment tool for surgical education is a multi-stage process, the ultimate aim of which must be to deliver a feasible, valid, and reliable system by which to measure performance of a given task or skill. This process may incorporate the following steps (figure 1.9):

- Needs analysis
- Metric selection
- Validity assessment
- Clinical correlation
- Dissemination
Figure 1.9. Stages of assessment tool development.
1.5.3.1 Needs analysis

Before developing a new tool, it is important to establish that this is also likely to fulfill a genuine need. Literature reviews, as have been described above, can be helpful in this regard, both to inform the researcher on previous work (or lack thereof) in the area of interest, as well as avoiding duplication of work already published.

Need maybe established by identifying new techniques which lack valid methods of assessment, or where existing tools suffer from significant disadvantages which could be improved upon.

Where need cannot be established from published literature, primarily in cases where there exists a paucity of data in a new field of research, researchers may consider canvassing experts and stakeholders in the relevant field instead. Establishing expert consensus can provide invaluable guidance not only on confirming need for further research, but can also help identify potentially useful metrics for incorporation into the tool being developed, and may be achieved through Delphi consensus,\textsuperscript{214} semi-structured interviews,\textsuperscript{215} or questionnaire / survey studies.\textsuperscript{216}

1.5.3.2 Metric selection

To ensure appropriate and valid measurement of the skill, or task, in question, the selection of appropriate metrics is necessary. Assessment tool metrics may take many forms, ranging from expert raters, to checklists, to computer-derived objective measurements. Particularly with the advent of computerised assessment, for example through virtual reality simulators,\textsuperscript{217} it is now possible to record a multitude of metrics for any given task. Identifying those which are most likely to represent accurate measures of skill (i.e. demonstrate validity) may be incorporated into studies of overall tool validity, or, ideally, identified in advance through feasibility or pilot studies, or expert opinion.
1.5.3.3 Validity assessment

The modern framework of validation espouses a process of accumulated validity evidence, as opposed to the model of establishing defined “types of validity” such as content or construct validity.\textsuperscript{218} Nevertheless, these terms remain useful to illustrate the fact that newly developed assessment tools must demonstrate that their chosen metrics are appropriate to measure task performance (content validity), and that these can correctly detect true differences in performance (construct validity). However, these should be considered in the context of a whole body of evidence and the context of validation studies, rather than as absolutes.\textsuperscript{219}

Researchers may consider conducting initial validity studies within a simulated environment. Simulated constructs of all types (e.g. virtual reality, animal models, benchtop, high-fidelity) have been successfully used for validation studies for a wide range of existing assessment tools.\textsuperscript{220-223} The use of simulation allows conduct of trials in a standardised fashion with replicable scenarios or tasks, within a safe environment, controlling for extraneous factors and potential confounders.

Repeated studies can help establish the feasibility and validity of assessment tools, but also give the opportunity to study reliability of the measurement tool. Rating scale-based tools, in particular, are at risk of observer or rater bias, and thus should routinely have a second rater (ideally blinded to the subject’s identity) for some, if not all, assessments. Inter-rater reliability may be statistically assessed using Cronbach’s alpha (ordinal data), or Cohen’s kappa (categorical data), as appropriate.

One must also consider acceptability of the tool for use by both the trainer and trainee. A valid assessment tool is of little value if it does not lend itself to easy use and is accepted by potential end users. Questionnaires may be used to collect formal feedback throughout
validation studies to ensure this, or to address any potential issues raised, and end-users should ideally be consulted throughout the development process. The degree of training or level of experience required to use the assessment tool must be clear, with consideration of how assessor training, if required, is to be delivered – whether in written format, or as part of a more interactive session, such as through workshops or simulation training.

Finally, cost implications of any new research tool must also be calculated. Whereas rating scales or checklists may be essentially cost-free – unless subject to intellectual property licensing – this may not be the case for technology-based methods, which may require initial outlay and / or per-use expenditure for equipment.

1.5.3.4 Clinical correlation

Having established feasibility and validity, the “gold standard” of assessment tool validation requires use in clinical practice, either alone or with correlation with simulator-based assessment of the same task. Such evidence further strengthens validity evidence for use of an assessment tool, without which use in higher stakes assessment is impossible.

1.5.3.5 Dissemination

Beyond the development and evaluation of a new assessment tool, consideration should be given to its further dissemination, and how to encourage its wider use. This may initially be driven through academic publication and conference presentations, whose impact will be augmented if appropriate steps have been taken during the development process to ensure that the tool meets an expressed need, and will want to be used by trainers, trainees, and other researchers. Replication of validation studies in other institutions, systems, or countries can provide additional strengthening evidence in the form of external validation.
Chapter 2

Identifying quality markers and improvement measures for ward-based surgical care
2.1 Introduction

Modern surgical expertise demands more than technical ability. The previously discussed variability present in failure to rescue rates is reflective of differences in the quality of postoperative patient care, rather than solely quality of surgery itself.\textsuperscript{194, 224} This variability, and its effect on surgical outcomes,\textsuperscript{61, 97} continues to drive research seeking to identify and improve underlying causative factors.\textsuperscript{67, 97, 104, 107} As presented in Chapter 1, this thesis has identified the potential of process improvement to reduce failure to rescue rates and increase survival. Despite this, it has reported the current paucity of data available on the assessment of routine care processes taking place at the bedside, and the improvement of postoperative care.

The surgical ward round (WR) represents the primary interface between the clinical team and the patient, typically involving a daily review of patient progress and prescription of the management plan for continued patient care. The WR is therefore integral to the quality of postoperative care, and the determination of patient outcomes, and is relies on a clinician’s mastery of not only skills on diagnosis and management, but also non-technical talents such as communication, leadership, and teamwork.

Recently, there has been a burgeoning recognition of the importance of non-technical skills for surgical training. Several preliminary scientific studies have reported the effectiveness of non-technical skills training programmes in improving surgical outcomes, with team training interventions linked to reductions in surgical errors, and morbidity.\textsuperscript{225-227} However, these have focused primarily on the operating room environment, rather than the post-operative phase of care.

In the UK, the previously described joint statement of the RCP and RCN in October 2012 specifically has called for a refocusing of priorities on WRs as a key point of care for quality
improvement.\textsuperscript{209} Crucially, however, this paper lacks specific guidance on how to address this training gap, reflecting the current dearth of knowledge in this area.

Traditionally, trainees have relied on the Halstedian model of experiential learning for conducting WRs, dating from the end of the 19\textsuperscript{th} century when William Halsted established the first surgical residency at Johns Hopkins Hospital in Baltimore in the style of an experience-based apprenticeship. “Teaching ward rounds” have been long noted for their variability in both quality and effectiveness.\textsuperscript{228} Particularly in surgery, WRs are often forced to compete with other clinical priorities, such as operating lists where delayed starts result in resource and financial wastage, or outpatient clinics with patients expecting to be seen in a timely fashion. Furthermore, in the past decade there has been a significant statutory limitation placed upon on duty hours, limited by the European Working Time Directive to 48 hours a week in Europe and the UK, and to 80 hours a week by similar legislation in the US. These limitations were intended to limit physician fatigue, improve patient outcomes, and improve the quality of life of residents. Recent systematic reviews have found that they have had little effect on patient outcomes,\textsuperscript{229} but may be contributing to improvement in residents’ lives.\textsuperscript{230} Despite this, however, significant concerns regarding the reduction of clinical training hours as a result of these regulations have been voiced by such senior figures as the president of the Royal College of Surgeons of England.\textsuperscript{231} Reinforcing these concerns, a recent US observational study of ward rounds observed the content and duration of medical rounds across four different hospitals ranging from local community to large academic teaching hospitals. Authors found that whilst standard assessment and care planning took place for most patients (96.7%), teaching activities were heavily relegated, with skills such as oral presentation skills (4.7% of patient encounters) or discussion of learner-identified topics (3.2%) taking place at near-insignificant levels.

Research to explore, assess, and improve surgical WRs must consider the large number of stakeholder groups involved – to address all sides of this multi-faceted process, as well as to
ensure consideration of potential confounders. While the primary surgeon will determine the course of patient care during the WR, they are not alone in determining outcomes. Decisions are made with input from junior clinicians, care delivered by nursing staff and allied health professionals, and experienced by patients, all of whom contribute significantly to the resulting patient outcome.

The aim of this study was to explore key issues pertaining to the surgical ward round, drawing on experiences of multiple stakeholder groups to define problems in current WR practice, and identify potential means of quality assessment and improvement.

2.2 Methods

A qualitative, semi-structured interview-based approach was adopted to explore issues around surgical ward rounds and post-operative care. An institutional research ethics committee approved the study, reference 13/WM/0260.

2.2.1 Subjects

Interviewees were purposely selected to represent the main multidisciplinary roles involved in a general surgical ward round, and to give voice to all major stakeholders in post-operative care: patients, nurses, junior trainees (house officers), senior trainees (specialist registrars), and consultants.

Clinical staff participants were sampled from several regional hospitals, and selected to include a broad range of backgrounds and experience levels. Patient interviewees were identified from the inpatient register of a single tertiary academic surgical centre and
considered for inclusion in the study if they had undergone a general surgical procedure during their current admission.

Patients were excluded from consideration if they were deemed to lack mental capacity (either acutely, or owing to chronic problems such as dementia), or did not possess adequate language skills to conduct a full interview in English. Patients were briefed on their right to refuse participation on any grounds. All interviewees gave written consent to their participation in the study.

2.2.2 Setting

In order to ensure patients would be well be enough to participate in an interview, and to minimise any concerns that the interview might affect their care, subject interviews were limited to the day of, or day prior to, planned discharge. Interviews took place in a private office on the general surgical ward, with an independent interviewer who was identified as a researcher, and not a member of the surgical care team.

2.2.3 Interview process

Subjects were interviewed using a pre-determined, semi-structured interview protocol, which was piloted with senior surgical trainees prior to use in final subject interviews (appendix 1). The interview protocol covered the following topics with reference to ward-based care in general surgery, with a focus on clinician-led processes:

- Introduction and demographic details
- Identifying current problems in ward-based surgical care (i.e. variability)
- Defining the surgical ward round
• Required skill set
• Identifying quality markers for ward rounds
• Improving surgical ward round quality

2.2.4 Data analysis

Recorded interviews were anonymised and transcribed by a third party not familiar with the participants. Identifying phrases and names were removed. The resulting anonymised transcripts were analysed by a member of the research team, with emerging themes recorded until thematic saturation was reached. A second blinded researcher acted as second reviewer, analysing 20% of transcripts selected using a computer generated randomisation sequence, to ensure reliability of theme extraction. Themes mentioned by each interview subject were coded by both researchers independently, overarching theme content unified between reviewers via consensus discussion within the research team, and reliability of thematic extraction statistically assessed using Cohen’s kappa.

Statistical analysis was performed in IBM SPSS 21 (IBM Corp, Armonk, NY). A p value of less than 0.05 was considered statistically significant.

2.3 Results

Twenty-five interview subjects were recruited from eight hospitals across southern England, UK – two tertiary academic centres (16/25 subjects, 64%) and six district general hospitals (9/25 subjects, 36%) – with varying levels of experience (table 2.1). Recruited consultants
comprised three colorectal surgeons, one bariatric and upper gastrointestinal surgeon, and one vascular surgeon.
Table 2.1 Participant demographics

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>M : F</th>
<th>Academic : DGH</th>
<th>Experience in post (median, range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consultants</td>
<td>5</td>
<td>4 : 1</td>
<td>4 : 1</td>
<td>9 (1 – 22) years</td>
</tr>
<tr>
<td>Registrars</td>
<td>5</td>
<td>3 : 2</td>
<td>2 : 3</td>
<td>PGY 7 (5 – 10)</td>
</tr>
<tr>
<td>House officers</td>
<td>5</td>
<td>2 : 3</td>
<td>1 : 4</td>
<td>PGY 1 (1 – 1)</td>
</tr>
<tr>
<td>Nurses</td>
<td>5</td>
<td>2 : 3</td>
<td>5 : 0</td>
<td>9 (2 – 20) years</td>
</tr>
<tr>
<td>Patients</td>
<td>5</td>
<td>3 : 2</td>
<td>4 : 1</td>
<td>21 (3 – 35) inpatient days</td>
</tr>
</tbody>
</table>

DGH: district general hospital, PGY: postgraduate year.
Inter-rater reliability of coding was good, as measured by Cohen’s kappa = 0.724 (p < 0.001).

Themes for each interview subheading were summarised and are presented below.

2.3.1 *Current problems in ward-based care*

Almost all subjects (23/25, 92%), including all clinical staff (20/20, 100%), believed that there was significant variation in the quality of surgical WRs, and that this had the potential to affect patient care (see table 2.2):

“...There are differences in quality and activity. Some [clinicians] are good with the complex amount of information which needs to be processed, some are not. Some are good with patients, some are good with staff, some are bad with both. So we see huge variation, yes.”

(Subject 24, consultant)
<table>
<thead>
<tr>
<th>Table 2.2</th>
<th>Interviewees’ opinions on current WR practice</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
</tr>
<tr>
<td>Variation exists and affects care</td>
<td>23 (92%)</td>
</tr>
<tr>
<td>Causes of variation in WR quality</td>
<td></td>
</tr>
<tr>
<td>Thoroughness of round</td>
<td>18 (72%)</td>
</tr>
<tr>
<td>Communication quality</td>
<td>12 (48%)</td>
</tr>
<tr>
<td>Structured approach</td>
<td>9 (36%)</td>
</tr>
<tr>
<td>Team approach</td>
<td>9 (36%)</td>
</tr>
<tr>
<td>Expectations of the WR</td>
<td></td>
</tr>
<tr>
<td>Patient assessment</td>
<td>25 (100%)</td>
</tr>
<tr>
<td>Define management plan</td>
<td>23 (92%)</td>
</tr>
<tr>
<td>Review progress</td>
<td>18 (72%)</td>
</tr>
<tr>
<td>Communication with patient</td>
<td>16 (64%)</td>
</tr>
<tr>
<td>Communication with team</td>
<td>13 (52%)</td>
</tr>
<tr>
<td>Required skill set</td>
<td></td>
</tr>
<tr>
<td>Communication skills</td>
<td>25 (100%)</td>
</tr>
<tr>
<td>Patient assessment / history taking</td>
<td>18 (72%)</td>
</tr>
<tr>
<td>Diagnostic ability / knowledge</td>
<td>17 (68%)</td>
</tr>
<tr>
<td>Teamwork / multidisciplinary cooperation</td>
<td>11 (44%)</td>
</tr>
<tr>
<td>Leadership and management skills</td>
<td>10 (40%)</td>
</tr>
</tbody>
</table>
The most common cause identified was the degree of thoroughness with which the WR was conducted (18/25, 72%), followed by poor communication (12/25, 48%) and the lack of a structured approach (9/25, 36%).

“If you’re not thorough, then you can miss things that are easy to sort out in the beginning, but if you let them drag on then it can be much harder to sort them out or the patient can get sicker before you really realize that there’s a problem.”

(Subject 10, house officer)

2.3.2 Defining the surgical ward round

In response to being asked to define their expectations of the surgical WR, the majority of subjects interviewed identified similar themes. The WR was defined as a process of patient assessment (25/25, 100%), during which patient progress was reviewed (18/25, 72%) and a management plan defined (23/25, 92%). It was also felt to represent a forum for communication with the patient (16/25, 64%) and for communication within the multi-disciplinary clinical team (13/25, 52%):

“[The purpose of the WR] is to assess the progress of the patient in regard to their treatment. That means picking up problems which may or may not be anticipated, the patient’s concerns, their worries. And all that information should be passed across... through the other members of the team.”

(Subject 24, consultant)
2.3.3 Required skill set

Interviewees identified clinicians’ skills which they deemed critical to the conduct of high quality WRs. Above all, communication was identified as the most important (25/25, 100%), with implications for both clinical performance, as well as patient experience:

“Communication is essential... professionalism generally follows”

(Subject 2, consultant)

“Some people, you really feel like you’re working as a team... sometimes you feel that you’re completely not, which is not very nice because you’re the [nurse] who’s there the whole time. And you pick up on the way the patients feel because they notice straight away... it’s not very nice to hear ‘oh they weren’t very nice were they?’”

(Subject 4, nurse)

Related skills such as history taking and patient assessment (18/25, 72%) were also cited, as well as teamwork (11/25, 44%), leadership and management skills (10/25, 40%).

2.3.4 Identifying quality markers for ward rounds

Subjects identified a number of processes to be included in the ideal ward round, in the context of assessing an unwell post-operative patient. These were felt to potentially be utilised as markers of ward round quality (see table 2.3). Processes were divided into processes of patient assessment and management. Assessment tasks included physical
examination (21/25, 84%), taking a recent history from the patient (21/25, 84%), and checking of vital sign charts (21/25, 84%). Management tasks included defining further investigations or interventions necessary (19/25, 76%), dietary status (14/25, 56%), and ensuring appropriate documentation (13/25, 52%).
<table>
<thead>
<tr>
<th></th>
<th>n (%)</th>
<th>Consultant</th>
<th>Registrar</th>
<th>House officer</th>
<th>Nurse</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient assessment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical examination</td>
<td>21 (84%)</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>History from patient</td>
<td>21 (84%)</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Status charts (i.e. vitals, fluid balance)</td>
<td>21 (84%)</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Prescription chart</td>
<td>19 (76%)</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Pathology and lab results</td>
<td>18 (72%)</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Review medical notes</td>
<td>17 (68%)</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Radiology results</td>
<td>17 (68%)</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>History from nurse</td>
<td>12 (48%)</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>VTE prophylaxis</td>
<td>9 (36%)</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Assess analgesia</td>
<td>6 (24%)</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td><strong>Patient management</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investigations or interventions</td>
<td>19 (76%)</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Dietary status</td>
<td>14 (56%)</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Appropriate documentation</td>
<td>13 (52%)</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Communicate plan to nursing staff</td>
<td>12 (48%)</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Appropriate prescribing</td>
<td>12 (48%)</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Blood tests if required</td>
<td>11 (44%)</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Discharge planning</td>
<td>10 (40%)</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Communicate plan to team</td>
<td>9 (36%)</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Communicate plan to patient</td>
<td>8 (32%)</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

VTE: venous thromboembolism
In reference to processes determining patient management during the WR, there was a marked divide between clinical staff and patients. Whereas there was broad agreement between clinical staff, which named a median of 6 themes (range 2-9), patients named far fewer tasks, 2 (1 – 4), possibly reflecting a lack of clinical knowledge by lay patients in this domain.

Most frequently named by patients was the expectation that the management plan be communicated to the patient (3/5, 60%). This varied between patients, however, with some patients less interested in the details of their ongoing care, whereas others wished to have full and active involvement:

“*I’m not too bothered* [about knowing what’s going on with me in hospital]”

(Subject 19, patient)

“I want to know everything. *Partly because if I know everything, I can push to keep the plan. Like I knew about* [my prescribed diet], *so I had to call the junior doctor because the caterer didn’t know anything about it. Every time it turns out they need to take blood, I say ‘what for?’*”

(Subject 21, patient)

None of the patient interviewees expressed a desire to explicitly not be told of their ongoing management.

2.3.5 Improving surgical ward round quality
All subjects agreed that measures to improve surgical ward round quality were needed, with a number of potential interventions identified (table 2.4).

The most commonly mentioned intervention was to implement scenario- or simulation-based training (13/25, 52%):

“You could simulate a ward round... the benefit is that you could reproduce the same patients so you could create the complex patients that you wouldn’t otherwise have, necessarily, when someone is watching you [during routine clinical practice]. And you’ve got to remember [during a clinical teaching WR] you’ve got a patient here, do you want to start scaring them... going through the vast array of complications you can get? It’s not really appropriate to do that in front of the patient.”

(Subject 8, registrar)

Checklists and proformas for ward rounds were also named as a potential intervention to improve WR performance (12/25, 48%). Some interviewees strongly advocated this approach, citing the empowering nature of checklists and the potential for these to ensure basic essential tasks are reliably completed (subject 3, consultant). However, others argued that checklists did not allow the necessary flexibility for different WRs between, for example, unwell postoperative patients or fully recovered patients about to be discharged (subject 7, registrar). Another concern raised was that checklists might limit clinicians’ own free interpretation of the patient by enforcing narrow prompts (subject 24, consultant).
Table 2.4 Recommendations for WR quality improvement and training

<table>
<thead>
<tr>
<th>Recommendations for Improvement</th>
<th>n (%)</th>
<th>Consultant</th>
<th>Registrar</th>
<th>House officer</th>
<th>Nurse</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulation- or scenario-based training</td>
<td>13 (52%)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Checklist or proforma</td>
<td>12 (48%)</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Definition of standardised approach</td>
<td>8 (32%)</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Regular formalised trainee assessment</td>
<td>7 (28%)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Top-down cultural change</td>
<td>7 (28%)</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Communication training</td>
<td>2 (8%)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Increase staffing and time available</td>
<td>2 (8%)</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
In addition to training interventions, the need for a formalised system of trainee WR assessment was specifically mentioned by 7/25 (28%) of subjects. Others mentioned the need for top-down cultural change to improve the quality of WRs and training overall (7/25, 28%). Modifications to structural factors were also identified, such as staffing levels to increase the amount of clinician time available for WRs (2/25, 8%).

2.4 Discussion

This study presents an exploration of the current problems, and solutions, pertaining to the practice of surgical ward rounds. It identifies potential quality markers which might be used to measure, and interventions with which to improve, this critical process of postoperative care. Additionally, this is the first time all stakeholder groups involved in WRs – attendings, residents, interns, nurses, and patients – have been included in such a study.

The results of this study indicate clear agreement, across interviewees from eight different hospitals and diverse professional backgrounds, that current WR practice contains unacceptable levels of variability, placing patients at risk. Reflecting the already described variation reported in failure to rescue literature and by surgical outcomes researchers in general, this places further emphasis on the need for assessment and improvement of WRs.

In seeking to address these issues, this study has been the first to attempt to define various stakeholders’ expectations of the WR, as well as the skill set required to meet these. The identified processes expected of a routine WR included patient assessment, management, progress review, and communication with both patient and clinical team, broadly in agreement with the definition of WRs set out by the Royal College of Physicians – Royal College of Nursing joint statement. In identifying a required skill set, emphasising non-
technical skills such as communication, teamwork, and leadership skills, this study reflects the growing perceptual shift of surgical training in general, which continues to expand beyond the traditional scope of operative skill alone.

A number of potential sub-processes which might be used to assess WR quality, were identified in this study, with broad agreement amongst interviewees. The identification of these key processes (table 2.3) suggests the use of thoroughness of assessment as a marker of WR quality. Certainly, not all processes may apply to all patients (i.e. not all patients might have radiological investigations to review, or require specific dietary planning to consider). However, though the degree of thoroughness of assessment during the WR will be naturally tailored according to perceived patient necessity and clinical priority, it is intuitive that patients will universally benefit from a comparatively more thorough assessment. The casual greeting of patients without physical examination or review of charts, for example, may indeed convey more information than immediately apparent – an appropriately responsive patient might be assumed to be adequately analgesed, possess sufficient cerebral perfusion, normal blood glucose levels, and be free of signs of acute confusion or delirium – but it cannot reveal the overnight pyrexia, the decreased oxygen saturation level, or the erythematous wound. Accordingly, a “lack of thoroughness” was identified in interviews as a major perceived source of variability in WR quality:
“More senior [clinicians], these ones tend to take a little more time looking at all the information. They do actually come up with more jobs for us to do as juniors... They can trouble-shoot problems before they really start and I think that’s what keeps [patients] safe.”

(Subject 14, house officer)

The processes identified in this study may thus act as a guide in defining WR thoroughness and the development of quality tools, which is discussed in detail in the following chapters of this thesis.

One area in which a lack of agreement was seen among interviewees was the discussion of the agreed management plan with the patient. 60% of patients (3/5) expressed a strong wish to be told of their management plan, contrasting with only 25% (5/20) of clinical interviewees who named this as a necessary part of WRs. In the modern era of patient choice and patient-related outcome measures, such a disconnect serves only to further highlight the need for increased awareness and training in this area of care.

Finally, this study identified a number of potential interventions for WR improvement, including simulation-based training. The benefits of simulation, allowing trainees to train in a high-fidelity, controlled environment, with the elimination of risk of patient harm, has resulted in the wide uptake of simulation-based training for technical skills in surgery, with an established body of evidence demonstrating its efficacy.\textsuperscript{32, 232, 233} A recent survey of North American program directors has supported the increased adoption of simulation generally to train surgeons’ non-technical skills.\textsuperscript{234} The results of this interview study suggest there is a perceived need for the application of simulation specifically for WR training. Additionally, other recommendations for improvement named by interviewees in this study, such as
regular trainee assessment and communication training, would be well suited for incorporation into a ward simulation-based training scheme or curriculum in future.

Checklists or proformas were named as a further potential intervention to improve care. However, despite support from 12/25 (48%) interviewees, several subjects also stated their objections to the use of checklists in this area of care. These cited the perceived necessary variation and flexibility of WRs between patients, according to their physiological status and level of clinical concern, which cannot be accommodated by uniform checklists. Checklists, in the form of daily goals sheets, have been effectively used to improve communication quality in the context of intensive care units (ICUs). However, as the ICU environment represents a highly standardised level of patient care, it is possible that principles which work in the ICU setting may be inappropriate to transfer to the context of surgical ward care.

One methodological strength of this study has been to include the full breadth of stakeholders involved the surgical WR. The RCP/RCN statement on WRs, for example, has particularly emphasised the necessity of multidisciplinary involvement for good WR practice. Our selection of interviewees across the relevant health professions has reflected this, but has resulted in a relatively small number of interviewees per group. Despite this limitation, thematic saturation was reached, with broadly similar recommendations divined from all groups. Though a larger sample size might have allowed greater differentiation between groups, this was not the primary aim of this study, which was to explore ideas and perceptions across groups on the topic of surgical WR practice.

2.5 Conclusion
In conclusion, this study confirms that clinical staff and patients alike perceive there to be significant variability in current surgical WR practice, which places patients at risk and affects care. It highlights non-technical skills, rather than medical knowledge or technical ability, as a target for improvement of surgical care. Interviewees identified perceived quality markers which may be used for assessment of WR quality. Further development of interventions, such as ward simulation, to assess and improve surgeons’ and surgical trainees’ performance is needed to standardise and improve both WRs patient care.
Chapter 3

The simulated ward environment and assessment tools for ward-based surgical care
3.1 Development of a simulated environment for ward round assessment and training

3.1.1 Introduction

Financial and political pressures on health care systems and surgical training have had the effect of reducing overall time spent in training and in the operating theatre, whilst the imperative for a high level of quality to be provided to patients by appropriately trained surgeons has remained the same.\textsuperscript{235} Simulation – defined as “the imitation of the behaviour of some situation or process by means of a suitably analogous situation or apparatus”\textsuperscript{236} – allows trainees to acquire and practice skills in a safe, high-fidelity environment without risk to patients,\textsuperscript{237} with enhanced retention of information and transfer of taught skills from the classroom to the clinical setting compared to standard didactic training.\textsuperscript{238} It enables trainees to learn in a dedicated educational environment, monitor their progress, receive feedback on performance, and to enhance patient safety.\textsuperscript{239}

The formalised use of simulation-based education dates back to almost half a century ago, to the development of task trainer manikins for cardiology trainees at the University of Miami.\textsuperscript{240} Since the 1970s, this has been greatly expanded across a number of domains of health education. Particularly within surgery, simulation has been broadly adopted for the training of technical skills, providing an analogue to patients on which trainees might practice and develop.\textsuperscript{241} Modern technical skills simulators range from benchtop simulators, such as simple foam suture pads, to animal or human cadavers, to complex virtual reality computer simulations.
The use of simulation extends also to non-technical areas of practice. Other, non-technical skills may also be taught through simulation, such as through the use of actors to train students in history taking, communication skills, or other clinical competencies.\textsuperscript{242} Additionally, there has been growing interest the specific training of “soft” skills such as teamwork, leadership, and communication, such as through dedicated team training sessions for entire clinical or operative teams.\textsuperscript{243}

The uptake of simulation training for surgical technical skills continues to gather pace, in parallel with the growing body of supporting evidence. Several reviews\textsuperscript{238, 244} and a meta-analysis,\textsuperscript{245} comparing surgical training schemes with and without enhancement through simulation, have demonstrated the benefits offered by surgical simulation. These have demonstrated improvements in technical skill, reduction of errors, and minimisation of operative times. In the United States, the Accreditation Council for Graduate Medical Education (ACGME), the national body responsible for accreditation of medical training programs and curricula, now mandates the inclusion of simulation in educational curricula, coupled with regular assessments of skills competency. Completion of the Fundamentals of Laparoscopic Surgery (FLS) simulation curriculum is now a prerequisite to completion of surgical residency. FLS, developed at McGill University in Montreal, Canada, is a comprehensive simulation-based training programme for basic laparoscopic skills and includes web-based study guides and a laparoscopic box trainer, and incorporates an assessment of performance on five technical tasks upon completion of the curriculum.\textsuperscript{246}

Studies have demonstrated the improvement of skills which may be achieved through the FLS curriculum, which have been shown to translate to improved performance in the operating theatre, as in Sroka et al’s study of 19 residents conducting a laparoscopic cholecystectomy following randomisation to either a control group, or one incorporating FLS training.\textsuperscript{247} Building on these successes, the Accredited Education Institutes programme of the American College of Surgeons has, since 2007, built a gradually expanding international
programme of accreditation for “comprehensive” and “focused” simulation centres, to encourage the further proliferation of high standard simulation training.

Meanwhile, the equivalent body of the ACGME in the UK, the Joint Committee on Surgical Training (JCST), has more recently signaled a shift towards a similar approach. In their most recent 5-year strategy document published in 2013, the JCST has placed much greater emphasis on the inclusion of simulation in surgical curricula. Whilst stopping short of formalising an actual mandate for adoption, it has specifically named as a priority “to maximise the use of simulation techniques in surgical training. We believe that this is an essential step to improving patient safety.”248

Whereas simulation in the past has focused primarily on technical skills training, it is now also being applied increasingly outside of the operative arena. Research has shown that the performance and interaction of trainees with trained actors as simulated patients is comparable to their performance with real ones,249,250 and simulated patients are now regularly used for training non-technical skills such as communication251 and teamwork in crisis management.252 Larkin et al employed simulated patients as part of a novel curriculum for first year residents specifically designed to address a training gap they had identified for communication, team working, and stress management. Video-based assessments of trainee performance were able to demonstrate significant improvements in empathic communication skills upon completion of the curriculum, taught across multiple half-day didactic and simulation-based sessions.251

In addition, modern simulated operating theatres allow the integration of both technical and non-technical skills simulation, allowing the simultaneous assessment and development of both.239,253 Moorthy et al described the validation of the simulated operating theatre environment, assessing separate cohorts of junior and senior trainees on their management of intraoperative crises (intraoperative haemorrhage).254 In this study, though significant
differences were found between groups in the assessment of technical skill, for which the (surrogate) primary outcome was taken to be volume of blood loss. However, no differences were found for non-technical skill, utilising a modification of the NOTECHS scale (which shall be discussed in greater detail later in this thesis), which authors suggested might be symptomatic of a lack of dedicated training in this area, and that training should consider both non-technical and technical skills in future.

The evidence examined thus far as part of this thesis have suggested the potential of simulation to play a significant role in assessment and improvement of ward-based care. Different from existing single-patient human or computerised human patient simulators which simulate case-based or crisis management, simulation of the ward as a whole may address global patient care processes and system error affecting all patients. However, as with any simulator, there are significant cost and resource implications associated with the development of a new simulation facility. This chapter aims to discuss the process of establishing a ward simulator, taking into account curriculum, users and facilities, and describing the experience of setting up a ward simulator as part of an American College of Surgeons Level 1-accredited education centre.
3.1.2 Developing a simulation ward

3.1.2.1 Identifying the end-user

It is important to identify the end-user in order to allow tailoring of resources to the intended audience. The intended user group will impact upon ward design, choice of equipment and indeed also the sources from which funding of the project might be sought. Experienced trainees are likely to have more advanced learning needs requiring more specialised equipment or complex ward scenarios than students being exposed to a clinical environment for the first time. Whilst the basic setup of any ward simulator should be the same, the need for details such as the inclusion of a high dependency environment, specialty monitoring equipment or interventional tools, for example, will need to be determined on the basis of the target audience.

3.1.2.2 Fidelity, facilities and equipment

The physical spaces and furnishings required to create a simulated ward merit careful consideration. Most simulated wards are likely to be setup within an existing skills centre or educational building with links to a training hospital or university. Space is often at a premium within pre-existing centres. However, it is important to provide a large enough space to set up a ward capable of accommodating multiple beds and associated equipment to ensure a high-fidelity environment. The demonstrable advantages of high-fidelity simulation over traditional didactic teaching mean simulated ward furnishings should be as close in style and setup to those used on actual wards as possible. For purposes of teaching and assessment, space for observational equipment including audio and video recording capability must be accounted for also.
Aside from the ward itself, access to formal lecture halls or classroom space as well as toilet facilities must be considered and are likely to be shared within existing buildings.

Appropriate electronic and networking connections, including internet or intranet access, should be provided. In addition, an internet web page will provide a mode of publicising courses and teaching material, as well as a means of potential candidates to register for courses and simulation sessions.

Staffing considerations which will include teaching, administrative / management and maintenance staff may be shared out as additional responsibilities to existing employees or to purposely created positions and must also be accounted for in considering budgeting and office facilities.
Figure 3.1 The simulated ward at Imperial College London – St Mary’s Hospital, London, UK
3.1.2.3 Scenarios

The use of simulation as an educational tool relies on realistic scenarios which can be carried out in the simulated environment and which are tailored to learning objectives set out at the beginning of each training course. Desired clinical topics for training are chosen based on the learning needs identified. The creation of appropriate scenarios must have real patients and their experiences as a starting point for simulated patient (actor) role development; the continuing involvement of real patients in the development and implementation of scenarios shall ensure their appropriate calibration to real world experiences and increase simulation fidelity. Running a simulated scenario entails costs associated with the hiring of patient actors. For St Mary’s Hospital, these are costed at £25 per hour per actor, or approximately £300 per 4-hour half day. These costs compare favourably to the cost of other simulators, such as benchtop disposable models for technical skills training (e.g. open inguinal hernia repair model, Limbs & Things, Bristol, UK, £286 per model), or the cost of running a simulated operating suite (up to £500 per day).

Versatility in adapting a simulator to a wide range of learning objectives and scenarios is one of the inherent advantages of simulation. A simulated ward has the potential to be applied to a large range of common clinical scenarios relevant to modern medical practice. This may include communication scenarios, with the potential assessment of inter-professional communication within the clinical team, as well patient-physician communication, but also clinical scenarios focusing on medical assessment and management. The management of the post-operative patient having undergone a specific procedure, for example, might be simulated with several simulated patients representing different stages of post-operative recovery. Within a simulated ward, one bed may have a patient immediately postoperative, another 2 days post-op and others either having developed complications requiring intervention and further management, or nearing discharge. Thus the entire spectrum of patients that are likely to be encountered following a certain procedure can be experienced
within a short period of time. An example scenario is detailed in figure 3.2. Such scenarios are ideally based on actual cases, whilst taking into account sufficient modifications necessary to address ethical questions and prevent patients being identified. Case design and complexity must be appropriate to the intended learning objectives and learner group, and developed in collaboration with clinical experts to ensure content validity.
Figure 3.2  Example simulated patient scenarios for 3-bed ward, with expected assessment
tasks by trainee reviewing a single patient through multiple stages of care

Simulated patient scenarios for Mr John Smith, aged 67, post left hemicolectomy.
PMHx: hypertension, hypercholesterolaemia

Bed 1: Day 0
Mr Smith has undergone a left hemicolectomy, following a diagnosis of a T3N0 tumour. The operation went well and he was transferred from the recovery room to the ward 2 hours ago. He is still tired from the operation, but easily rousable. The nursing staff report that he has appeared comfortable since returning from theatre though he has intermittently complained of nausea. He has not retched or vomited. He has an epidural in situ for analgesia, as well as a urinary catheter and IV fluids running.
- Check drug chart: regular antiemetics, laxatives, appropriate analgesia
- Consider oral fluids as tolerated if appropriate
- Consider need for IV fluids
- Consider removal of catheter
- Early consideration of discharge plans / social circumstances
- Patient to be mobilised with physiotherapists

Bed 2: Day 1
Mr Smith is recovering from his left hemicolectomy which he underwent successfully yesterday. Following the operation he was sat out in his chair for several hours, and managed to drink several cups of water yesterday, with supplemented IV fluids overnight. His catheter was removed yesterday following the operation and he has been passing urine with the aid of a bottle. He has not yet opened his bowels or passed flatus. His nurse reports that he has been comfortable throughout and the epidural appears to be functioning appropriately.
- Commence full diet
- Mobilisation
- Discontinue IV fluids
- Ensure discharge plan in place

Bed 3: Day 2
Mr Smith is now day 2 post-op and has been progressing according to plan. He has been mobilised along the corridor with the aid of the physiotherapists yesterday. He has returned to a full diet and the dietician is satisfied with his oral intake. He reported passing some flatus yesterday evening.
- Early removal of epidural with appropriate oral analgesia
- Mobilisation
- Prepare discharge
3.1.2.4 Assessement Tools

Gauging the effectiveness of any learning tool requires a means of assessment to quantify performance.\textsuperscript{259} Objective and reproducible assessment provides feedback to the student, identifying weaknesses and charting progress.\textsuperscript{260} Appropriate to the selection of simulation modules and scenarios, availability and choice of assessment method must also be considered and will depend on the learning goals at hand. Assessment of performance in surgical procedures, both in real-world and simulated environments, is achieved primarily by breaking down a given procedure into a checklist of a number of key tasks.\textsuperscript{261, 262} Within a simulated ward environment, task-based checklists may be effective to assess performance for simpler scenarios or learning goals, such as ensuring that a ward round is conducted appropriately with all relevant charts being checked for each patient.\textsuperscript{263} For the more summative assessment of complex tasks, however, more thorough assessment frameworks are required. Examples include the Observational Teamwork Assessment for Surgery (OTAS)\textsuperscript{264} and Non-technical Skills for Surgeons (NOTSS),\textsuperscript{265} which assess overall behavioural and communication performance for non-technical tasks, or the assessment of the Patient Assessment and Management Examination (PAME)\textsuperscript{266} to evaluate integrated patient assessment and management. In areas where appropriate methods of assessment do not exist, it may be necessary to pursue the development of the same. Video recording of student performance in the simulator can be used for self-assessment and to emphasise feedback to the student, but also to create further educational and promotional media. Finally, involvement and consultation of real patients in the creation and implementation of scenarios can help to achieve a high level of fidelity.\textsuperscript{258} User and patient feedback on the ward environment itself will also serve to improve quality of the simulator, suggest improvement and direct further development of the simulated ward.
3.1.2.5 Multidisciplinary involvement

Training within the simulated ward should reflect the multidisciplinary nature of real-life wards rather than be limited to medical students or trainees, though controlling for additional variables by limiting training to specific groups may be beneficial particularly for junior trainee groups, or multidisciplinary groups of significantly different experience levels. Breakdowns in teamwork and team communication, particularly between medical and nursing staff, are major sources of medical error.\textsuperscript{1,267} In the Quality in Australian Health Care study, for example, Wilson et al found communication breakdowns between members of the clinical team to be leading cause of adverse events and medical harm, responsible for over twice as many deaths as clinical decision failures.\textsuperscript{2} Donchin and colleagues reported the results of an observational study of error on an intensive care unit, wherein a 24-hour period of direct beside observation was conducted on 46 randomly selected patients. 554 errors were recorded in this period. Significantly, despite communications between nursing staff and doctors making up on 3\% of all recorded activities, they were implicated in 37\% of all errors.\textsuperscript{267}

Training involving entire multidisciplinary teams has been effectively implemented in the operating theatre environment. Wolf et al reported the results of a team training intervention conducted for operating theatre teams. Up to two years following the intervention, authors reported significant reductions in delays, equipment issues, antibiotic administration, and hand-off problems.\textsuperscript{268} Similarly, fully realised curricula for WRs should therefore seek to involve allied healthcare professionals from a multitude of disciplines. Just as simulated operating suites have proven an excellent modality for training communication and other non-technical skills in the operating room,\textsuperscript{269,270} ward simulation has the potential to do the same, within the context of post- and peri-operative care.
3.1.2.6 Specialist wards

In accordance with intended end-users and learning objectives, the adaptation of ward simulators for the simulation of specialist ward environments is also possible. Examples of this may include paediatric, intensive care or psychiatric wards. Whilst the basic layout and bed space within the simulated ward may remain the same, each of these specialist wards will have unique furnishing and equipment requirements to enable realistic scenarios to be implemented and maintain a high level of fidelity. To ensure the ability to make use of the full potential breadth of application of the simulator, these requirements should be considered when setting up a simulated ward.

3.1.2.7 Cost

The expense of setting up and maintaining a surgical ward simulator depends on a multitude of factors, many of which have already been touched upon. In addition, costs to be considered include those for procuring or preparing appropriate space along with the associated facilities maintenance costs, audio-visual and information technologies equipment and connections, staffing of both the ward itself as well as administration thereof, not forgetting the furnishing of the ward with appropriate hospital equipment. Facilities and staffing costs are likely to be shared within a larger centre with other simulation or training units. Fixed costs unique to the ward simulator will be incurred in the appropriation of ward furnishings; the cost to furnish a 4 bed ward simulator is approximately £5,000, with a breakdown of sample costs detailed in table 3.1.
<table>
<thead>
<tr>
<th>Item</th>
<th>Quantity</th>
<th>Cost (£)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chart boards</td>
<td>4</td>
<td>48</td>
</tr>
<tr>
<td>Folding screen separator</td>
<td>2</td>
<td>552</td>
</tr>
<tr>
<td>Alcohol gel dispenser</td>
<td>4</td>
<td>32</td>
</tr>
<tr>
<td>Apron dispenser</td>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td>Glove dispenser (3 column)</td>
<td>1</td>
<td>85</td>
</tr>
<tr>
<td>Intravenous drip stand</td>
<td>4</td>
<td>251</td>
</tr>
<tr>
<td>Medical records trolley</td>
<td>1</td>
<td>303</td>
</tr>
<tr>
<td>Case file pockets</td>
<td>1</td>
<td>40</td>
</tr>
<tr>
<td>Rubbish bin (non-clinical waste)</td>
<td>1</td>
<td>86</td>
</tr>
<tr>
<td>Rubbish bin (clinical waste)</td>
<td>1</td>
<td>86</td>
</tr>
<tr>
<td>Bedside cabinet</td>
<td>4</td>
<td>1,794</td>
</tr>
<tr>
<td>Bedside table</td>
<td>4</td>
<td>1,020</td>
</tr>
<tr>
<td>Ward equipment storage cabinet</td>
<td>1</td>
<td>400</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td><strong>4,724</strong></td>
</tr>
</tbody>
</table>
3.1.3 Conclusion

The medical imperative of *primum non nocere* cannot be fully realised by relying on the traditional method of supervised learning on actual patients. Increasing evidence supports simulation as an efficient and cost-effective way in which to shorten the learning curve. The American College of Surgeons has endorsed simulation in surgical training and in 2006 begun the formal accreditation of surgical simulation Education Institutes. This has contributed to the growth in numbers of simulation centres, with over 60 institutions now Level I and Level II accredited worldwide. These are mandated to provide surgical education for both procedural and non-procedural skills, with simulated wards one of the potential tools in a simulation centre’s arsenal.

There is a need to investigate the potential benefits of ward simulation and to expand the currently small body of published evidence. An effective assessment tool for complex ward simulations such as simulated ward rounds must be determined. Analogous to studies which have been conducted for other forms of simulation, once validated assessment tool and scenarios have been developed, the next step is to validate the simulated ward concept by studying skills transfer to clinical practice.

The use of the simulated ward, however, need not be limited to surgical or medical trainees alone. Already, nurses and nursing students are being incorporated into simulated ward-based introductory training. With the advent of revalidation in the UK, requiring regular reviews of fitness to practice, practitioners of all training grades may benefit, in addition to other allied health professionals, and patients themselves.
3.2 Validation of the simulated ward environment and Surgical Ward care Assessment Tool

3.2.1 Introduction

Surgery has evolved over the past 20 years in pursuit of objective assurances of quality and developing a culture of safety. Varied interventions to improve surgical practice in the operating theatre have been successfully implemented, from ensuring safe minimum operative volumes to error-reduction and skills-training strategies from other high risk professions such as the aviation industry. Surgical simulation remains a leading example of this, enabling trainees to acquire skills on high fidelity models, whilst benefitting from the advantages of a dedicated educational environment which eliminates risk to patients or trainees.

To date, simulation in surgery has been widely used for procedural skills but in only a limited capacity for non-technical skills and scenarios outside of the operating theatre. To address this, the development of ward simulation and its rationale have been previously discussed (chapter 3.1).

The use of simulated ward environments is not necessarily a new concept. Varying degrees of ward simulation have been previously utilised by several exploratory studies, typically involving the introduction of junior doctors or nurses to the clinical environment. Mollo and colleagues, for example, describe the running of introductory simulation courses for surgical ward rounds, in which 50 US medical students participated in simulated rounds of surgical patients. In feedback questionnaires – the study’s primary endpoint – students reported interest in the course and an increase in clinical confidence. A similar study, introducing
nursing students to clinical environments and scenarios, is described by Liaw et al, which reported similar increases in student confidence and satisfaction. Within the UK, the use of ward simulation facilities has also been described by Stirling and colleagues at the University of Dundee, Scotland.

The unmet need for ward-based simulation is demonstrated by these studies, and has been reflected by the findings of the interview study as part of this thesis (chapter 2). However, despite these reports, there has remained a dearth of empiric data on the effectiveness or fidelity of ward simulation. Analogous to the many well validated laparoscopic or benchtop models for technical skills simulation, it is imperative that any novel simulator undergoes thorough scientific assessment and evaluation of its validity. Validated assessment tools to quantify trainee performance must be developed. Without such data, the impact and validity of any findings is greatly weakened, and the implications for clinical practice unclear.

The aims of this study were to develop and assess the feasibility of a simulated ward environment. Within the simulated ward we sought to investigate differences between novices and experts in the conduct of ward rounds, utilising a novel assessment tool (construct validation). Fidelity of the simulation was assessed using a questionnaire (face validity).
3.2.2 Methods

3.2.2.1 Subjects

General surgical trainees (specialist registrars) were recruited from a single academic surgical unit. For the purposes of this study ST3 - ST4 level trainees (first or second year general surgical specialty trainees) were defined as junior trainees, with ST5 – ST8 level trainees defined as senior trainees. More junior trainees were excluded, as these do not conduct independent ward rounds in common practice.

3.2.2.2 Setting

The study was conducted using the simulated ward at the Imperial College London Surgical Innovation Centre (figure 3.3). The ward contains three standard hospital beds, one of which can be set up with additional monitoring equipment on ceiling-mounted equipment arms to act as a high dependency unit (HDU) bed. The ward is fully furnished with standard hospital equipment. Documentation in the form of medical notes, vital sign charts and drug charts appropriate to the patient scenarios was available. Simulated medical documentation was not limited to that relevant to the current admission, but included documentation of previous unrelated admissions and other paperwork such as nursing care pathway proformas, social assessments or pre-operative assessments, as would be found in actual medical notes. A ward computer was available to search for radiology and blood test results for each patient as required.
Figure 3.3 Setup of the simulated ward

In background right, high dependency bed with ceiling arm for monitoring equipment and HDU chart, patient being examined by lead clinician with intern and staff nurse present. Foreground right, notes trolley with medical notes folders. Left, regular ward patients with intravenous fluid drips attached.
3.2.2.3 Simulated scenarios

Three patient scenarios were designed in collaboration with a surgical consultant to ensure clinical accuracy and validity (figure 3.4):

- Patient 1: postoperative elderly male who had undergone an anterior resection of the rectum, demonstrating signs of sepsis secondary to a probable anastomotic leak.
- Patient 2: young female patient acutely admitted with right iliac fossa pain, with a suspected diagnosis of appendicitis.
- Patient 3: middle-aged female patient, admitted with epigastric pain and blood results suggestive of pancreatitis.

The scenarios were intended to reflect a complex case mix of patients requiring further evaluation and intervention, whilst ensuring that each case was a commonly presenting problem that trainees of all levels would be expected to be familiar with and manage appropriately. Case details were derived from real patients, modified to preserve patient anonymity, and underwent consultant review to ensure realistic and clinically appropriate content. All reflected diagnoses and management scenarios described in the UK Intercollegiate Surgical Curriculum Programme. Each simulation case was based on a real-life scenario and implemented in the simulated ward following a single pilot session to identify any missing information such as paperwork or pathology reports.

Professional medical actors, appropriate to age and gender of each simulated patient were briefed both verbally and with written actor cards before each session in regards to their scenario, answers to any potential questions and how they would be expected to react upon physical examination. Each actor had intravenous drips, central lines, epidural and urinary catheters and wound dressings applied, as appropriate. All were recruited from a pool of experienced actors with years of experience in medical patient simulation, and had taken
part in previous educational and exam sessions for medical students at Imperial College London.

Two initial pilot sessions were conducted to fine-tune the scenarios. The first was conducted by a member of the research team, to ensure completeness of the scenarios. This gave an opportunity to thoroughly assess the clinical documentation and patient details, and allowed troubleshooting of errors such as incomplete anaesthetic charts, inconsistent labeling of documents, or a lack of additional documents such as clinic letters and auxiliary paperwork unrelated to the acute scenario as would be expected in actual clinical notes. Following on from this, a second trial was conducted by a senior trainee who did not take part in the subsequent study itself. Again, minor gaps were identified such as details in the history of the patients' histories or comorbidities. These were modified and completed prior to the actual trial to ensure an immersive and high fidelity environment.
**Patient 1**

78 year old male, day 2 post anterior resection

**Background:** Difficult operation, laparoscopy converted to open. Patient complaining of increasing abdominal pain and feels unwell. Diabetic, hypertensive.

**Clinical:** Increased inotropic requirements, pyrexial, tender and guarded abdomen. Reduced urine output with dark urine in urine bag.

**Investigations:** Raised inflammatory markers.

**Patient 2**

28 year old female, acute admission with lower right iliac fossa pain

**Background:** 36 hour history of right iliac fossa pain with nausea and vomiting, which the patient believes may be due to food poisoning. History of epilepsy, well controlled on regular medication.

**Clinical:** Localised peritonism and rebound tenderness at McBurney’s point.

**Investigations:** Raised inflammatory markers, pregnancy test result unclear, radiology outstanding.

**Patient 3**

58 year old female, acute admission with epigastric pain

**Background:** Sudden onset of severe epigastric pain 12 hours ago, nausea and vomiting. History of reflux, hypercholesterolaemia.

**Clinical:** Very tender with guarding in epigastrium.

**Investigations:** Blood tests not transcribed to patient notes yet, but available on computer if checked. Raised amylase and liver function tests, modified Glasgow pancreatitis score = 3 based on available results. Chest and abdominal radiographs normal.
3.2.2.4 Study protocol

Each subject was asked to conduct a ward round of three patients. A member of the research team, acting as the on-call night registrar conducting a weekend morning handover, conducted a pre-round standardised handover. This situation was intended to represent a typical situation in which the subjects would have a number of patients handed over to review, with whom they would not be familiar. A verbal description of patients and a written handover sheet was given to the subject, as would be the case in standard clinical practice.

The subject was free to conduct the ward round in the manner of their personal practice and was accompanied by a ward nurse and a house officer, both played by simulation confederates, who had been briefed with pre-scripted clinical information to provide about patients if asked. The ward nurse was additionally given a written nursing handover sheet to draw upon, containing basic information about all three patients.

The subject was unobtrusively observed and assessed by a member of the research team who was not part of the simulation. Subjects were unaware of the intended purpose of the study or the nature of assessments being conducted. Additionally, all ward rounds were recorded using stationary digital video cameras. To counter rater bias, 40% of observations were assessed by a second rater blinded to the level of training of the subject.
Figure 3.5 Study protocol overview, with subject flow and assessment methodologies detailed.
3.2.2.5 Development of the Surgical Ward care Assessment Tool

The Surgical Ward care Assessment Tool (SWAT) was developed to assess ward round performance. SWAT is a checklist-based assessment tool, whose structure was based in part on a previously validated checklist for clinical handovers, the Postoperative Handover Assessment Tool (PoHAT). PoHAT is a checklist-based assessment tool for the assessment of surgical handovers, developed by Nagpal et al. Selection of the items included in PoHAT, such as whether or not patient name, age, and medical history are described in the handover, were identified through a combination of task analysis, interviews, and literature review. The development of SWAT proceeded along a similar framework, taking into account the necessary elements of assessment tool development previously described (chapter 1.5.3).

The selection of processes to be incorporated into the tool was based upon a task analysis of the ward round, pertaining to management of the surgical patient. This considered all clinical information available to the clinician, categorised into ten items such as patient history, drug charts, radiology, or physical examination (table 3.2). Seven generic management processes applicable across all cases were identified. The processes selected for inclusion in the final iteration of the SWAT tool (appendix 2), were those deemed most relevant to generic surgical patient assessment, and which corroborated with the quality markers identified through majority consensus in the interview study presented in chapter 2. Each was included in the final checklist, to be scored as either done or not done, with elements of patient assessment (SWAT-A) and management (SWAT-M) separated into two subsections. A combined final score for each, as well as overall score, is represented by the % completion rate of each checklist.
Table 3.2 Checklist of generic assessment and management processes for the surgical ward round

<table>
<thead>
<tr>
<th>Assessment: Assess and acknowledge available sources of clinical information</th>
<th>Management: Consider following aspects of surgical care</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient</td>
<td>Analgesia</td>
</tr>
<tr>
<td>Nursing staff</td>
<td></td>
</tr>
<tr>
<td>Medical records</td>
<td>IV fluids</td>
</tr>
<tr>
<td>Vital signs chart</td>
<td></td>
</tr>
<tr>
<td>Prescription chart</td>
<td>Further blood tests</td>
</tr>
<tr>
<td>Physical examination: chest</td>
<td></td>
</tr>
<tr>
<td>Physical examination: abdomen</td>
<td>Blood cultures if pyrexial</td>
</tr>
<tr>
<td>Mechanical or pharmacological VTEP</td>
<td></td>
</tr>
<tr>
<td>Radiology</td>
<td></td>
</tr>
<tr>
<td>Blood tests</td>
<td></td>
</tr>
</tbody>
</table>

VTEP: Venous thromboembolism prophylaxis, IV: intravenous
3.2.2.6 Development of the W-NOTECHS tool

Non-technical skills such as teamwork, leadership and communication were assessed using an adapted NON TECHnical Skills (NOTECHS) scale. The original NOTECHS, and subsequently developed OTAS (Observational Teamwork Assessment in Surgery) scales, were developed for use within the operating room, based on concepts borrowed from the aviation industry.264,265 Through a series of expert interview and literature review, five categories of non-technical skills were identified – comprising situation awareness, decision making, task management, leadership, and communication / teamwork – with good and bad exemplar behaviours defined for each on a Likert scale ranging from 1 to 5 (total score 5-25). As such, NOTECHS serves to assess team non-technical interactions in the operating room, and has been since more thoroughly validated in further studies.227,290-292

Since its initial development, NOTECHS has since been further adapted for use with clinical trauma teams: T-NOTECHS.293 Though this adaptation involves only minor changes to the context or wording of exemplar behaviours, these changes are necessary to maximise the applicability and validity of the rating scale for the context in which it is applied. Based on these methods as described by Steinemann et al, the T-NOTECHS scale was modified and contextualised for use on the surgical ward round (W-NOTECHS). Thus, some descriptive behaviours were modified, such as for the Assessment and Decision making domain, where “primary and secondary surveys complete” was changed to become “full patient assessment,” for example, whereas others were not, with the Communication and Interaction domain largely unchanged: “Communications clear and audible to all / All key findings verbalized to all.” The resulting score is presented in appendix 3.

3.2.2.7 Additional assessment tools
In addition to using the SWAT and W-NOTECHS, the time taken for the subject to review each patient and to conduct the complete ward round was documented. Adverse events were identified from post-hoc analysis of video recordings.

Upon completion of the ward round, each subject filled in a fidelity questionnaire, rating a series of statements on a Likert scale of 1 to 5, with additional space for free text comments.

3.2.2.8 Statistical analysis

As this study represents a first pilot study into the assessment of surgical ward rounds, no previous data on quantification of performance, or differences in performance between trainees of different grades, was available. An estimation of the required sample size was therefore not possible, or performed. Rather, data from this study might inform such calculations for other studies in future.

Data was analysed using Microsoft Excel 2010 (Microsoft Corp, Redmond, VA) and IBM SPSS Statistics 19 (IBM Corp, Armonk, NY). Appropriate non-parametric tests were used, with Mann-Whitney U used to compare WR times, SWAT and W-NOTECHS scores between junior and senior trainee groups. The Spearman rank test was used to test for correlation between time, SWAT, and W-NOTECHS scores. Cronbach’s alpha, a statistical measure of inter-rater agreement, wherein values of 0.8 or greater indicate good agreement, was used to measure reliability of the assessments.

3.2.3 Results
Twenty subjects were recruited, ten senior and ten junior trainees. Two subjects were unable to attend the scheduled study sessions, which could not be rearranged due to practical limitations on the employment of patient actors for the sessions. As a result, a total of nine seniors and nine juniors completed the study over the course of five sessions, spread over two weeks with an equal ratio of male-to-female trainees (table 3.3).
Table 3.3  Subject demographics

<table>
<thead>
<tr>
<th></th>
<th>Junior Trainees</th>
<th>Senior Trainees</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Male : female ratio</td>
<td>7 : 2</td>
<td>7 : 2</td>
</tr>
<tr>
<td>Level of training: n</td>
<td>ST3: 8</td>
<td>ST5: 4</td>
</tr>
<tr>
<td></td>
<td>ST4: 1</td>
<td>ST6: 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ST7: 1</td>
</tr>
</tbody>
</table>
3.2.3.1 Time taken

All results are reported as median (interquartile range, IQR), unless otherwise noted. The time taken for each ward round was 36 (14) mins vs. 36 (14) mins for the senior and junior groups, respectively. There was no statistically significant difference between groups (Mann-Whitney U test, $p = 0.19$), for both overall time taken and time taken to see each individual simulated patient: Patient 1: 12 (5) vs. 10 (3) mins ($p = 0.91$), patient 2: 10 (5) vs. 8 (3) mins ($p = 0.30$), patient 3: 16 (7) vs. 12 (7) mins ($p = 0.38$).

3.2.3.2 SWAT scores

Senior trainees performed a significantly greater number of assessment processes (SWAT-A) compared to junior trainees. Of the assessment processes observed for each patient, seniors performed 70 (29)% vs. 60 (22)%, Mann-Whitney U test, $p = 0.016$ (figure 3.6). Considering specific individual processes, the greatest difference was observed in the review of medical records, 88% vs. 59% (Chi square test, $p = 0.013$) and checking for appropriate mechanical and/or pharmaceutical venous thromboembolism prophylaxis, 40% vs. 11% ($p < 0.001$).

Senior trainees also performed a greater number of management processes (SWAT-M), 71 (43)% vs. 57 (37)%, though this was borderline statistically insignificant, $p = 0.053$ (figure 3.7). For individual tasks, seniors ordered the appropriate further investigations more frequently, 67% vs 33% ($p = 0.007$).
Figure 3.6 Percentage of assessed patient assessment tasks completed for individual patient scenarios and cumulative performance for junior trainee (dark grey) and senior trainee (light grey) groups. Boxplot with outliers noted.
Figure 3.7 Percentage of assessed patient management tasks completed for individual patient scenarios and cumulative performance for junior trainee (dark grey) and senior trainee (light grey) groups. Boxplot with outliers noted.
Considering processes for each patient scenario individually, for patient 1 senior trainees consulted nursing staff more often, 89% vs 44% (p = 0.046) and checked VTEP more frequently, 44% vs 0% (p = 0.023), though there was no statistically significant difference in overall number of assessment, 89 (22)% vs. 78(22)%; p = 0.085, or management processes, 71 (21)% vs. 57 (29)%; p = 0.091.

For patient 2, there was no significant difference for either overall completion of assessment (60 (5)% vs. 60 (15)%; p = 0.31) or management (50 (33)% vs. 50 (42）%; p = 0.21) tasks, or for specific tasks.

For patient 3, VTEP was checked more frequently by senior trainees (56% vs. 0%, p = 0.009). There was no significant difference in overall assessment (80 (20)% vs. 70 (15)%; p = 0.067). A significantly greater number of senior trainees ordered appropriate further imaging (in this case, ultrasound scan), 100% vs. 56% (p = 0.023). Overall, a greater number of management tasks were completed by senior trainees (100 (10)% vs. 80 (10)%; p = 0.020).

### 3.2.3.3 W-NOTECHS

To ensure rating reliability, seven subjects (7/18, 39% of total sample) were scored by a second observer blinded to the level of training of the subject. Inter-rater reliability was high, with Cronbach’s alpha = 0.941.

Total W-NOTECHS scores (out of a maximum of 25) were significantly different between the two groups, with senior trainees scored 22 (2.8) vs. junior trainees 16.5 (4.5); p = 0.017. Individual W-NOTECHS domain scores were significantly different for “leadership,” 5 (0.3) vs. 4 (0.8), p = 0.03, and “communication and interaction,” 5 (1.0) vs. 3 (2.0), p = 0.014 (figure 3.8).
Fig 3.8 W-NOTECHS scores for junior trainees (grey) and senior trainees (black).
Statistically significant differences for leadership (p = 0.030) and communications (p = 0.014) as well as cumulative overall score (p = 0.017, not shown).
3.2.3.4 Potential adverse events

Potential adverse events were identified through post-hoc video analysis. Treatment delays, diagnostic delays, and harm events were considered separately.

Seven events likely to cause delay to treatment were identified. One senior trainee ordered a CT scan for patient 1, suspecting an anastomotic leak, but explicitly allowed the patient to eat and drink as tolerated, which would have caused a delay to theatre had a leak been confirmed. For patient 2, four juniors and two senior did not specifically specify that the patient was to be nil by mouth despite planning to take them to theatre for appendicectomy.

One event likely to result in delayed diagnosis was identified, with a junior trainee ordering a chest x-ray and urine dip test for patient 1, suspecting sepsis, but not ordering any further imaging to consider an intra-abdominal cause of sepsis.

For patient 2, a number of events occurred in which case the patient was put at avoidable risk of an adverse event. Two junior trainees and one senior trainee did not confirm pregnancy test status, as would be required for a woman of childbearing age both with abdominal pain, and prior to being considered for a general anaesthetic. Furthermore, patient scenario 2 was presented as a known epileptic on antiepileptic therapy, who was unable to take anything orally due to nausea and vomiting. However, only a single junior trainee (11%) considered prescribing a parenteral antiepileptic instead, whereas the majority (56%) of senior trainees did so.

Altogether, 23 adverse events were identified, with senior trainees responsible for 8 vs. junior trainees responsible for 15 events (p < 0.001).

3.2.3.5 Fidelity questionnaire
Responses to the fidelity questionnaire are detailed in table 3.4. 100% of subjects responded positively (positive response defined as Likert score 4 or 5, on 1-5 Likert scale) in regards to the fidelity of the ward, patients, and clinical scenarios.

Additional free-text space was allocated for comments on the most and least realistic aspects of the ward. 43 comments praised the realism of the simulation, in particular the immersiveness of the simulation, specifically due to the presence of the full ward, realistic documentation, quality of the actors and convincing clinical signs. An example comment, given by a senior (ST7) trainee was

“I was actually surprised at how immersive it was as an exercise. Very impressive.”

There were twelve negative comments. Subjects commented on the relative background silence and calm of the ward, which would be dissimilar to a real ward round. They also commented on the fact that a full respiratory exam could not be conducted. Within the study, if a subject indicated that they wished to auscultate the chest, the house officer (a simulation confederate) offered to examine the chest instead and reported findings as scripted for the scenario. Where this was not possible, and the subject themselves auscultated the patient, they were told of the intended findings by the observing member of the research team. Two subjects commented on a lack of a live vital signs monitor for the HDU bed. Only two commented on the fact that they were being watched as being an unrealistic aspect of the simulation.
Table 3.4 Fidelity questionnaire responses (Likert scale, range 1-5)

<table>
<thead>
<tr>
<th>Questionnaire statement</th>
<th>Mean score (range)</th>
<th>% positive response*</th>
</tr>
</thead>
<tbody>
<tr>
<td>The quality of the ward round can affect patient care outcomes</td>
<td>4.9 (4 - 5)</td>
<td>100</td>
</tr>
<tr>
<td>The simulated ward is realistic to work in</td>
<td>4.4 (4 - 5)</td>
<td>100</td>
</tr>
<tr>
<td>The simulated patients are realistic to examine</td>
<td>4.5 (4 - 5)</td>
<td>100</td>
</tr>
<tr>
<td>The case scenarios and clinical information provided are realistic</td>
<td>4.4 (4 - 5)</td>
<td>100</td>
</tr>
<tr>
<td>The handover information was appropriate</td>
<td>3.3 (1 - 5)</td>
<td>56</td>
</tr>
<tr>
<td>The handover information was realistic</td>
<td>3.9 (3 - 5)</td>
<td>78</td>
</tr>
<tr>
<td>The simulated ward would be useful to assess conduct of clinicians during ward rounds</td>
<td>4.2 (3 - 5)</td>
<td>78</td>
</tr>
<tr>
<td>It is important to be able to assess clinicians conducting ward rounds</td>
<td>4.5 (3 - 5)</td>
<td>89</td>
</tr>
<tr>
<td>Clinicians should train in the simulated ward before conducting unsupervised rounds</td>
<td>3.3 (1 - 5)</td>
<td>28</td>
</tr>
</tbody>
</table>

* Positive response defined as score of 4 or higher on a Likert scale of 1-5
3.2.4 Discussion

This study describes the successful implementation of a simulated ward environment, and presents the SWAT score, a novel tool for use in the assessment of performance of surgical trainees in the conduct of ward rounds. Perceived realism (fidelity) of the simulation was high with 100% of subjects responding positively to this part of the questionnaire. In establishing differences between senior and junior trainees utilising the assessment methodology described, this study has demonstrated construct validity of SWAT.

Senior trainees were more thorough in their assessment of the patient, with a greater number of assessment tasks completed. These differences tended to be greatest for sources of information which did not necessarily pertain to the patient’s acute surgical condition, such as thromboembolism prophylaxis. Though this was not formally assessed by this study, it is possible that this reflects a greater cognitive load and lesser degree of automaticity in junior trainees with reference to the WR, which might cause them to focus on acute pathology at the expense of other aspects of care. Though all trainees came to the correct primary diagnosis after reviewing each patient, juniors also more frequently neglected to note secondary errors such as drug prescription errors, or issues pertaining to relevant comorbidities (such as management of epilepsy in patient 2). Such findings suggest that future interventions to improve WR performance must, among other things, place emphasis upon a “holistic” and structured approach to the patient as a whole.

In addition to a more thorough patient assessment, senior trainees also performed a greater number of management processes compared to juniors. Though this was statistically insignificant (p = 0.053), it is possible that this was due to a lack of statistical power, as one would expect more experienced trainees to be more capable in their management of patients compared to inexperienced trainees. The fact that a significantly larger number of
potential adverse events, causing delay or placing patients at risk, were committed by junior trainees supports this hypothesis.

Though senior trainees performed a greater number of both assessment and management tasks, the time taken to see patients was not different between groups. Seniors were therefore more efficient and structured in their approach to the patient, as partially reflected in the difference in W-NOTECHS scores.

This study demonstrates the value of simulation for the assessment of clinicians in the ward setting. It cannot address all aspects of postoperative care, such as poor handover practice and communication between teams. However, this study effectively demonstrates how variability in performance between clinicians, in this case due to the greater experience of senior trainees compared to junior trainees, may account for a significant proportion of variability in post-operative care outcomes (i.e. failure to rescue).

Simulated clinical environments may be utilised as both assessment and training tools. Moorthy et al have previously reported the use of simulating operating theatres to improve surgical performance in the management of operative crises. The simulated ward provides an analogous environment in which the management of post-operative events may be considered. By repeating past adverse events in a simulated environment, for example, errors may be safely addressed and learned from to improve future patient care. The results of this study establish the simulated ward as a valid option to ensure trainees have the skill to conduct a thorough patient assessment with a systematic, structured approach to the ward round.

The Royal College of Surgeons Membership Exam syllabus and American College of Surgeons Surgical Skills Curriculum for Residents specify modules on post-operative care, communication and management skills – the mandate to ensure not only the appropriate training and assessment of future surgeons’ operative skill, but also their skills of ward based
patient management, is clear. This approach was reflected in the responses from study participants, with 100% recognising the importance of ward rounds in the clinical management of patients and 89% agreeing that clinicians should be formally assessed in their performance of ward rounds. With the advent of annual revalidation procedures in the UK, the validated objective assessment of key clinical skills gains further importance still. The simulated ward provides an immersive environment where everyday clinician practice in form of ward rounds may be assessed; this may complement existing exam modalities such as the Patient Assessment and Management Examination (PAME), an examination format developed at the University of Toronto by MacRae and colleagues, which assesses trainee performance, but is limited to individual case scenarios in a serialised exam setting.

Though 89% of subjects agreed that the ability to assess clinicians’ ability to conduct effective ward rounds was important and 100% agreed the simulation was realistic, only 28% agreed that simulated ward round training should be undertaken before conducting rounds on real patients. This disconnect was not explored in detail, however it may reflect the disinclination to engage with the assessment process which has been reported in other areas with clinicians undergoing re-validation. Curnock et al reported the results of a semi-structured interview study in which views of doctors resistant to participation in a peer appraisal scheme was explored. Central to this was a perceived lack of “value” in the system, involving subjective peer reporting, with a lack of objective and validated measures.

Having demonstrated the validity of SWAT and W-NOTECHS in this study, it is possible that such barriers to WR assessment may be overcome with education and information, as such mechanisms of assessment remain necessary if surgical training is to advance with proficiency-based curricula.

This study has some limitations to be taken into account. The relatively small sample size represents a limitation and may have affected our ability to draw statistically significant
conclusions from some of the results, such as the lack of significantly different SWAT-M scores for patient management. SWAT also does not assess the contributions of other clinical staff (nurses and other allied health professionals). However, it is the lead clinician who is responsible for most, if not all, patient assessment and management decisions; aspects of team performance as whole, furthermore, are captured by W-NOTECHS.

During the WR simulation, subjects were not set any time limitation in their assessment of the patient. However, as all patients presented in the simulation were acutely unwell, their assessment would be presumed to take priority over other clinical duties in real practice, such that placing an explicit limit on time would have been unrealistic. Despite this, had a time stressor been introduced, it is possible that error rates would have been higher still. Future studies involving the simulated ward will seek to further improve fidelity. In response to the comments received from this study’s fidelity questionnaire, the simulated ward at St Mary’s Hospital has been further improved through incorporation of background noises and distractors as would be present on a real ward round, simulated live patient monitoring and enabling clinicians to perform a full respiratory examination through the use of programmable digital stethoscopes capable of reproducing audible chest pathology (Ventriloscope, Lecat’s Simply Sim, Ohio, USA).

3.2.5 Conclusion

Simulation has long established its place within the surgical curriculum for technical skills, with clear advantages for both trainees and patients. With attention shifting from the surgeon’s performance within the operating theatre to their conduct without and from technical to non-technical skills, the simulated ward provides the ideal environment to facilitate training and assessment of essential ward-based skills. Providing an immersive,
high-fidelity environment within a dedicated educational context, the simulated ward represents a valuable tool to enhance clinical performance and post-operative care for surgeons and patients, respectively. Having demonstrated feasibility and validity of the simulated ward, the next step is to demonstrate its clinical relevance. Future studies will also seek to further explore the potential of ward simulation as a training and assessment tool.
Chapter 4

Surgical ward round quality and impact on variable patient outcomes
4.1 Introduction

Surgical culture and training has changed dramatically in recent decades, with the modern surgical profession espousing a culture of safety, quality improvement, and outcomes measurement. In recognition of this, measures such as national surgical databases and broadly implemented interventions such as the WHO surgical safety checklist have contributed to quality control, error reduction and standardisation in operative patient care. These advances to improve care in the operating room, unfortunately, have not been matched by changes to care outside of it.

This thesis has previously discussed the body of literature surrounding studies of adverse events, which suggest that over half of all events take place on the ward rather than in the operating room, as well as the concept of “failure to rescue,” evidence which increasingly now suggests that short-term surgical outcomes are determined more by the quality of post-operative care than the success of the operation itself. This would appear to imply that medical errors and variability in the processes of giving inpatient, peri-operative care, are at the root of poor outcomes. The underlying mechanisms which determine the successful or unsuccessful rescue of patients, however, are unclear, preventing targeted intervention to reduce failure to rescue rates.

As elucidated in previous chapters of this thesis, the primary care process in determining ward-based care, and instrumental in the assessment and management of the patient on a routine basis, is the surgical ward round (WR). Leading a ward round, accompanied by clinicians, nurses, and allied health professionals requires not only clinical knowledge, but also skill in communication, decision making, leadership and management. During the WR, the clinician’s task is to synthesise a multitude of information sources, ranging from verbal accounts to charts and physical examinations, to assess what progress or setbacks in care the patient may have experienced since the last clinical evaluation. Failing to account for
any of these information sources potentially constitutes an error of omission, risking oversight of valuable clinical information which may affect patient care and result in adverse events as a result.

The varied nature of WR performance which exists, even within the standardised environment of the simulated ward, has been described in the previous chapter of this thesis. The recent RCP / RCN statement on ward rounds has served to further highlight this perceived lack of standardisation in care, which may result in a high potential for error and harm. However, no empirical data currently exists as to the true variability in clinical conduct of WRs, or its potential influence (if any) on patient outcomes.

The previously described development and validation of SWAT and W-NOTECHS scores for the measurement of clinician WR performance represent, for the first time, the opportunity for objective WR quality measurement. The availability of such tools should allow the evidence-based approach required to apply science to what has too long been merely an art. By capturing clinical data through objective measurement the true effect of WRs on surgical outcomes can be clarified, and attempts to improve practice initiated.

_The aim of this study was to observe variability of in vivo surgical WRs, to identify sources of error in care, and investigate the possible relationship between WR quality and clinical outcome._
4.2 Methods

4.2.1 Setting

This study was conducted at St Mary’s Hospital, an academic tertiary referral centre in London, England, following review and approval by a regional ethics committee (reference 12/LO/0617).

The study was conducted on the surgical high dependency unit (HDU), which occupies a dedicated bay of the general surgical ward. The HDU is similar to a surgical intensive care unit, consisting of a 4-bed unit with a 1:2 nurse:patient ratio. It provides care for patients requiring close monitoring, basic inotropic and respiratory (continuous positive airway pressure) support or invasive monitoring, but does not include provision for renal replacement therapy or active ventilation.

Patients on the HDU are cared for and have WRs conducted by the same clinical team as on the ward, under the care of a named consultant member of the general surgical team. At the time of this study, the general surgical department consisted of four colorectal (benign and neoplastic resections) and five upper gastrointestinal surgeons (including bariatric, benign and neoplastic oesophago-gastric resections).

4.2.2 Patients

Only patients admitted to the HDU either directly from the emergency department or immediately post-operatively (either planned or unplanned) were included in the study. In this manner, a patient cohort which was acutely unwell, or had undergone major, high-risk, surgery, with an anticipated high risk of complications, was selected. Existing ward inpatients that were moved to the HDU following complications or other clinical
deterioration were excluded, as part of the purpose of the study was to capture developing morbidity and these patients would have, by definition, already developed complications.

4.2.3 Study protocol

Consecutive HDU ward rounds were observed following informed consent from patients. WRs were rated on their quality using a derivative of the SWAT score, a checklist-based measure of thoroughness of patient assessment and management. All available sources of clinical information (SCI) were take into account. Clinicians’ actions were directly observed to determine whether clinical information items were considered or not, with a quality score awarded based on the percentage of available SCI taken into account. Whether or not the WR included examination of the prescription chart, physical examination, or checking of drain bag contents were readily observable, and completion recorded regardless of whether or not a verbal or written comment to this effect was given. If tasks were delegated to other team members by the lead clinician, this was considered as having been completed.

Non-technical performance was assessed using the W-NOTECHS scale, which assesses team, communication and leadership skills performance across five behavioural domains, with exemplar behaviours modified for the ward round setting. Each domain is rated on a Likert scale of 1-5, resulting in a final score of 5-25. The development and validation of both SWAT and W-NOTECHS were previously discussed in chapter 3.

Routine WRs consist of a lead clinician (surgical consultant or registrar), accompanied by other clinician staff (registrars, senior house officers, and house officers), as well as nursing staff or other health professionals such as pharmacists. The task of determining patient assessment and management is the responsibility of the lead clinician, who may undertake these himself or herself or delegate to other members of the team.
4.2.4 Patient outcomes

Patient demographics and clinical information were retrieved through casenote review of patient records on discharge; any complications or adverse events were recorded. Complications were recorded according to the clinical documentation. These were also crosschecked with available clinical information to ensure their appropriateness, in accordance with established guidelines where appropriate (i.e. presence of radiographic infiltrates combined with two of pyrexia, leucocytosis or leucopenia, and purulent secretions for the diagnosis of pneumonia). Adverse events and errors were classified for severity, in line with previously published methodology and clinician consensus, as incurring a risk of harm, possible harm, or definite harm.2, 3

Defining preventable complications was based on Agency for Healthcare and Research Quality Patient Safety Indicator definitions for preventable morbidity and included: post-operative pneumonia, urinary tract infections, wound infections, acute renal failure, and central venous catheter-related sepsis.305

Where patients developed post-operative morbidity, WR data was analysed to determine whether the adverse event in question had been preceded by a related medical error of either commission or omission – had the affected organ system been appropriately attended to, or was there other evidence to suggest earlier deterioration which could have led to an earlier diagnosis or alternative management, for example?

4.2.5 Statistical analysis

Inter-clinician variability of WR quality was analysed with ANOVA. To analyse the effect of WR quality on outcomes, patients were divided into two groups based on SWAT performance scores above or below the median value. For these groups, incidence of all complications, adverse events, and preventable complications was compared using Chi-
square tests, and effect size of WR quality calculated. Statistical analysis of data was performed in IBM SPSS Statistics 20 (IBM Corp, Armonk, NY, USA). A p-value of less than 0.05 was considered statistically significant.

4.3 Results

4.3.1 Patient demographics

A total of 50 patients admitted to the general surgical HDU were followed during the period of June – October 2012, totaling approximately 55 hours of clinical observation over the course of 37 observation days.

Planned HDU admissions following major elective surgery constituted the majority (54%) of patients, with the remainder representing emergency admissions, either following emergency surgery, or unstable surgical patients requiring higher-level observation and care who did not undergo an operative procedure (see table 4.1).
<table>
<thead>
<tr>
<th>Table 4.1 Patient demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total number of patients, n</strong></td>
</tr>
<tr>
<td>Male : female ratio</td>
</tr>
<tr>
<td>Age (years)</td>
</tr>
<tr>
<td>ASA grade</td>
</tr>
<tr>
<td><strong>Type of admission</strong></td>
</tr>
<tr>
<td>Emergency</td>
</tr>
<tr>
<td>Elective</td>
</tr>
<tr>
<td><strong>Operative procedure (n)</strong></td>
</tr>
<tr>
<td>Colectomy</td>
</tr>
<tr>
<td>Gastro-oesophageal resection</td>
</tr>
<tr>
<td>Laparotomy without resection</td>
</tr>
<tr>
<td>Laparoscopy without resection</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>No operative procedure</td>
</tr>
<tr>
<td><strong>HDU length of stay (days)</strong></td>
</tr>
<tr>
<td><strong>Total length of stay (days)</strong></td>
</tr>
<tr>
<td>**Complications (%) **</td>
</tr>
<tr>
<td>Pneumonia</td>
</tr>
<tr>
<td>Ileus</td>
</tr>
<tr>
<td>Sepsis</td>
</tr>
<tr>
<td>Urinary tract infection</td>
</tr>
<tr>
<td>Wound infection</td>
</tr>
<tr>
<td>Anastomotic leak</td>
</tr>
<tr>
<td>Acute renal failure</td>
</tr>
<tr>
<td>Other</td>
</tr>
</tbody>
</table>

Results reported as mean ± standard deviation unless stated, ASA: American Society of Anesthesiologists, “other” operative procedures include stoma reversal, radiological interventions, *median(range), **total percentage greater than overall morbidity due to multiple complications for some patients
4.3.2  Ward round quality

69 patient WRs were observed, constituting 72% of all HDU WRs for the 50 patients in question. WRs were led by 15 different clinicians of varying grade, including consultants (30%), specialty registrars (64%), or senior house officers (6%). WR quality, as measured by the percentage of SCI assessed during the WR, varied significantly between clinicians, range 9-91%, mean 55% ± 17% SD, ANOVA p = 0.025. Similar differences in non-technical performance were observed, with significant variation in W-NOTECHS scores, range 16-25, mean 19.7 ± 0.35, ANOVA p < 0.001. The assessment of individual SCI varied, with verbal assessment of the patient and checking of vital signs the most frequently completed tasks (96%), and examination of the respiratory system least frequent (14%) (see table 4.2).
Table 4.2 Assessed sources of clinical information during surgical ward rounds

<table>
<thead>
<tr>
<th>Assessed item</th>
<th>n</th>
<th>% assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient verbally assessed</td>
<td>50</td>
<td>96%</td>
</tr>
<tr>
<td>Vital signs</td>
<td>50</td>
<td>96%</td>
</tr>
<tr>
<td>Nurse verbal handover</td>
<td>50</td>
<td>82%</td>
</tr>
<tr>
<td>Drains</td>
<td>26</td>
<td>76%</td>
</tr>
<tr>
<td>Wound</td>
<td>41</td>
<td>72%</td>
</tr>
<tr>
<td>Abdomen examined</td>
<td>50</td>
<td>68%</td>
</tr>
<tr>
<td>Stoma</td>
<td>14</td>
<td>54%</td>
</tr>
<tr>
<td>Lines and catheters</td>
<td>39</td>
<td>32%</td>
</tr>
<tr>
<td>Medical notes</td>
<td>50</td>
<td>30%</td>
</tr>
<tr>
<td>Prescription chart</td>
<td>50</td>
<td>30%</td>
</tr>
<tr>
<td>Thromboprophylaxis</td>
<td>48</td>
<td>22%</td>
</tr>
<tr>
<td>Chest examined</td>
<td>50</td>
<td>14%</td>
</tr>
</tbody>
</table>
4.3.3 Complications and errors

The observed morbidity rate was 60% (30/50), with 74% (35/46) of all complications occurring on the HDU. The other 26% of complications occurred after stepping down to standard ward care. Retrospective analysis determined that for 41% (19/46) of all complications, development of morbidity was preceded by suboptimal patient assessment. In these cases, failure to consider relevant clinical information was felt likely to have resulted in delayed diagnosis and treatment, or failure to prevent the complication. The most common example was the development of post-operative pneumonia in patients where the respiratory system was not previously attended to during the WR. Others included the delayed diagnosis of a wound infection where allied health professionals had documented early signs of cellulitis, with failure to review the medical notes during the WR, or delayed catheter removal and subsequent urinary tract infection (see table 4.3). Additionally, retrospective review revealed 20 further risk events which were without clinical sequelae, these related mostly to prescribing errors (table 4.3).
<table>
<thead>
<tr>
<th>Type of event</th>
<th>Clinical example</th>
<th>SCI overlooked</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Preventable complication (possible harm)</strong></td>
<td>Chest not attended to, subsequent pneumonia</td>
<td>Chest</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Failure to remove urinary catheter as per protocol, subsequent UTI</td>
<td>Lines and catheters</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Suggested treatment documented in notes but not acted on, chest not attended to, subsequent pneumonia</td>
<td>Chest</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Wound discharge documented in notes by AHP, wound not attended to, subsequent wound infection</td>
<td>Wound</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Failure to administer prescribed chest therapy documented in notes but not noted, subsequent pneumonia</td>
<td>Medical notes</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>CVC-related sepsis, insertion site cellulitis noted during CVC removal</td>
<td>Lines and catheters</td>
<td>1</td>
</tr>
<tr>
<td><strong>Prescribing error (risk of harm)</strong></td>
<td>Prescription error relating to dose or indication</td>
<td>Prescription chart</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Failure to restart regular medications appropriately</td>
<td>Prescription chart</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Recommended therapy documented in notes, not prescribed</td>
<td>Prescription chart</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Antiembolism stockings not worn</td>
<td>Thromboprophylaxis</td>
<td>3</td>
</tr>
</tbody>
</table>

SCI: source of clinical information, CVC: central venous catheter, AHP: allied health professional, UTI: urinary tract infection
4.3.4 Outcome analysis

Patients were divided into groups of high and low quality WRs (defined as %SCI assessed above or below the median, respectively) (figure 4.1). Those patients assessed by low quality WRs demonstrated a greater incidence of complications which were either preventable (as defined by AHRQ criteria) or whose treatment was suboptimal (delayed diagnosis or treatment based on expert review), 83% (10/12 patients) vs. 39% (7/18 patients), Chi-square p = 0.034, odds ratio 6.43 (95%CI 1.05, 39.3). There was no difference in the incidence of preventable errors (OR 1.40, 95%CI 0.36 – 5.41), or of complications in general (OR 3.06, 95%CI 0.80 – 11.73).

Patients who developed complications had a longer length of stay on HDU, 4.1 ± 1.7 days vs. 2.6 ± 1.5 days (t-test, p = 0.050). There was no significant relationship between complication or error rates and any of age, ASA score, W-NOTECHS score, or type of admission (emergency vs. elective), all p-values > 0.2.
Figure 4.1 Effect sizes of ward round quality for error incidence, all complications, and preventable complication rates.

<table>
<thead>
<tr>
<th>ID</th>
<th>ES (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preventable error</td>
<td>1.40 (0.36, 5.41)</td>
</tr>
<tr>
<td>All complications</td>
<td>3.06 (0.80, 11.73)</td>
</tr>
<tr>
<td>Preventable complications</td>
<td>6.43 (1.05, 39.30)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>High quality WR</th>
<th>1</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>Low quality WR</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
4.4 Discussion

This study documents the variability of care and incidence of error and adverse events in post-operative surgical care. It represents the first time, to our knowledge, that post-operative care and surgical ward rounds in the clinical environment have been subject to empirical examination in this manner. Existing literature is sparse and has considered limited aspects of the WR such as use of medical records or communication, but not the WR process as a whole. Whilst other important aspects, such as clinician handoff and information transfer, also factor into the quality of post-operative care, it is during the WR that patients are dependent on their clinician’s ability to assess, appropriately diagnose, and effectively manage problems or complications which may develop. Despite its importance, the results of this study illustrate the variable levels of performance in the conduct of WRs that currently exist, even in a tertiary academic centre, and a lack of standardisation in approach to patient assessment across clinicians of all experience levels.

Perhaps most importantly, this study demonstrates the potential for such variability, and resulting errors or omissions, to adversely affect patient care and outcomes. It is an unfortunate, but accepted, reality that medical error is a common occurrence. Fortunately, most errors, in the majority of cases, do not lead to actual patient harm. However the events examined here strongly suggest links between oversights in patient assessment and a potential for subsequent deteriorations in patient condition. The results presented here cannot definitively demonstrate causality; however, this would likely be impossible, certainly without much greater patient numbers, given the multifactorial nature of post-operative complications.

Despite this, many of the recorded errors and their potentially associated complications represent well-documented entities. The failure to remove unnecessary urinary catheters is a well-established risk factor for urinary tract infections (UTI), for example, and likely
contributed to development of this complication in 10% of patients. Almost one in three
patients where the catheter was delayed, contrary to local policies for catheter removal or
lacking any clear indication, developed a UTI. Furthermore, only 14% of pre-morbid ward
rounds attended to the patient’s respiratory system, though pneumonia was the most
common complication on this unit (30%). Diagnosis of pneumonia was for the most part
made once the patient developed systemic symptoms of sepsis. Auscultation and
examination of the patient’s chest may have led to an earlier diagnosis or identification of
early treatable warning signs such as atelectasis or impeded sputum clearance. Published
evidence is clear on the importance of early identification and amelioration of complications
to improve outcomes310,311 – however without thorough patient assessment in the context
of a high quality WR, this is not possible. Other examples of error included the failure to
review the medical notes, in which any events of the last 24 hours would have been
documented. Though a handover from attending nursing staff was requested in almost all
(96%) cases, overlooking the written record led to delays in diagnosis or treatment on
several occasions, when concerns had been previously documented or advice given by other
consulting specialty teams.

Statistical analysis conclusively demonstrated the greatly increased incidence of
complications which were either preventable, or could have been managed differently, in
patients who were subject of poorer quality WRs. It is also notable that the most common
preventable errors (prescribing errors) and complications (pneumonia) were related to the
least commonly assessed aspects of the patient (prescription charts, respiratory system),
illustrating the importance of thorough patient assessment.

Though modern hospital care includes a number of other health professionals involved in
the prevention and management of many of the observed complications, ultimate
responsibility lies with the primary clinician. Provision of chest physiotherapy may be useful
in patients with respiratory problems, for example, but can take place only if prescribed.
Communication and team working skills are also needed to ensure appropriate understanding within the team and agreement of tasks with nursing staff for such orders as catheter removals or dressing changes.

This study has several weaknesses to consider. This was an observational study of a single unit within a single centre, representing a potential source of sample bias. However, it must be noted that the site in question is a tertiary academic and accredited trauma centre, recently ranked fourth out of 123 hospital trusts across England in terms of standardised and adjusted hospital mortality ratios. As such, it is unlikely that the results of this study represent poor performance on the part of the unit in question - rather it is probable that other centres would exhibit higher error rates still.

Though not all hospitals provide care in a surgical HDU, the HDU environment was chosen for its care complexity and therefore greater potential for error. The same clinical team as on the ward looks after HDU patients, and though patients potentially benefit from greater nursing ratios, the increased complexity of unwell HDU patients overwhelms this as a potential confounder. This is reflected in the fact that 74% of all complications experienced by the observed patient cohort took place on the HDU, despite a median HDU length of stay of only 3 days. Furthermore, should higher levels of nursing care prove a significant confounder after all, this would indicate greater levels of error still on the ward, and would place further importance on the generalisability of these findings still. Illustrating this point, Jones et al studied demographically equal patients who underwent postoperative care with or without access to an HDU, with the non-HDU group developing a significantly greater number of complications.

Observation of the WR provided only a “snapshot” of care, by a single observer, without observing any other care interactions throughout the day. Though the WR is undoubtedly the primary interaction between doctor and patient, other assessments will have taken
place. However, the outcome of most, if not all, of these were captured via subsequent review of medical records and were factored into a number of identified errors (e.g. where a documented assessment of patient deterioration by a nurse or junior doctor was not acted on appropriately). Though blinded assessment was not possible, assessments were performed by a single observer with substantial experience in skills assessment, validated in multiple previous experiments which included analysis of inter-observer reliability.

Finally, as demonstrating causal links is not possible, given the multifactorial nature of postoperative morbidity, this study instead suggests circumstantial links between events and complications. The high complication rate must be considered in the context of the fact that only HDU patients were included, which by definition were critically unwell and had undergone major or emergency surgery. In addition, a number of patients had undergone gastric or esophageal resections, which are associated with a known high risk of postoperative complications – in this context, the reported rate of complications is within previously described ranges.148,314,315

Ensuring quality WRs and thorough patient assessment requires adequate training, assessment and subsequent quality control. One obstacle has in the past been a lack of appropriate training tools. Complex patients such as those on an HDU are clinically inappropriate and ethically dubious subjects for training, and other priorities mandated for surgical trainees, such as attendance in the operating room, means that training time on the ward is often limited from the outset. Ward simulation has been preliminarily explored, though current published uses with clinicians are few.274,284 As this develops further it may present a viable alternative to in situ training.32

Whilst the large number of clinicians observed resulted in a low number of WRs per clinician in this study, and prevented meaningful analysis of clinician-specific analysis, the effect of clinician experience on WR quality has been previously explored in chapter 3.18 When
presented with an unwell patient, more senior clinicians were shown to be significantly more thorough and detailed in their examination of the patient, assessing and integrating greater amounts of clinical information, and committing less errors in their diagnosis and management of patient problems. To shorten the learning curve, particularly for junior clinicians, the development of evidence-based curricula and assessments have been successfully pursued and implemented in training for technical skills – it is now time to do the same for post-operative care.

4.5 Conclusion

In their current form, ward rounds are traditionally learned by experience and emulation, resulting in a lack of standardisation in patient assessment. This study is the first to assess the resulting variability of quality of surgical ward rounds and demonstrate the consequences borne by patients, with poor quality WRs placing patients at an up to six-fold risk of developing preventable complications. It demonstrates the pervasiveness of human error which has been demonstrated in other areas of medicine. Whilst errors are unavoidable, it is important to act to minimise their effects. To improve care, common errors and omissions identified in this study must be addressed by further research to develop targeted interventions and training methods to improve care.
Chapter 5

Development and implementation of a ward round training curriculum
5.1 Development of an evidence-based curriculum for training of ward-based surgical care

5.1.1 Introduction

Thus far, this thesis has described the validation of a tool for the assessment of WR quality, and has demonstrated the current variability of ward-based care in clinical practice. It has established the greatly increased risk of preventable complications which results with poor WRs and substandard care. With post-operative morbidity rates as high as 30-50% for some complex procedures, the importance of ensuring trainees’ ability to conduct thorough patient assessment and initiate appropriate management during the surgical WR is clear.

Currently, trainees conducting WRs rely primarily on the Halstedian method of “learning by doing” for this task. This outdated method raises a number of issues. Firstly, the use of patients to teach complex case management poses potential ethical questions, when the most unwell patients also become the best learning opportunities. Furthermore, from a practical perspective, opportunities for apprenticeship-style learning have been greatly reduced with the now widespread introduction of statutory limitations on working hours, such as in North America and Europe. The gradual loss of the traditional trainer-trainee relationship has lessened opportunities for knowledge transfer in this manner, and fewer hours spent on the wards also mean trainees are potentially exposed to fewer patients overall. Finally, the “traditional” approach fails to address the complex demands of leading a WR, which extend beyond learning the basic skills of patient assessment and management, to critical non-technical skills such as teaching, leadership and multi-professional communication. The lack of trainees’ graded progression, leading to assured
proficiency, potentially exposes patients to undue risk of mismanagement with dire consequences. Clearly, alternative training methods are needed.

In order to improve the quality of WRs the formalised means for standardisation of practice, training, and assessment must be considered. Current surgical curricula have begun to recognise the importance of WRs and non-technical skills in general, such as the Intercollegiate Surgical Curriculum Programme in the UK, which singles out WRs as a curricular requirement.\textsuperscript{318} Crucially, however, it does not provide recommendations on how skills such as the conduct of surgical WRs are to be acquired. The review of published literature presented in the introduction to this thesis (chapter 1.5) reveals the lack of evidence which currently exists for this domain of surgical training. In the present context of the known variability of ward-based care, there is a clear need for the development of an evidence-based, comprehensive curriculum to address this training gap.

The purpose of a curriculum is to confer knowledge and skill, through ongoing feedback and assessment with progression toward a predetermined level of proficiency.\textsuperscript{319} Simulation has proven itself as an effective means of achieving this, and has been widely incorporated in surgical technical skills training,\textsuperscript{320} with increasing use for non-technical skills too.\textsuperscript{321} The development of simulation for ward-based skills, described in chapter 3, is a relatively recent concept and presents the advantages of a high-fidelity, controlled environment in which reproducible assessment and training scenarios can be created, without risks to patients or clinical staff.\textsuperscript{322}

\textit{To allow the implementation of appropriate, evidence-based training for WRs, the aim of this chapter is to develop a comprehensive simulation-based curriculum for WR training based on modern training methods and best published evidence.}
5.1.2 Methods

The design of a simulation-based curriculum for WR training was modeled on previously validated methods, and designed in accordance with the recommendations of a recent international expert consensus framework for the creation of simulation-based surgical curricula. Broadly, these identify a three-step process, consisting of (1) predevelopment analysis, (2) curriculum development, and (3) curriculum validation.

Predevelopment analysis should incorporate a needs and resource assessment, as well the optimisation of resource-related prerequisites to enable implementation of the curriculum. The curriculum development stage advocates deconstruction of the procedure into component tasks and individual domains of cognitive, psychomotor and team-based aspects of skill, ensuring each is individually addressed. Finally, delivery and validation of the curriculum requires selection of appropriate simulation models for training, modular skill-based tutorials, selection of validated assessment tools, and structured de-briefing.

5.1.2.1 Predevelopment analysis

The consideration of appropriate space and resources, and the development and validation of the high-fidelity simulated ward at St Mary’s Hospital have been previously discussed (chapter 3). Additional classroom space and teaching resources were already in place, as part of the surgical skills centre in which the simulated ward is located.

As part of a pre-developemental analysis, an open questionnaire was circulated to general surgical trainees who had participated in the previous ward simulation validation study described in chapter 3. As trainees, and having volunteered to participate in a ward simulation-based study, these were deemed to have demonstrated significant interest in WR education, and therefore deemed suitable subjects of a stakeholder needs analysis.
Trainees were asked to name the most important factors they felt were responsible for determining the quality of a WR, and which could be improved, thereby identifying points for improvement that might be addressed. Questionnaire responses were transcribed into an anonymised database (Microsoft Excel, Microsoft Corp, Redmond, VA, USA) and were subjected to emergent themes analysis. The resulting topics were combined with information derived from existing surgical curricula to derive cognitive, team-based and psychomotor learning objectives.

5.1.2.2 Curriculum design

A modular tutorial design was adopted to address learning goals identified in the predevelopment analysis. A review of published literature was undertaken in each case to identify suitable, validated training and assessment tools.

5.1.3 Results

5.1.3.1 Predevelopment analysis

The questionnaire was issued to 20 general surgical trainees (specialist registrars level ST3 - ST8), 18 responses were received. Emergent theme analysis of the free text answers given identified eight factors which trainees deemed important determinants of WR performance (table 5.1). These included staffing-related issues (67%), documentation (56%) and time management (56%). Saturation of themes, the point at which no further new themes were identified, was reached within the number of questionnaire responses given. Following review of available literature, published guidance and existing curricula for other domains
of surgical training, these eight factors were condensed into three themes: (1) patient assessment and management, (2) communication skills and (3) team working. Broadly mirroring the three prescribed domains of cognitive, psychomotor and team-based skill, respectively, these formed the basis for further curriculum development.

Table 5.1  Factors critical to WR quality: results of trainee questionnaire theme analysis (by percent of respondents)

<table>
<thead>
<tr>
<th>Theme</th>
<th>Example</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staffing</td>
<td>Presence of staff and fulfilment of team role</td>
<td>67%</td>
</tr>
<tr>
<td>Documentation</td>
<td>Appropriate documentation in medical notes</td>
<td>56%</td>
</tr>
<tr>
<td>Time management</td>
<td>Efficient use of time to see patient and chase results of investigations</td>
<td>56%</td>
</tr>
<tr>
<td>Communication at handover / debriefing</td>
<td>Accurate handover and communication between teams</td>
<td>50%</td>
</tr>
<tr>
<td>Systematic approach</td>
<td>Thorough patient assessment in systematic manner</td>
<td>22%</td>
</tr>
<tr>
<td>Team approach</td>
<td>Team working and good leadership</td>
<td>22%</td>
</tr>
<tr>
<td>Work load</td>
<td>Managing work load and number of patients under clinician’s care</td>
<td>11%</td>
</tr>
<tr>
<td>Patient location</td>
<td>Efficiency of WR when visiting patients on outlier wards</td>
<td>11%</td>
</tr>
</tbody>
</table>
5.1.3.2 Curriculum development

To better allow integration with existing surgical training curricula, the WR curriculum was structured as a half-day training course (figure 5.1). This consisted of:

- An assessment of knowledge and confidence
- Didactic session
- Simulated WR and performance assessment
- Debriefing and feedback

Following a literature search, peer-reviewed, validated methodology was used in the design of each module (table 5.2).
Fig 5.1 Curriculum flow chart with projected course timeframe

- Pre-test knowledge and confidence assessment: 30 min
- Didactic session: 90 min
- Simulated WR: 30-45 min
- Debriefing and feedback: 60 min
- Post-test knowledge and confidence assessment: 30 min
<table>
<thead>
<tr>
<th>Skill or module</th>
<th>Evidence-based tool used</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowledge and confidence</td>
<td>Likert-scale questionnaire based on published model(^{225})</td>
<td>Examines trainee’s previous experience, current perceived skills and attitudes towards WR</td>
</tr>
<tr>
<td>Management and assessment of patient</td>
<td>Surgical Ward care Assessment Tool (SWAT)(^{326})</td>
<td>Validated tool gives score based on thoroughness of patient assessment, management</td>
</tr>
<tr>
<td>Team-based and non-technical skills</td>
<td>W-NOTECHS(^{293, 326})</td>
<td>Validated tool, scores non-technical behavioural domains</td>
</tr>
<tr>
<td>Clinical decision making</td>
<td>Compare actual to expected management for specific patient</td>
<td>Each scenario validated with surgical experts for expected (gold standard) treatment</td>
</tr>
<tr>
<td>Debriefing and feedback</td>
<td>Objective Structured Assessment of Debriefing (OSAD)(^{223, 327})</td>
<td>Structured approach to debriefing optimising learner engagement and benefit</td>
</tr>
<tr>
<td></td>
<td>Video feedback(^{328})</td>
<td>Self and expert video review</td>
</tr>
</tbody>
</table>

WR: Ward round
5.1.3.3 Knowledge and confidence

At the start of training session, trainees are asked to respond to a series of statements on a standard Likert scale in a pre- and post-test questionnaire, based on a validated design and previously published questionnaire by Flin et al.\textsuperscript{325} The questionnaire is designed to evaluate trainee knowledge, confidence and attitudes in the independent conduct of surgical WRs (figure 5.2).
### SimWard training day questionnaire

<table>
<thead>
<tr>
<th>Candidate # ____________</th>
<th>Do you routinely lead ward rounds in your current job?</th>
<th>Y / N</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly disagree</th>
<th>Unsure</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I am confident in my ability to thoroughly <em>assess</em> surgical patients</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>I can confidently <em>diagnose</em> post-operative complications</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>I can confidently <em>manage</em> complications of surgical patients on the ward</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>I communicate well within a multidisciplinary clinical team</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>I am confident in my leadership abilities to lead and manage a multi-disciplinary team</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>I am happy to routinely lead surgical ward rounds (with a more senior clinician)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>I am happy to routinely lead surgical ward rounds independently (without senior help)</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Ward rounds are important in determining care of surgical patients</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>The quality of a ward round varies between clinicians</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Patients should always be thoroughly examined during a ward round</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Medical notes should always be reviewed during a ward round</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>The quality of the ward round can determine patient outcomes</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>
5.1.3.4 Didactic session

The didactic session focuses on generic skills for surgical WRs and published recommendations for WR structure. The session encourages trainee discussion of their own experiences in conduct and to conduct WRs, to better relate learning point to trainees’ current practice. Describing current up-to-date published evidence, the session emphasises a systematic approach to the patient with conscious consideration of non-technical and team-based skills (figure 5.3).

5.1.3.5 Simulated Ward Round

WR performance is observed in a high-fidelity simulated ward, with trainees performing a WR of three simulated patient scenarios, played by trained medical actors. These standardised scenarios, previously validated, reflect commonly presenting surgical pathologies such pancreatitis, appendicitis, or colectomy, whose appropriate management could reasonably be expected of junior surgical trainees. Each scenario is based on a real patient, with the trainee required to devise an appropriate management plan during the WR, following thorough patient assessment, appropriate decision making, and inclusion of the multi-disciplinary team consisting of an intern and staff nurse (both course confederates). Conduct of the simulated WR is recorded using unobtrusive ceiling-mounted digital video cameras (smots™, Scotia UK plc, Edinburgh, Scotland, UK), for use in assessment, debriefing, and feedback.
Figure 5.3 A systematic approach to conduct of surgical WRs

1. Pre-brief team and assign roles
2. Check available results including
   - notes
   - operation note
   - vital signs
   - drug chart
3. Examine patient
   - abdomen
   - chest
   - drains and lines
4. Clear communication to scribe of findings and plan, prescribe treatment
5. Debrief with nursing staff and patient
   - plan
   - diet
   - mobilisation
   - referrals
6. Debrief with team at end of WR: prioritise and assign tasks
5.1.3.6 Assessment Tools

The WR is assessed using the Surgical Ward care Assessment Tool (SWAT)\textsuperscript{326} for psychomotor aspects of the WR, involving patient assessment and management. For non-technical skills such as team work, communication and leadership, multiple validated scales exist: the Non-Technical Skills scale (NOTECHS), Objective Teamwork Assessment for Surgery (OTAS), and Non-Technical Skills for Surgeons (NOTSS).\textsuperscript{224} Of these, the NOTECHS scale has been adapted to multiple clinical environments\textsuperscript{293} and is the only to have been validated in a version adapted specifically for WRs (W-NOTECHS, as described in chapter 3);\textsuperscript{326} it further benefits from its simplicity and its usability by minimally trained assessors.

5.1.3.7 Debriefing feedback

Following WR assessment, both summative and formative feedback is given. Based on principles outlined by Arora and colleagues,\textsuperscript{223, 327} the trainee receives specific feedback about performance based on the assessment tools as described as well as specific clinical information detailing their management of individual cases. Additionally, video of the trainee’s WR is reviewed for self- and expert-led feedback.\textsuperscript{328} At course end, each trainee is again assessed for confidence and knowledge with the same questionnaire (figure 5.2, above).

5.1.3.8 Cost of curriculum

Some initial expenditure is required for the design and development of simulated patient scenarios. For the scenarios designed in this pilot curriculum, this required the acquisition of medical stationary for the patients’ paper records as well as medical devices such as
intravenous catheters, urinary catheters, intravenous fluids, and wound dressings, all of which were reusable. These initial costs were estimated at GBP £100.

Running costs of this curriculum are limited to the cost of medical actors to play patient roles and are costed at GBP £50 per trainee, based on a three-patient WR for a standard four-hour half-day course for six trainees. This does not include any fees for faculty, room, or equipment hire, which are accounted for as part of general running costs of the surgical skills centre at St Mary’s Hospital, but would require consideration if implemented at other centres without similar facilities.

5.1.4 Discussion

Despite increasing calls for the improvement of post-operative surgical care,\textsuperscript{209, 329} as well as mounting evidence suggesting a lack of standardisation in this domain,\textsuperscript{65, 66} recommendations on how to address this have been lacking. In order to address an identified training gap, any curriculum must be based upon valid, evidence-based measures to be credible,\textsuperscript{323} and integrate with existing training schemes and resources to be implementable. As such, necessary evidence and tools to build a curriculum for surgical WRs have only recently been developed, and are here collated in the form of a comprehensive curriculum for the first time.

One of the key findings of the previously presented study for the validation of the simulated ward (chapter 3), was the differences in approach to the patient between senior trainees and less experienced juniors. Whilst all clinicians made the correct acute diagnosis for patients in the simulation, junior trainees almost without exception neglected to recognise secondary medical problems or attend to pre-existing comorbidity. It is this “holistic”
approach to the patient, through a structured and methodological WR approach to the patient, underpinned by an understanding of the evidence base, which this curriculum aims to address. Designed according to a validated framework and incorporating evidence-based principles across each stage of development, it provides a comprehensive, robust and affordable approach to training for ward-based surgical care. It represents a multi-modal approach to training, involving traditional didactic sessions, high-fidelity simulation, structured feedback, and debriefing. This is combined with assessment strategies designed to quantify performance of each of the many skills required to demonstrate proficiency in the multi-faceted skill of conducting a surgical WR.

Structured as a half-day course, this curriculum is designed to integrate easily into existing surgical curricula. At a cost of £66 ($110) per trainee (which includes actor hire but not any potential faculty fees), it is not subject to one of the often-cited drawbacks of simulation curricula: cost. Minimal capital expenditure to set up a ward simulator, or avoidance of such costs altogether by utilising in-situ simulation, means it is potentially much more affordable than many other comparable simulation-based training programs for technical skills training.

Whilst this curriculum was based on design for a single half-day course, full implementation with multiple interval training sessions would be desirable to assess and maintain proficiency, learner retention and feedback. Further courses could be tailored to encourage trainee progression with introduction of progressively more difficult or complex cases, with certification of proficiency at appropriately benchmarked levels of performance.

5.1.5 Conclusion
Improved WR performance should result in improved assessment of the patient, as well as management by a multi-disciplinary team of health professionals. Through better or earlier detection of complications or unwell patients, appropriate management may be implemented earlier with improved surgical outcomes. Further research will seek to assess the implementation of the curriculum with pre- and post-trainee assessment, as well as transfer of skills into clinical environment.
5.2 Ward simulation to improve surgical ward round performance: a randomised controlled trial

5.2.1 Introduction

Surgical outcomes do not depend on the quality of surgery alone. As described in chapter 1.2.2, Ghaferi and colleagues have repeatedly demonstrated the significant variations in post-operative mortality which exist across a large number of hospitals despite similar risk-adjusted morbidity rates (i.e. failure to rescue). These differences appear to implicate failings in peri- and post-operative care, with deficiencies in the diagnosis and management of complications resulting in an increased risk of avoidable death.

The surgical WR is the key process in ward-based patient care, setting the course of patient management for the following 24 hours, and representing the primary opportunity for detection, amelioration and prevention of post-operative complications. Skills required of the clinician include cognitive and psychomotor aspects such as the thorough assessment and appropriate management of the patient, but also include critical non-technical skills such as communication skills, leadership, and the ability to effectively manage a multi-disciplinary team.

In addition to the growing body of failure to rescue literature implicating the ward-based phase of care, evidence presented in this thesis has explicitly examined WRs and found a need to address differences in WR quality and patient care.

This chapter has proposed and developed a novel intervention for WR training. The development and implementation of a simulation-based curriculum has significant
precedent in the arena of technical skills, where such an approach has demonstrated great success.\textsuperscript{32, 232} Palter and colleagues recently reported the result of a randomised trial of a technical skills curriculum, wherein the intervention group took part in a curriculum incorporating theoretical preparation, case-based learning, virtual reality and laparoscopic training, followed by the performing of laparoscopic cholecystectomies in the operating theatre.\textsuperscript{232} Compared to a control group exposed to the standard (non-simulation-enhanced) curriculum, the intervention group performed significantly better on the first four cholecystectomies, as measured by the validated rating score used by the authors of the study. The differences in scores between the two groups were shown to narrow over the course of four procedures, with no differences between groups measurable by the fifth cholecystectomy, demonstrating the ability of such a simulation-based curriculum to flatten the initial learning curve otherwise which might otherwise be experienced by patients.

The development and validation of simulated wards as a training modality for ward-based surgical care has provided an opportunity to pursue the same path for non-technical skills such as WRs.\textsuperscript{201, 322, 331} Whilst the simulated ward has been validated as an effective assessment tool, its potential as a training tool to improve practice has not previously been evaluated. However, the controlled environment of a high-fidelity surgical ward simulator allows reproducibility, dedicated educational feedback, and safety for both patients and trainees in a realistic environment.

\textit{The aim of this study was to test the efficacy of a simulation-based curriculum to enhance ward-based surgical care.} This study aimed to evaluate whether a simulation-based curriculum could improve key WR skills including patient assessment, patient management, and non-technical team performance.
5.2.2 Methods

5.2.2.1 Subjects

Junior surgical trainees (core surgical trainees, CT1 and CT2) were recruited to take part. Recruitment took place via an e-mail invitation from the program director for North Thames Core Surgical Training to all trainees to take part in a half-day training programme for ward-based surgical care. Trainees of this level were intentionally recruited for this trial, as they represented the stage in training immediately before that at which trainees typically assume responsibility for daily WRs (i.e. registrar level, ST3 or higher).

5.2.2.2 Study design

This study was structured as a randomised controlled trial (figure 5.4), conducted at a tertiary academic centre. This was run as a half-day simulation-based training session in groups of six trainees at a time, with trainees block-randomised using a computer generated sequence to control or intervention groups. Five sessions were run over the course of two consecutive weeks.
**Fig 5.4** Study protocol CONSORT diagram. Only the intervention group completed the ward round (WR) curriculum prior to the WR performance assessment, the control group conducted the WR according to their own standard practice. Controls received feedback and completed the curriculum afterward (shaded boxes).
On arrival, all trainees were asked to fill in a pre-test confidence questionnaire. Thereafter, the intervention group completed the didactic portion of the simulation-based curriculum. This was immediately followed by an assessed simulated WR of three standardised patients. At the end of the WR, subjects in the intervention group received individualised formative and summative feedback, which included video review of their performance and that of their peers (peer review), though they were allowed to opt out of this if they did not wish their peers to view their performance. At the end of the training day, they completed a second (post-test) confidence questionnaire, as well as a standard course feedback form.

The control group did not receive any specific educational intervention, but were given access to WR-related literature including the Royal College of Physicians / Royal College of Nursing joint statement on ward rounds. They were assessed on a WR of the same three patients as the intervention group, before also filling in a post-test confidence questionnaire. Following this they completed the educational aspect of the curriculum as part of the training session; this was post-assessment and did not form part of the study results.

5.2.2.3 Simulation-based curriculum

The development of an evidence-based curriculum for surgical WRs and ward-based care has been described in the previous section of this chapter. Briefly, the curriculum is structured as a multi-modular half-day training course, incorporating evidence-based training and assessment tools for the conduct of surgical WRs and the assessment and management of surgical patients. Trainees take part in lecture and discussion sessions where best-practice principles are introduced. Emphasis is placed upon generic care processes such as a structured approach to patient assessment and management, and the role of the primary clinician to lead the round, communicate and delegate appropriately. Video examples of good and bad quality WRs are given during the didactic sessions, which
are followed by a simulated assessed WR. Trainees also take part in a structured debriefing session, receiving video-based feedback and faculty-led formative and summative assessment. The intervention (curriculum) group did not receive any specific information regarding the assessment tools being used, or specific behaviours being assessed, beyond the discussion of basic principles of thorough patient assessment and management, nor were they familiar with the simulated WR prior to conducting the assessed WR.

5.2.2.4 Simulated ward environment

The WR was conducted in the simulated ward at the Surgical Innovation Centre, St Mary’s Hospital, UK, a high-fidelity environment which includes realistic background “white noise” recorded from an actual ward and integrated digital camera recorders. Subjects were accompanied by a nurse and an intern, both faculty confederates, and assessed three standardised patient scenarios portrayed by trained medical actors matched to each patient’s age and gender. The scenarios used were identical to the ones used in the previous validation study (chapter 3), thus ensuring content validity. In addition to history taking and examination, trainees were able to access realistic computer-based pathology results, radiology, and paper medical records as required.

Scenarios reflected common general surgical presentations which trainees would be expected to be able to appropriately manage. Patient 1 was an elderly patient five days post colectomy with sepsis and probable anastomotic leak. Patient 2 was a young female patient emergently admitted with appendicitis, and patient 3 a middle-aged female with acute pancreatitis. To introduce a realistic element of temporal stress, trainees were also given a time limit for completion of the WR. This limit was based on data from the previous validation study (chapter 3). In the previous study, subjects had taken a median of 33 minutes to complete a WR of the same scenarios at their own pace, therefore trainees in
this study were given a limit of 30 minutes. At the start of the WR, trainees were told that they were required in theatre half an hour, with verbal warnings given with 10 and 5 minutes left.

5.2.2.5 Assessment methods

WR performance was assessed using the previously described and validated Surgical Ward care Assessment Tool (SWAT). Teamwork and non-technical skills were assessed with the NOnt-TEcnical Skills score for Ward based care (W-NOTECHS), a validated scoring system which measures non-technical performance across five behavioural domains such as communication and leadership.

In addition to a primary rater with expertise in non-technical skills assessment, 30% of all WRs (selected via a computer-generated randomisation sequence) were also rated via video review by a second rater blinded to the subject’s group allocation.

SWAT and W-NOTECHS scores were also compared against performance benchmarks (defined as the median group score) set by the senior trainee (ST5-8 level registrars) group in the previous validation study (chapter 3) incorporating the same scenarios, to gauge the learning effect of the simulation-based curriculum compared to experiential practice.

Post-hoc identification of clinical errors or risk events was performed by two clinician observers, with discrepancies resolved via consensus. Pre- and post- test Likert-scale confidence questionnaires (described in previous section – see figure 5.2) were also used to examine trainee skills and attitudes.

Likert-scale and free-text responses to the course feedback form (figure 5.6) from trainees in the intervention group were summarised. Positive responses were considered a score of 5 or higher, on a Likert scale of 1-7.
SimWard training day

Course questionnaire

Training grade  CT1 / CT2  Intended surgical specialty
____________________________

Gender  M / F  Qualification year ____________

Current hospital ______________________________

<table>
<thead>
<tr>
<th></th>
<th>Strongly disagree</th>
<th>Neutral</th>
<th>Strongly agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>The course covered the outlined aims and objectives</td>
<td>1  2  3  4  5  6  7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The course material and objectives are appropriate for my level</td>
<td>1  2  3  4  5  6  7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I found the course useful and applicable to my own practice</td>
<td>1  2  3  4  5  6  7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The course will help me improve my practice</td>
<td>1  2  3  4  5  6  7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I enjoyed the course</td>
<td>1  2  3  4  5  6  7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The course on the whole is well structured</td>
<td>1  2  3  4  5  6  7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The teaching material used was</td>
<td>Very poor</td>
<td>Average</td>
<td>Very good</td>
</tr>
<tr>
<td></td>
<td>1  2  3  4  5  6  7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What were the **best parts** of the course?

What were the **worst parts** of the course?

What can we do to improve the course next time?

Would you recommend the course to a friend / colleague?  

<table>
<thead>
<tr>
<th></th>
<th>No</th>
<th>Unsure</th>
<th>Definitely</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

Thank you for your time.

**Figure 5.5**  Course feedback form
5.2.2.6 Statistical analysis

All data was anonymised and entered into an Excel (Microsoft Corp, Redmond, VA) spreadsheet and analysed in SPSS Statistics (IBM Corp, Armonk, NY). Mann-Whitney U test was used to compare data between groups. Chi-square test was used to compare the number of scores above and below benchmark values in each group. A p-value of less than 0.05 was considered statistically significant.

5.2.3 Results

5.2.3.1 Study participants

A total of thirty trainees were recruited. One was unable to attend the planned session and not included in the study. Twenty-nine trainees were randomised to either control (n=15) or intervention (n=14) arms.

There were no differences between control and intervention groups for gender, male:female ratio 9:6 vs. 8:6, p = 0.88, or training grade, CT1:CT2 ratio (first or second year of surgical training), 9:6 vs. 8:6, p = 0.35.

5.2.3.2 WR performance

All results reported as median (IQR) unless otherwise indicated. SWAT scores (table 5.3) were significantly higher in the intervention group for patient assessment (control 64.4 (11.6)% vs. intervention 78.8 (10.6)%, p = 0.002) and management (58.3 (20.4)% vs. 69.4 (18.5)%, p = 0.014) components.
Table 5.3  Surgical Ward care Assessment Tool (SWAT) results for assessment (A) and management (M) of individual patient scenarios and overall results. Mann-Whitney U test, values expressed as median (IQR).

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Intervention</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SWAT-A</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient 1</td>
<td>75 (25)%</td>
<td>83 (25)%</td>
<td>0.034</td>
</tr>
<tr>
<td>Patient 2</td>
<td>64 (27)%</td>
<td>82 (27)%</td>
<td>0.012</td>
</tr>
<tr>
<td>Patient 3</td>
<td>55 (18)%</td>
<td>73 (18)%</td>
<td>0.004</td>
</tr>
<tr>
<td>Overall assessment</td>
<td>64.4 (11.6)%</td>
<td>78.8 (10.6)%</td>
<td>0.002</td>
</tr>
<tr>
<td><strong>SWAT-M</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient 1</td>
<td>67 (33)%</td>
<td>67 (44)%</td>
<td>0.149</td>
</tr>
<tr>
<td>Patient 2</td>
<td>44 (22)%</td>
<td>67 (11)%</td>
<td>0.038</td>
</tr>
<tr>
<td>Patient 3</td>
<td>62 (37)%</td>
<td>75 (25)%</td>
<td>0.032</td>
</tr>
<tr>
<td>Overall management</td>
<td>58.3 (20.4)%</td>
<td>69.4 (18.5)%</td>
<td>0.014</td>
</tr>
</tbody>
</table>
Similar differences were seen for overall SWAT scores (60.6 (13.0)% vs. 72.9 (12.0)%, \( p = 0.001 \)) (figure 5.8).

Non-technical skills were assessed using W-NOTECHS. 10/29 (34%) of subjects were rated by a blinded second clinician via video review, with excellent inter-rater reliability, Cronbach’s alpha 0.814. Non-technical skills performance (see table 5.4) was significantly higher in the intervention group (19 (4) vs. 23.5 (2), \( p < 0.001 \)). For individual behavioural domains of W-NOTECHS the greatest differences were seen for leadership, cooperation/resource management, and communication/interaction, with no significant differences for assessment/decision making or global awareness/stress management.
Figure 5.6 Boxplot of performance for control (gray) and intervention (white) groups for SWAT assessment (SWAT-a), management (SWAT-m), total SWAT, and W-NOTECH scores with benchmark performance scores indicated for each (dashed line). W-NOTECHS normed to percentage of maximum value for purposes of this graph.
Table 5.4  Surgical ward-based non-technical skills score (W-NOTECHS) results, with individual component and global scores for control and intervention groups. Mann-Whitney U test, values expressed as median (IQR), range.

<table>
<thead>
<tr>
<th>Component</th>
<th>Control</th>
<th>Intervention</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leadership</td>
<td>4 (1), 2-4</td>
<td>5 (0), 4-5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cooperation and resource management</td>
<td>3 (1), 2-4</td>
<td>5 (0), 4-5</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Communication and interaction</td>
<td>4 (2), 3-5</td>
<td>5 (0), 4-5</td>
<td>0.001</td>
</tr>
<tr>
<td>Assessment and decision making</td>
<td>4 (1), 1-5</td>
<td>4 (0), 3-5</td>
<td>0.063</td>
</tr>
<tr>
<td>Global awareness and stress management</td>
<td>4 (0), 1-5</td>
<td>4.5 (1), 3-5</td>
<td>0.186</td>
</tr>
<tr>
<td><strong>Total W-NOTECHS</strong></td>
<td><strong>19 (4), 11-22</strong></td>
<td><strong>23.5 (2), 20-25</strong></td>
<td><strong>&lt;0.001</strong></td>
</tr>
</tbody>
</table>
5.2.3.3 *Comparison to senior trainee cohort*

Trainees in this study completed identical scenarios, under the same conditions, as trainees in the validation study described in chapter 3.2. When comparing junior trainee performance against the performance benchmarks set by senior registrars in this study\(^{126}\) (figure 5.7), 27% of control vs. 79% of intervention group trainees exceeded benchmark levels for patient assessment SWAT scores (\(p = 0.005\)). Similar results were seen for the patient management component of SWAT (0% vs. 43%, \(p = 0.004\)) and overall SWAT scores (13% vs. 64%, \(p = 0.005\)), as well as W-NOTECHS scores (7% vs. 85%, \(p < 0.001\)).

5.2.3.4 *Clinical errors*

Post hoc analysis of video recorded performances identified clinical errors, which placed patients at risk of harm or may have delayed diagnosis or treatment. Examples included the prescription of penicillin-based antibiotics despite documented allergy, failure to confirm pregnancy status in a young female presenting with abdominal pain, or not specifying a dietary status of nil per oral in a patient to be scheduled for the operating room. Though fewer errors per WR were recorded for the intervention group, this was not statistically significant (control 5(4) vs. intervention 4(4), \(p = 0.112\)).

5.2.3.5 *Confidence questionnaire*

97% (28/29) of subjects completed both pre- and post-test confidence questionnaires. No significant differences between groups were found, all \(p\) values > 0.1.

5.2.3.6 *Course feedback*
All subjects in the intervention group completed the feedback questionnaire (14/14, 100% response rate). Feedback was excellent across all statements (table 5.5). All respondents felt that completion of the curriculum was relevant and would improve their clinical practice. 100% (14/14) of respondents enjoyed the training session and would recommend it to colleagues. In free-text responses naming the best parts of the session, residents cited the realism of the simulation (86%, 12/14 responses), quality of the didactic session (36%, 5/14), and video feedback (3/14, 21%). A single negative response to the curriculum was given, where trainees were asked, “What was the worst part of the course?” one respondent answered, “simulation.” However, this subject also subsequently listed the realism of performing a full WR in a training environment as the best part of the course.
Table 5.5 Results of course questionnaire. Positive responses are “agree” to “strongly agree,” or “good” to “very good” on Likert-scale questionnaire responses, range 1-7.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Response (mean ± SD)</th>
<th>Positive responses* (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Course covered stated aims and objectives</td>
<td>6.0 ± 0.7</td>
<td>100% (14/14)</td>
</tr>
<tr>
<td>Course material was appropriate for level of training</td>
<td>6.1 ± 0.7</td>
<td>100% (14/14)</td>
</tr>
<tr>
<td>Course was applicable to my own practice</td>
<td>6.1 ± 0.7</td>
<td>100% (14/14)</td>
</tr>
<tr>
<td>Course will help me improve my clinical practice</td>
<td>6.1 ± 0.7</td>
<td>100% (14/14)</td>
</tr>
<tr>
<td>I enjoyed taking part in the course</td>
<td>6.1 ± 0.7</td>
<td>100% (14/14)</td>
</tr>
<tr>
<td>The course was well structured</td>
<td>6.0 ± 0.8</td>
<td>93% (13/14)</td>
</tr>
<tr>
<td>Rate the quality of teaching material</td>
<td>5.6 ± 0.8</td>
<td>100% (14/14)</td>
</tr>
<tr>
<td>I would recommend the course to a colleague</td>
<td>yes/no only</td>
<td>100% (14/14)</td>
</tr>
</tbody>
</table>

*Positive response considered score of 5 or higher on Likert scale of 1-7.
5.2.4 Discussion

This study represents the first trial to empirically examine the effectiveness of training interventions for the delivery of ward-based surgical care. It establishes the feasibility of using a simulation-based training curriculum within the context of existing surgical training programs, and demonstrates how the introduction of a comprehensive training curriculum for WRs can result in significantly improved performance in junior trainees, with regard to the thoroughness of patient assessment and management plan, and quality of interaction within the clinical team.

The introduction of a systematic approach to the patient, as part of the simulation-based WR curriculum described previously in this chapter, resulted in a more thorough assessment as reflected by higher SWAT assessment scores. Results presented earlier in this thesis have previously demonstrated how poor quality WRs, with lack of a structured approach to thorough patient assessment, may result in neglect of secondary patient conditions (as demonstrated by the junior trainee cohort of the validation study presented in chapter 3.2), and increased risk of preventable complications (as seen in the poor quality WR group of the observational study data presented in chapter 4). As such, the improvement of SWAT scores seen in this simulation-based trial has the potential to lead to earlier detection and treatment of complications in the clinical environment.

Whilst basic concepts of patient management, such as determination of further investigations or treatment, and appropriate documentation, are discussed, the curriculum implemented in this study focuses primarily on patient assessment, advocating a structured, standardised approach, which may be applied to patient generically. As management of patient conditions naturally requires a more diagnosis-specific approach, this is less suited to a generic “core skills” training curriculum for WRs. An approach to improve this aspect of patient WR care is considered separately in the following chapter.
Despite the curriculum’s focus on patient assessment, however, the management of patients did, in fact, improve, as evidenced by higher SWAT management scores. But as the mechanisms to achieve improvement in assessment and management are inextricably linked, this is not necessarily surprising. Firstly, improved patient assessment may simply lead to increased detection of morbidity, resulting in more appropriate management. Secondly, if non-technical skills can be improved, this may in turn improve overall performance, e.g. through appropriate delegation of tasks to reduce distractions to the primary clinician, or the enabling effect of improved communication on the execution of management plans by the rest of the team. Finally, adopting a structured approach to assessment and management may also have the effect of reducing the cognitive burden on the clinician, increasing their ability to engage in other activities without requiring significant attentional resources – a characteristic which has been termed automaticity. In other words, by having a set structure for the WR in mind, trainees in the intervention group were required to devote less time to thinking about what to do next as part of the WR, and more considering the implications of their findings, or the management of other aspects of the patient beyond the primary diagnosis. The effects of simulation-based training to automaticity have been demonstrated in a randomised trial published by Stefanidis et al. Intervention group participants “overtrained” on a simulation model, whereby proficiency was not only measured on previously defined simulator metrics, but also on the ability to complete a secondary visual-spatial task at the same time. Automaticity was demonstrated though trainees’ ability to complete the secondary task despite safely performing laparoscopic suturing on a simulator at the same time. Subsequent assessment of a Nissen fundoplication on a live porcine model demonstrated superior performance by the automaticity group compared to the same group’s training at “proficiency” stage, or conventionally trained control group.
Teamwork, leadership and managerial skills are crucial elements of WRs, but are potentially also the skills least cultivated in junior trainees. Whilst they will be accustomed to independently assessing patients and establishing basic management plans, the transition to a more senior role in which they must lead a WR and multi-disciplinary team can be a daunting one. The importance of non-technical skills has been demonstrated in other contexts, with non-technical skills interventions shown to result in improved teamwork and error reduction in the operating room environment. As part of a comprehensive approach, the WR training curriculum purposefully avoided focusing on (sub-)specialty- or patient-specific skills. Rather, the didactic session stressed generic non-technical skills, and a structured and thorough approach to WRs in general, backed by current evidence and expert guidelines, resulting in an improvement in non-technical skills as measured by W-NOTECHS.

A recent study by Birkmeyer et al reported that higher ratings of teamwork were associated with decreased complication rates following bariatric surgery. Data from a statewide bariatric surgery collaborative in Michigan, USA, was compared to results of staff responses to validated patient safety surveys. Overall safety ratings, especially as rated by nursing staff, were strongly related to adjusted rates of serious complications (risk ratio 1.73 (1.35, 2.23), p < 0.0001). Significant relationships were also found for responses relating to a number of individual survey statements relating to communication and cooperation within the hospital and the operating theatre. Authors concluded that patient safety culture was significantly associated with rates of complications, particularly highlighting classroom and simulation-based team training as a potential solution. Such data highlights the importance of improving WR quality and the need to replace the current Halstedian system of learning, and is particularly poignant for junior trainees. Underscoring the effectiveness of simulation-based learning, in this study many junior trainees were able to achieve levels of WR performance similar, or even superior, to trainees at least three years more senior, despite
their lack of experience. Comparison of scores between this study’s cohort and the senior trainee (registrars of level ST5 or greater) group from the previous validation study (chapter 3) demonstrated the potential of a curriculum for WRs to flatten the learning process. After exposure to the simulation-based curriculum, almost two-thirds of junior residents met benchmarks previously set by the senior trainee group for the same WR scenarios.

Several limitations of this study must be considered, including a lack of baseline WR testing, as well as a lack of further WR assessment following debriefing and feedback. However, for this study the desire was to implement a curriculum which could be realistically incorporated into existing curricula. As such, it was adapted to a currently utilised model for technical skills training, which incorporates half-day training sessions, meaning that further assessment sessions were not feasible to implement on this occasion. The provision of a second WR assessment would ideally be considered in future studies. Furthermore, the number of participants, the fact that all were junior trainees with minimal WR experience, randomised and blindly allocated, and that all were within a 12-month range of each other in clinical experience can be expected to minimise the risk of any occult selection bias.

Finally, pre-test confidence questionnaires were statistically equal.

Though full blinding of assessments was not possible in this randomised trial, the SWAT is by definition an objective checklist measure of performance and vulnerability to bias is low. To avoid bias in subjective W-NOTECHS ratings, the second assessor reviewed video recordings without knowledge of the trainee’s group allocation, and did not know the trainees personally.

Competing clinical priorities, such as operating room schedules and routine interruptions through pages and phone calls, can result in the surgical WR being de-prioritised.\textsuperscript{197} Considering the known variability in ward-based care and resulting surgical outcomes,\textsuperscript{67} combined with critical role of WR process in daily management of patient, the explicit
inclusion of WRs in the modern paradigm of surgical training is past due. Trainees commented that they appreciated the opportunity to receive formalised feedback in an area they would not receive training on ordinarily, reflecting that trainees may be unprepared to take on the responsibility of conducting WRs independently.\textsuperscript{201} With proficiency-based progression and training for technical skills now taking hold,\textsuperscript{336} it is time to consider and implement these for ward round performance too.

\subsection*{5.2.5 Conclusion}

This study shows ability of a simulation-based curriculum to assess, inform and improve clinician performance in surgical WRs. Along with these promising results, it is important to note that it was universally well received by trainees who participated. Further assessment of the curriculum and its effect on skill in both simulated and clinical environments is required, including consideration of skill retention and generalisability to other centres. On the basis of this initial cohort, however, conducted in a high-fidelity, validated environment, this study presents an important tool to assess and improve practice, and potentially outcomes, in a critical area of the surgical pathway of care.
Chapter 6

Development and implementation of checklists for the management of postoperative complications
6.1 Development of surgical care checklists to optimise patient care following postoperative complications

6.1.1 Introduction

Despite continuing advances in surgical care, postoperative complications remain a common and accepted risk of surgery.\textsuperscript{67,148} Interventions such as educational curricula to improve surgical WRs have been shown to potentially improve the assessment of patients, and may thus lead to earlier detection of complications. It is also, of course, the management of these complications in the postoperative phase which is vital to the successful recovery of patients who suffer postoperative morbidity. As had been discussed in the preceding chapters of this thesis, however, the quality of management of complications has been shown to vary greatly, as reflected by differing rates of death following the development of complications between institutions, i.e. failure to rescue.\textsuperscript{65,66}

Investigation of the factors underlying this variability (chapter 1.3) has identified certain structural factors, such as nurse : patient staffing ratios.\textsuperscript{104} However, this has accounted only for a small proportion of clinical variability seen. In a recent survey of 7,906 American surgeons, over 70% of respondents attributed witnessed medical errors to individual-level (i.e. process), rather than system-level (structural), factors.\textsuperscript{337}

Studies of error and patient safety have indicated that the majority of human process errors result from passive errors, or “errors of omission.”\textsuperscript{338-340} The reduction and mitigation of such failures, Reason has suggested,\textsuperscript{341} are best addressed through checklists or protocols. Functioning as mental aids or prompts, checklists have been adapted for use in surgery following the example of other high-complexity industries, such as nuclear safety and
In these industries, checklists exist to guide the management of specific crises, such as engine flameouts or a failure to deploy landing gear. Thus far, checklists in surgery have been most widely employed to address generic care processes, such as with the World Health Organisation Safer Surgery Checklist\textsuperscript{345}, or in a broader sense to guide patient postoperative care in general, in the form of enhanced recovery protocols.\textsuperscript{95} As previously discussed in greater detail (chapter 1.1), care outcomes have been shown to improve significantly in relation to levels of compliance with care protocols defining best-practice care.\textsuperscript{140}

More recently, diagnosis-specific checklists have been trialed for the management of operating room crises.\textsuperscript{143, 343} The standardised management of critical events was significantly improved with the introduction of crisis checklists in a randomised trial by Arriaga and colleagues.\textsuperscript{343} In this study, multidisciplinary theatre teams were placed in a high-fidelity simulated operating theatre environment and exposed to a number of crisis scenarios such as operating room fires, or failed airways. Checklists were developed to guide the management of these critical events, and made available to teams the intervention group, resulting in a significant reduction of missed critical steps (6% vs. 23%, $p < 0.001$). 97% of participants indicated they would want the checklists used in the event that they themselves experienced such an event.

It stands to reason that a similar approach may be required for postoperative crises too. Whereas intra-operative crises are, fortunately, rare (with an incidence of less than 1.5%),\textsuperscript{344} this stands in stark contrast to the high incidence of postoperative complications, which may affect as many as half of patients following major surgery.\textsuperscript{67, 148}

The variability in the management of postoperative complications occurs despite established, evidence-based guidelines for the treatment of many of the most common postoperative complications, such as catheter-associated urinary tract infections.\textsuperscript{345} It was hypothesised
that the development of checklists for common postoperative complications might therefore have a significant impact on reducing error and standardising management of these conditions.

This study aims to identify the most common complications occurring after complex gastrointestinal surgical procedures, and presents a structured framework for the development and validation of treatment checklists for patient management.

### 6.1.2 Methods

A multi-phase, iterative design process was adopted. This was informed by previously published methodologies for surgical checklist design,¹⁴³, ³⁴⁶-³⁴⁸ and involved initial literature review, followed by expert assessment, casenote-based validation, and end-user feedback (figure 6.1).
Figure 6.1 Iterative checklist development process

Phase 1
Literature review

Identify most common post-operative complications

Identify evidence-based management of complications

First draft of checklist (1\textsuperscript{st} iteration)

Phase 2
Expert review

Reviewed by experts in surgery and critical care

Revision of checklist (2\textsuperscript{nd} iteration)

Phase 3
Casenote review

Recent cases of morbidity identified, casenotes reviewed

Validated against current practice

Phase 4
End-user review

Reviewed by residents and nursing staff

Final checklist (3\textsuperscript{rd} iteration)
6.1.2.1 Literature review

An online search of published literature was performed, limited to the last 2 years to reflect current surgical practice. The PubMed online database was searched for studies published 2012 – 2013, which reported postoperative outcomes for gastrointestinal surgical procedures. The following search terms were used: (“postoperative outcome” or “complications”) and “surgery” and (“esophagectomy” or “gastrectomy” or “colectomy” or “pancreat*”), including MeSH terms “surgical procedures, operative,” “gastrectomy,” and “colectomy.” Both benign and neoplastic resections were considered, with studies reporting at least 1000 cases or more included for final analysis.

Data on incidence of postoperative complications was extracted from included publications. Reported complications were ranked in order of incidence, rather than using absolute complication rates, to control for the variable risks of morbidity for different surgical procedures. Mean of ranks across all included studies were calculated, with the most common (i.e. highest ranked) complications included for checklist development.

6.1.2.2 Checklist development

Initial checklists were drafted for the management of each identified complication based on best published evidence and guidelines published by national and international specialty associations. These were then combined with local protocols, where applicable (e.g. local antibiotic prescribing guidelines), to ensure integration with local practice and culture.

6.1.2.3 Clinical expert review

Each checklist was subsequently reviewed by a panel of clinical experts, who were asked to comment upon checklist usability, content, and practical applicability to their personal
practice. In addition, they were asked to identify any common clinical errors they felt were not addressed by the checklist which might merit inclusion. Finally, they were asked whether they would want their trainees using the checklists in the management of patients under their care. Feedback gained from the expert panel resulted in revision and second iteration of the checklists.

6.1.2.4 Casenote review and clinical validation

To assess checklists’ applicability to current practice, and potential to improve care, a retrospective casenote analysis was performed. The last forty consecutive postoperative complications to which the checklists applied were identified from a prospectively maintained database of gastrointestinal surgery patients at an urban tertiary academic centre in London, UK. Complications were defined on an intention-to-treat basis as recorded by the clinical team. Patient case notes were retrieved and management of complications compared to checklist items. Completion of checklist items was recorded.

A clinician researcher experienced in casenote review recorded delays or process errors occurring in the management of complications. A second clinician independently reviewed 30% of cases, with inter-rater reliability on checklist compliance and incidence of errors calculated using Cohen’s Kappa, a statistical measure of the likelihood of inter-rater agreement occurring by chance alone. Errors were defined according to Reason, as any action which failed to achieve its aims, or an incorrect action for a given aim. Delays were defined as failures of prescribed actions or treatments to take place at the intended time.

6.1.2.5 Statistical analysis
The relationship between quality of management of complications, as measured by the number of checklist items completed, and subsequent patient outcomes was analysed. Regression analysis, using a forward logistic model, was performed for the risk of developing multiple morbidity, with checklist compliance, age, ASA, and gender considered as independent variables (IBM SPSS Statistics 21, IBM Corp, Armonk, NY).

For all analyses, a p value of less than 0.05 considered statistically significant.

6.1.2.6 End-user review

A penultimate version of the checklists was presented to a multi-disciplinary group of end-users, comprised of residents, interns, and nursing staff. These were asked to review format, readability, and practicability of use. They were asked to identify any desired changes or errors, and whether they would use checklists in their own clinical practice. This feedback was taken into account and a final iteration of checklists completed.

6.1.3 Results

6.1.3.1 Literature review

The literature search as described returned six publications matching the pre-defined inclusion criteria (see table 6.1), describing postoperative outcomes for 81,936 patients following colectomy, esophagectomy, gastrectomy, and pancreatectomy. Pooled results demonstrated an overall postoperative complication rate of 17 – 43%.
Calculating the mean ranked incidence of all complications, wound infection, hospital acquired pneumonia, sepsis, intra-abdominal sepsis / anastomotic leak, and urinary tract infection were identified as the five most common postoperative complications, and considered for checklist development. Cardiac and renal failure were the sixth and seventh most common complications, but were excluded as these descriptions cover a potentially broad range of underlying pathologies (e.g. congestive heart failure vs. acute myocardial infarction) and were therefore not suitable for a standardised approach to treatment without a more specific diagnosis. Surgical bleeding, as the next most common type of morbidity, was included.

The six complications selected for checklist development (wound infection, pneumonia, sepsis, anastomotic leak, urinary tract infection, and bleeding) accounted for 92.2% of all reported morbidity. Pulmonary embolus and all other reported postoperative complications each made up less than 1% of all reported morbidity and were thus not further considered.
Table 6.1: Type and incidence rate of most frequently reported postoperative complications.

<table>
<thead>
<tr>
<th>Study</th>
<th>Data source</th>
<th>Procedure</th>
<th>n</th>
<th>Overall</th>
<th>Wound infection</th>
<th>HAP</th>
<th>Sepsis</th>
<th>IA sepsis</th>
<th>UTI</th>
<th>Cardiac</th>
<th>Renal</th>
<th>Bleed</th>
<th>PE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kwaan, 2013. USA.</td>
<td>NSQIP</td>
<td>Colectomy</td>
<td>4,875</td>
<td>17</td>
<td>7</td>
<td>2.3</td>
<td>4.3</td>
<td>2.8</td>
<td>2.9</td>
<td>0.4</td>
<td>0.7</td>
<td>0.4</td>
<td>0.7</td>
</tr>
<tr>
<td>Molena, 2013. USA.</td>
<td>NIS</td>
<td>Esophagectomy</td>
<td>18,966</td>
<td>n/r</td>
<td>18.25</td>
<td>5.92</td>
<td>5.92</td>
<td>1.46</td>
<td>6.19</td>
<td>3.75</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yasunaga, 2013. Japan.</td>
<td>National database</td>
<td>Gastrectomy or colectomy</td>
<td>30,765</td>
<td>15.5</td>
<td>5.1</td>
<td>2.3</td>
<td>1.2</td>
<td>1.7</td>
<td>0.9</td>
<td>2.2</td>
<td>0.1</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>Merkow, 2012. USA.</td>
<td>NSQIP</td>
<td>Esophagectomy</td>
<td>1,600</td>
<td>43</td>
<td>20</td>
<td>30</td>
<td>26</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cone, 2012. USA.</td>
<td>NSQIP</td>
<td>Colectomy</td>
<td>24,730</td>
<td>n/r</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grant, 2012. USA</td>
<td>NSQIP</td>
<td>Pancreatectomy</td>
<td>1,030</td>
<td>n/r</td>
<td>9</td>
<td>6</td>
<td>1.8</td>
<td>2.7</td>
<td></td>
<td>1.8</td>
<td>1.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td>81,936</td>
<td>9.92</td>
<td>3.59</td>
<td>4.63</td>
<td>2.49</td>
<td>1.75</td>
<td>1.58</td>
<td>0.74</td>
<td>0.15</td>
<td>0.04</td>
<td></td>
</tr>
</tbody>
</table>


n.B. Only most commonly reported complications shown, therefore percentages of all reported complications do not sum to 100%.
6.1.3.2 Checklist development

A search of published peer-reviewed and grey literature was conducted on guidelines for the management of the six identified complications, and combined with local prescribing and treatment protocols in the development of the first iteration of Post-Operative Surgical Checklist for Hospital care (POSCH).

6.1.3.3 Clinical expert review

A group of five clinical experts in surgery (n = 3) and critical care and anaesthesia (n = 2) were consulted on the use of the POSCH checklists in clinical practice. All (5/5, 100%) agreed that the checklists were useful and relevant in the management of postoperative complications, and that they would encourage their trainees to use them.

The expert panel recommended minor modifications which were incorporated in the second iteration of the checklists. For example, experts identified a common clinical error – the delay to initiation of treatment when antibiotic therapy was initiated, but no loading dose prescribed. Relevant checklist items were changed to include the prescription of loading doses to address this. In addition, minor changes were made to phrasing and formatting (i.e. informing of a senior member of team, previously included as a recommendation, made a mandatory checklist item).

6.1.3.4 Casenote review and clinical validation

Of the forty identified cases, three clinical records could not be retrieved, resulting in a total of 37 casenotes undergoing review (see table 6.2). 30% (11/37) were reviewed by an independent second clinician reviewer, with very good inter-rater reliability (Cohen’s Kappa = 0.775, statistically highly significant, p < 0.001). Overall checklist compliance was 60 (24)%,
with full compliance (100% of checklist items met) seen in only 16% (6/37) of cases. All summary results are reported as median (IQR) unless otherwise indicated.

65% (24/37) of cases were associated with delay or error in treatment (see table 6.3). 37% (9/24) of errors related to general process error, such as failure to remove a urinary catheter associated with infection. Errors relating specifically to prescription errors, generally relating to a failure to comply with local antibiotic guidelines, were seen in a further 37% (9/24). In 25% (6/24), errors related to communication failures, such as where a suspected diagnosis was made but not acted upon due to delays in communicating to an appropriate member of the clinical team. These errors resulted in median delays of 6 (5.4) hours, (range 1 – 36 h) to the patient receiving appropriate treatment. All documented errors were considered preventable with the appropriate use of the developed checklists.
Table 6.2 Results of casenote review

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>37</td>
</tr>
<tr>
<td>Age (years)</td>
<td>73  [58-82]</td>
</tr>
<tr>
<td>M:F ratio</td>
<td>24:13</td>
</tr>
<tr>
<td>Checklist compliance</td>
<td>60 [24]%</td>
</tr>
<tr>
<td>100% compliance achieved</td>
<td>16 % (6/37)</td>
</tr>
<tr>
<td>Delays to treatment</td>
<td>6 [5.4] hours</td>
</tr>
<tr>
<td>Results reported as median [IQR]</td>
<td></td>
</tr>
<tr>
<td>Event</td>
<td>n (%)</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------</td>
</tr>
<tr>
<td>Process error</td>
<td>9 (37%)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Prescribing error</td>
<td>9 (37%)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication failure</td>
<td>6 (25%)</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

UTI: urinary tract infection
6.1.3.5 Statistical analysis

The relationship between checklist compliance and outcomes were analysed. Logistic regression analysis, considering checklist compliance, ASA grade, age, and gender as factors for the risk of developing further (multiple) morbidity revealed checklist compliance of less than 50% to be to only significant determining factor, OR 6.75 (95% CI 1.11, 41.00), p = 0.038.

6.1.3.6 End-user review

A group representing potential end-users (n = 10) of the checklist, comprised of residents (n = 4), interns (n = 3) and nurses (n = 3) reviewed the checklists. All (10/10, 100%) responded positively to the question of whether they would wish to use them in their own practice. On the recommendations of this end-user group, minor changes to formatting and layout were made to improve usability and readability in a final revision of the checklists (see appendix 4 for final version). In addition to printed reference panels, they were designed as adhesive stickers to be entered into the written medical record upon use. No further content changes were deemed necessary.

6.1.4 Discussion

The incidence of postoperative morbidity following major surgery remains high in current practice. Optimal management of complications requires appropriate treatment to maximise the patient’s chance of recovery and prevention of further deterioration, or death (failure to rescue). However, variations in care quality and the unavoidable nature of human error mean that the management of postoperative complications remains inconsistent.37
To reduce error and enhance standardisation of care, this study describes the development and validity evidence of checklists for common gastrointestinal surgical morbidity, with the six developed checklists encompassing over 92% of reported postoperative complications. By integrating evidence-based care processes with local procedures and culture, and involving end-users and clinical experts in the development process, this study presents a framework for the development of an intervention for the optimisation and standardisation of an important aspect of postoperative care.

Errors in management place patients at risk of receiving inappropriate treatment and delays, which may negatively affect outcomes. In this study, we report that failure to adhere to best-practice principles is significantly associated with an increased risk of further morbidity. Analysis of current practice through casenote review revealed significant potential to improve care through appropriate checklist implementation, with checklist compliance only 60%, and full compliance with checklisted items in only 16% of cases. Though checklists were revised following expert and casenote review, this resulted in modifications to the 38 items included in the final six checklists only (e.g. explicit instruction to prescribe loading dose of antibiotics as part of prescribing regimen item), rather than additional discrete items, and as such did not artificially affect checklist compliance rates. Rather, these figures are reflective of poor adherence to best-evidence mandated care alone. Whereas POSCH checklists could address most process errors, a significant number of treatment delays also resulted from communication errors, rather than to omission errors correctable through inclusion as a specific checklist item. Furthermore, as only those communication errors explicitly documented in the medical record could be recorded in this study, either through specific documentation of errors which had occurred, or obviously contradictive entries, their true incidence is likely to be have been higher still. The prevalence of communication errors in surgical care, and their relation to adverse events, is well documented. Increasingly, however, evidence suggests that the improvement of
communication is in fact one of the primary mechanisms through which checklists affect change, as concluded by a recent systematic review by Russ and colleagues investigating the effect of operating theatre checklists on teamwork and communication.\textsuperscript{357} By mandating a set care pathway to be followed by clinical staff, studies have shown checklists capable of flattening clinical hierarchies and improving safety attitudes as they improve outcomes.\textsuperscript{358} Implementation of POSCH checklists could be expected to exert a similar effect. With all senior clinicians and end-users involved in the development process universally supportive of the use of checklists in clinical practice, these may encourage junior members of the team to initiate appropriate care earlier, and empower nursing staff to query any unchecked items.

One strength of the checklist development process presented in this study is the inclusion of local care protocols and clinicians in the finalisation of the checklists. In this manner, applicability to local practice and buy-in from staff could be assured. Barriers to the successful adoption of interventions in the past, such as the WHO Safer Surgery Checklist, have included a lack of understanding of their use and purpose, highlighting also the need for training in their use prior to introduction.\textsuperscript{359} A recent trial of operative crisis checklists piloted their use in a simulated setting,\textsuperscript{343} with the use of simulation for the training and certification of proficiency for technical skills already well established.\textsuperscript{32,232}

Atul Gawande and colleagues, building on the considerable success of the WHO Safer Surgery Checklist, have published a “checklist for checklists” as part of a greater drive to develop surgical safety checklists for other areas of care, Project Check.\textsuperscript{348} This checklist, one of the frameworks which informed the development of the POSCH checklists, describes steps of needs analysis, checklist design (e.g. font, colour, uncluttered format), and validation – specifically recommending repeated simulated or pilot clinical trials to allow for refinement and modification prior to full implementation.
The introduction of POSCH checklists should endeavour to a similar approach, involving high-fidelity ward simulation\textsuperscript{326} to ensure staff understanding of, and proficiency in, their use prior to clinical implementation. Dissemination of checklists to other centres would also need to be adapted to local needs through a similar process as described in this study to take into account local policies and resources (ie. integration into electronic records systems).

The results of this study must be considered in the context of several limitations. The initial literature search was non-exhaustive, and limited to recent, large cohort studies only. However, this was intended to reflect current practice, and across these publications, there was broad agreement in the type of complications reported, regardless of procedure. The validation of the checklists was a single-centre, retrospective casenote review, representing a source of selection bias. To limit this, this study included consecutive patients receiving treatment in a tertiary academic centre, rated as one of the top five centres in England according to standardised hospital mortality indicators.\textsuperscript{360} Any bias is therefore likely to have been in a positive direction, with levels of adherence to best-practice in other centres likely to be lower still.

It is important to note that in the majority of cases, though care was imperfect according to the checklist-defined criteria, all patients recovered fully, with no deaths in the assessed patient cohort. Nevertheless, there was a high incidence of treatment delays and other errors.\textsuperscript{361} These were associated with an increased risk of further morbidity and were likely to have resulted in a prolonged length of stay, preventable patient suffering, and excess healthcare expenditure – all of which, this analysis has suggested, could have been prevented with use of the checklists. In the modern era of economic austerity, patient-related outcome measures, and patient choice, the optimisation of the patient experience and other associated metrics are increasingly seen as fundamental metrics of care, second only to morbidity and mortality. POSCH checklists may present a further tool to achieve this.
Finally, though this study discusses implementation and training strategies, a lack of in vivo clinical data means that there may be obstructions to effective clinical use that have not yet been encountered. However, checklists represent an amenable intervention, whose effectiveness in other areas of care have been thoroughly and well documented. Use of the checklists in other centres would require adaptation to local culture through a process similar to that described in this study, incorporating end-users in the process to maximise staff buy-in. With appropriate implementation, there can be little doubt that improvement of clinical care, in line with accepted international guidelines, can improve care.

6.1.5 Conclusion

This study presents a development and validation of checklists to guide management of complications following gastrointestinal surgery for the improvement of postoperative care. It presents six initial checklists for management of the most commonly reported types of morbidity, based upon best evidence and adaptable to local practice and clinical culture. This is the first study to assess patterns of error in the management of these same common complications, suggesting broad scope for improvement even in a tertiary academic centre, and validating the applicability of checklists in this context. Future research to assess feasibility, clinical implementation, and effect on patient outcomes, is required.
6.2 Surgical care checklists improve postoperative care in a simulated environment: a randomised controlled trial

6.2.1 Introduction

The brief literature review performed as part of the development of the POSCH checklists (chapter 6.1) suggested that many of the most frequently reported events, such as pneumonia or wound infection, are common to all patients undergoing gastrointestinal surgery, irrespective of the specific procedure. Despite their relatively common nature, however, the high degree of variability in the management of postoperative complications has been repeatedly evidenced by both failure to rescue literature and the assessments of WRs reported in preceding chapters of this thesis. The casenote review of 40 consecutive patients suffering postoperative complications presented in chapter 6.1 has contributed further evidence to this hypothesis, demonstrating the significant variation in management of common complications, with poor adherence to optimised care guidelines set out by various specialist medical bodies, which exists in current care.

Prevention and management of complications is subject to the appropriate conduct of the surgical ward round, but is subject to the pressures of time, resource limitation, and the competing prioritisation of other clinical activities. Coupled with natural tendencies for human error, the ward-based phase of surgical care is one rife with potential for error and adverse events. In keeping with Reason’s recommendations, POSCH checklists were developed (chapter 6.1) to serve as a critical barrier layer of defences to prevent error and improve care.

Checklists in surgery have been widely adopted in recent years, with their use becoming the accepted standard of care in many areas of surgical care. Evidence also suggests that
checklists may act beyond the checklisted processes themselves to improve safety culture as a whole, through the empowerment of individuals and flattening of clinical hierarchies.\textsuperscript{357, 358}

At the same time, however, the wider uptake of checklists has served to highlight some of the potential pitfalls which may be encountered when attempting to change established process and clinical culture.\textsuperscript{366} The clinical effectiveness of checklists is dependent on more than the content of the checklist alone. Without effective staff involvement and training, implementing such interventions may be at best, ineffective,\textsuperscript{367} or at worst, a source of staff frustration and discord.\textsuperscript{368}

To address this issue, interventions have been shown to benefit from initial trialing in simulated environments to assess feasibility and clinical effectiveness, prior to embarking upon full clinical implementation.\textsuperscript{343} Such an approach benefits from the known advantages of simulation, with controlled and reliable scenarios within a dedicated assessment space, and elimination of the potential risk to patients.

The aim of this study was to assess the feasibility and clinical effectiveness of POSCH checklists in the management of common postoperative complications, within a high-fidelity simulated ward environment.

6.2.2 Methods

6.2.2.1 Study participants

General surgical registrars were recruited from a single academic hospital. These were in their third year of specialty training (ST3) or higher, reflecting the level at which residents typically assume responsibility for routine patient assessment during the ward round (WR) in
typical practice in the United Kingdom. A sample size calculation assuming 90% power and
an alpha level of 0.05 was performed, utilising the rates of surgical WR error from the
observational study previously discussed in chapter 4, and the expected effect size of
surgical process checklists as reported in the study by Arriaga et al, which examined process
checklists in the management of operating theatre crises, and was also performed in a
simulated environment. This calculation suggested a necessary sample size of only three
subjects in each group, which was increased to ten to maximize data collection and increase
the validity of results, and account for any increased variability in our sample.

6.2.2.2 Setting

Simulated WRs took place within the same high-fidelity three-bed simulated ward described
previously in this thesis. Again, professional medical actors were employed to simulate
patients and exhibit realistic clinical signs. An integrated audio-visual system digitally
recorded all activity and played realistic background “white noise” recorded from a real
clinical environment.

6.2.2.3 Study design

Trainees were exposed to a series of simulated scenarios in which patients experienced one
of the following postoperative complications: pneumonia, anastomotic leak, wound
infection, urinary tract infection, postoperative bleeding, and intra-abdominal sepsis. These
complications were previously identified as the most common types following major
gastrointestinal surgery, with development of checklists for their management previously
described (chapter 6.1).
All trainees conducted a baseline WR of three patients within the simulated ward according to personal practice (see figure 6.2). WRs were supported by a staff nurse and a house officer (both played by simulation confederates not familiar to the trainee). Trainees were blinded to study design, purpose, and assessment endpoints. To create a realistic sense of time pressure, trainees were told they were required to finish each WR within thirty minutes, based on previous data of similar scenarios. The order in which patient scenarios were assigned to each trainee’s WR were determined through a computer-generated randomisation process, with all six scenarios being assessed by each resident over the course of the study.

Following the baseline WR assessment, each trainee was randomised (using a further computer-generated randomisation sequence) to either control or intervention groups. Residents in the intervention group received a didactic session, with written and verbal instruction, on use of the previously developed checklists for management of surgical complications. They then completed their second assessed WR, during which the checklists were available for optional use. Checklists were formatted as sticker labels to be checked, signed, and placed in the medical record, and were available at each bedside.
Figure 6.2 CONSORT study flowchart.

- Trainees recruited (n=20)
- Baseline WR assessment (3 patients per WR)
  
  Allocation
  
  Randomised to control
  n = 10
  
  Second simulated WR
  (3 patients per WR)

  Randomised to intervention
  n = 10
  
  Checklist training
  
  Second simulated WR
  (3 patients per WR)
  
  Feedback questionnaire
Trainees in the control group conducted a second WR according to their own personal practice, thus acting as controls for any potential learning effects resulting from repeated exposure to the simulator. Baseline and final WRs were conducted in separate sessions on different days for both groups over the course of two weeks.

6.2.2.4 Assessment methods

Primary endpoint for the study was the rate of failure to execute critical management steps for each postoperative complication. Each complication management checklist involves five to eight management processes, such as the ordering of blood cultures and administration of antibiotics for sepsis, or arranging a blood transfusion and preparing a patient for return to the operating theatre in case of postoperative haemorrhage. Completion or omission of each checklist item was recorded in a binary manner through direct observation.

Secondary endpoints included technical and non-technical performance in conduct of the WR. Technical performance was assessed using the Surgical Ward care Assessment Tool (SWAT), a checklist-based, validated tool to assess thoroughness of patient assessment (SWAT-A) and management (SWAT-M). Non-technical performance was assessed using the W-NOTECHS scale, a validated tool which assesses five behavioural domains on a Likert scale, resulting in a final score of 5-25. Prescribing errors were also recorded.

Each WR was digitally recorded using two integrated ceiling cameras within the simulated ward, with 20% of WRs rated by a second experienced rater blinded to each subject’s group allocation, to assess potential for observer bias. Inter-rater reliability was assessed using Cronbach’s alpha.

Following completion of the final observed WR, residents in the intervention group were also surveyed regarding their perception of the checklists, with questionnaire responses graded from 1 (strongly disagree) to 5 (strongly agree) on a Likert scale.
### 6.2.2.5 Statistical analysis

Demographic and WR performance data for failure rate, SWAT, W-NOTECHS, and time taken was compared between groups, in addition to within-group comparison of baseline and final WRs. Appropriate non-parametric tests were used, with Mann-Whitney U and Wilcoxon Sign Rank tests for inter- and intra-group analysis, respectively. Chi-square test was used to compare categorical data.

Secondary post-hoc analyses were performed to assess potentially confounding factors. To investigate differences in management of different complications, failure rates were stratified by scenario and differences in means assessed using ANOVA. A linear regression model was used to adjust for trainee factors (gender and years of experience) when interpreting the effect of trial group (control vs. intervention) on failure rates.

Questionnaire responses were collated and reported as mean and standard deviation. Responses given as a 4 or 5 on a Likert scale of 1 to 5 were considered positive.

All data was entered into an anonymised table, with analysis was performed using SPSS Statistics version 21 (IBM Corp, Armonk, NY). P values of less than 0.05 were considered statistically significant. All results are reported as medians and interquartile ranges unless otherwise indicated.

### 6.2.3 Results

#### 6.2.3.1 Study participants
Twenty trainees were recruited; each completed two WRs of three patients each, resulting in a total 120 patient assessments included in the final analysis. Each trainee completed all six patient scenarios, with equal randomisation to control and intervention groups. There were no differences in gender or experience levels of participants between groups (table 6.4).

6.2.3.2 Ward round performance

Recorded videos of 20% of all ward rounds were assessed by a second blinded rater, with excellent inter-rater reliability, Cronbach’s alpha = 0.893.

Considering the rates of failure to complete critical processes in management of postoperative complications (figure 6.3), there was no difference between groups for baseline assessment [control 58(37)% vs. intervention group 67(19)%], p=0.988. For the final WR, process failure rates were significantly different, with near-elimination of errors in the intervention group [60(35)% vs. 0(0)%], p<0.001. All residents in the intervention group elected to use the optional checklists for the final WR.
### Table 6.4 Subject demographics

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Intervention</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>10</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>M:F ratio</td>
<td>9:1</td>
<td>7:3</td>
<td>0.264</td>
</tr>
<tr>
<td>Years of surgical training</td>
<td>3 (2)</td>
<td>3 (1.5)</td>
<td>0.796</td>
</tr>
</tbody>
</table>

Years of training reported as median (IQR).
Figure 6.3 Failures to adhere to critical processes in the management of postoperative complications. Boxplot comparisons of performance between control and intervention groups for baseline (left) and final (right) ward rounds. Circles denote outliers.
There were no statistically significant differences between groups at baseline for SWAT-A, SWAT-M, or W-NOTECHS metrics (table 6.5). In the final WR, the intervention group scored significantly higher than the control group in patient management, SWAT-M [control 0.63(0.25) vs. intervention 0.83(0.25), p=0.005], though not for patient assessment (SWAT-A) [p=0.094] or non-technical skill (W-NOTECHS) [p=0.218]. There was no significant difference in the time taken for each WR between groups or between WRs.

Considering intra-group comparisons, a small but significant improvement was seen in both groups for SWAT-A (both p < 0.001). However for patient management (SWAT-M), improvement was only seen in the intervention group (p < 0.001) and not the control group (p = 0.571). Similarly, total W-NOTECHS scores improved only in the intervention group (p = 0.043, vs. control group p = 0.809). There were no significant differences in means for individual W-NOTECHS domains scores. There were significantly fewer prescription errors in the intervention group (3 vs. 14, Chi-square p = 0.046). These related, for example, to the prescription of penicillins in patients with documented penicillin allergy and failure to appropriately reduce medication dosages in patients with acute renal failure.

There were no differences in performance between scenarios (table 6.6), p = 0.796.

Controlling for resident experience and gender through inclusion as independent variables in a logistic regression model had no effect on results, with trial group allocation the only significant predictor of process failure rates (OR (95% CI): 0.537 (0.457, 0.617), p < 0.001).

### 6.2.3.3 Questionnaire responses

All (10/10, 100%) intervention group study participants returned completed questionnaires. Responses were very positive, with 100% of respondents believing the checklists were easy to use, and would improve clinical practice (table 6.7). All indicated that they would want the checklists used if they themselves were to experience a postoperative complication.
### Table 6.5 Simulated ward round performance metrics

<table>
<thead>
<tr>
<th></th>
<th>Baseline assessment</th>
<th>Final assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Control</td>
<td>Intervention</td>
</tr>
<tr>
<td>Simulated patients, n</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>SWAT-A</td>
<td>0.62 (0.24)</td>
<td>0.73 (0.18)</td>
</tr>
<tr>
<td>Intragroup (baseline vs. final), p-value:</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SWAT-M</td>
<td>0.63 (0.25)</td>
<td>0.56 (0.21)</td>
</tr>
<tr>
<td>Intragroup (baseline vs. final), p-value:</td>
<td>0.571</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>W-NOTECHS</td>
<td>19.5 (5.0)</td>
<td>17.0 (4.5)</td>
</tr>
<tr>
<td>Intragroup (baseline vs. final), p-value:</td>
<td>0.809</td>
<td>0.043</td>
</tr>
<tr>
<td>WR time (mins)</td>
<td>26.5 ± 5.1</td>
<td>26.8 ± 5.1</td>
</tr>
<tr>
<td>Intragroup (baseline vs. final), p-value:</td>
<td>0.349</td>
<td>0.111</td>
</tr>
<tr>
<td>Prescription errors</td>
<td>10</td>
<td>22</td>
</tr>
<tr>
<td>Intragroup (baseline vs. final), p-value:</td>
<td>0.298</td>
<td>0.020</td>
</tr>
</tbody>
</table>

WR time reported as mean ± SD. All other figures median (IQR).
Table 6.6 Baseline WR performance stratified by scenario. ANOVA p = 0.796.

<table>
<thead>
<tr>
<th>Clinical scenario</th>
<th>Failure rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pneumonia</td>
<td>65 ± 26%</td>
</tr>
<tr>
<td>Wound infection</td>
<td>58 ± 13%</td>
</tr>
<tr>
<td>Anastomotic leak</td>
<td>61 ± 13%</td>
</tr>
<tr>
<td>Sepsis</td>
<td>61 ± 13%</td>
</tr>
<tr>
<td>Urinary tract infection</td>
<td>70 ± 33%</td>
</tr>
<tr>
<td>Postoperative haemorrhage</td>
<td>68 ± 18%</td>
</tr>
</tbody>
</table>

Failure rate reported as mean ± SD.
Table 6.7 Checklist users’ responses to use of checklists for management of postoperative complications. Responses received from 10/10 intervention group subjects. Possible scores range from 1 (strongly disagree) to 5 (strongly agree) on a Likert scale.

<table>
<thead>
<tr>
<th>Questionnaire statement</th>
<th>Response score mean (range)</th>
<th>Positive response (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The checklists were easy to use</td>
<td>4.7 (4-5)</td>
<td>100%</td>
</tr>
<tr>
<td>The checklists cover clinical conditions relevant to my practice</td>
<td>4.8 (4-5)</td>
<td>100%</td>
</tr>
<tr>
<td>I found this checklist applicable to my practice</td>
<td>4.1 (3-5)</td>
<td>80%</td>
</tr>
<tr>
<td>The checklists may improve management of complications</td>
<td>4.4 (4-5)</td>
<td>100%</td>
</tr>
<tr>
<td>I would want this checklist to be used for me if I experienced a postoperative complication</td>
<td>4.3 (4-5)</td>
<td>100%</td>
</tr>
<tr>
<td>I would consider using the checklists in my routine practice</td>
<td>4.1 (3-5)</td>
<td>80%</td>
</tr>
</tbody>
</table>
6.2.4 Discussion

To date, the implementation of checklists has been limited to the domain of routine surgical care. However, it is patients who suffer complications in the postoperative phase who are most at risk of adverse events, and most vulnerable to variations in their care. In the management of postoperative morbidity, failures to adhere to principles of best practice place unwell patients at risk of avoidable harm and poorer outcomes. As such, adherence to care process-oriented guidelines, such as expert body recommendations or antimicrobial stewardship policies, may contribute significantly to the standardisation of care and optimisation of patient management. The development and implementation of POSCH checklists, as seen in this study, represents a means to achieve this.

The results of this study demonstrate again the significant variability which exists in the management of postoperative complications, despite their common nature. Crucially, it also demonstrates how this may be essentially eliminated through a simple checklist-based intervention. In this study, the introduction of checklists to guide patient management, once the diagnosis of a postoperative complication had been made and a decision to treat taken, led to the standardisation of treatment in line with evidence-based guidelines, and reduction of median error rates from 60% to 0%. Despite their simple nature, or perhaps because of it, implementing checklists in the manner of our study has the potential to greatly increase the standardisation of ward-based care, and may act to significantly reduce failure to rescue rates and improve outcomes.

The experimental, simulated context in which they were this study was conducted represents a potential confounder in the interpretation of its results. However, the overwhelmingly positive effect achieved in this study through the introduction of checklists is in agreement with previously published data. Pronovost and colleagues’ landmark study saw the introduction of checklist-based protocols for antiseptic central venous catheter
placement eliminate catheter-associated sepsis altogether over the course of 3 months (with a reduction of infection rates from 2.7 to 0 per 1000 catheter-days). Similarly, Arriaga et al’s previously described simulation-based trial assessing the management of operating room crises reported a reduction of error rates by 75% when checklists were used to guide the operating room team’s response to rare events such as asystole, malignant hyperthermia, or unstable arrhythmias. Such data attests to the disproportionate impact on clinical care which may be achieved through such simple, cost-effective interventions.

Patient management without the aid of checklists placed patients at risk of critical errors such as failure to prescribe antibiotics for a patient with an anastomotic leak, insufficient intravenous access in a patient suffering from haemorrhage, or failure to obtain blood cultures in a patient with pyrexia and sepsis. The use of checklists saw a near elimination of these errors, standardising care without compromising clinicians’ decision-making or patient care. It is important to note that the checklists were introduced as aides to guide care only, with clinicians free to use them (or not) as they felt appropriate, dependent upon the suspected diagnosis and intended treatment. Though they represented an additional piece of paperwork, the checklists were well received, with all reporting that they were easy to use, and improved their practice. Checklist use did not prolong the WRs, with no significant difference in time taken between groups. Perhaps most significantly, all residents reported that they would want checklists used in their own care, were they to require care for postoperative complications themselves.

Introduction of the checklist also led to improvements in non-technical skill. The ability of checklists to flatten clinical hierarchies and improve communication has been previously seen in improved safety attitudes and staff questionnaire responses with the use of the WHO Surgical Safety Checklist, but this is the first time that the effect of a checklist on non-technical performance has been directly observed and measured in a clinical context.
The positive feedback seen in this study represents a crucial aspect of checklist implementation, and illustrates the potential benefits of simulation-based trialing of such interventions prior to introduction into clinical practice. Centres have previously reported significant difficulties associated with attempts to change established clinical process.\textsuperscript{366-368} By incorporating local staff (who were not involved in this study or aware of its content) into the development process,\textsuperscript{365} we were able to increase the likelihood of staff buy-in with the checklists. Any negative feedback would have been captured in the simulation and could have been addressed prior to clinical implementation.

The results of this study, intended as an exploratory study of the effectiveness of checklists for postoperative care, must be considered in the context of its limitations. First, this study assessed the performance of surgical residents only. Other health professionals including nursing staff and interns were not included, as their recruitment was not feasible for this trial. However, as it is residents who are commonly responsible for routine patient review, and are most likely to make the first diagnosis of a postoperative complication and initiate treatment, this is unlikely to represent a major confounder. Interaction with other members of the clinical team, though important, would be secondary to clinician’s diagnosis and decision to treat. Second, this was a single centre study of a simulation-based trial.

Localization and adaptation to local protocols and practices is a critical component of successful checklist implementation,\textsuperscript{36,373} which would require adapted checklists for implementing centres. Simulation has become a thoroughly accepted environment in which to assess clinical processes, wherein standardised scenarios may be controlled for and patient risk eliminated.\textsuperscript{374} Use of a high-fidelity simulated ward environment has been shown by previous studies presented in this thesis to reflect the realities and variation seen in clinical practice.\textsuperscript{322,370} The present study design also controlled for any potential learning effects of the simulator by including repeated WR assessments for both intervention and control groups. Future pragmatic trials should involve all caregivers and implementation of
checklists in the multiple clinical environments. Third, this study does not evaluate the longevity or retention of changes to practice, which may experience degradation of checklist use over time. Other groups implementing checklist-type interventions have variably reported both deterioration\textsuperscript{375} and improvement\textsuperscript{84} of performance over time. Analogous to the mastery of technical skill through repeated training,\textsuperscript{233} it should be unsurprising that repeated training sessions may need to be considered for such non-technical interventions also. This could be investigated through further longitudinal assessments in future.

6.2.5 Conclusion

This study demonstrates the effectiveness of checklist-driven care to standardise treatment according to established best-practice standards and guidelines. With an increasing emphasis on safety and accountability in surgical culture, postoperative checklists represent a simple, but important, treatment aid which is easy to use, accepted by clinicians, and amenable to audit and governance. Previous research has assessed checklists for routine care\textsuperscript{94} or rare crises.\textsuperscript{143} Considering the common nature of complications, however, improving care of patients with postoperative morbidity has perhaps the greatest potential for impact upon reduce failure to rescue rates, and may act to standardize treatment, and improve outcomes.
Chapter 7

Conclusion
7.1 Background synopsis

The pursuit of improved patient outcomes and avoidance of harm is the basic tenet which drives all surgical research. Avedis Donabedian’s belief in quality control through the reporting of relevant outcomes has developed into the burgeoning field of medical outcomes research that exists today, with efforts to define optimal outcomes for surgical care quality continuing unabated.

The evolution of outcomes research has been in part related to the continued evolution of surgical practice itself. Whereas coarse mortality figures were in the past seen as the primary indicators of quality of care, improvements in medical and surgical care have meant that mortality rates have been significantly reduced over time. Procedures such as elective colectomy, for example, once considered a major surgical procedure with significant associated morbidity and mortality have been transformed through the introduction of laparoscopy, peri-operative checklists, and enhanced recovery protocols. Where patients were once routinely expected to remain in hospital for up to 10 days following surgery, they are now being discharged after 48 hours,\textsuperscript{92} with some surgeons even advocating cases to be performed as day-case procedures in future.\textsuperscript{376} In such context, mortality loses its significance as a quality indicator, and has been replaced by more refined measures such as the incidence rates of adverse events, morbidity, or failure to rescue.

The consideration of these outcomes, in particular, has brought previously overlooked realms of care to the forefront of quality and safety research. Whereas traditionally, major interventions to improve surgical care were targeted at the operating theatre and the surgical procedure itself, analyses of adverse events have highlighted the dangers of the postoperative phase of care instead, wherein the majority of events and errors appear to take place on the ward, rather than in the operating theatre.
The growing body of evidence surrounding failure to rescue appears to confirm these findings. Surgery, it suggests, is relatively comparable across modern medical institutions. Adjusting for patient-related variables, patients can generally be expected to survive a given procedure, and stand and equal chance of developing postoperative complications regardless of in which hospital they receive care. In patients developing complications – the risk of which, for some complex procedures, approaches 50% – differences in care provided by hospitals are revealed, with some hospitals more than twice as likely to “rescue” patients than others.

At this point, it is prudent to again recall Donabedian’s paradigm of process, structure, and outcome. Beyond the study of outcomes alone, Donabedian emphasised the need to consider the links to structures and processes which may influence the outcome in question. The quantification of the outcome of care quality must be followed by structure and process interventions to improve it, or it will have achieved little.

Analyses of failure to rescue data, as reviewed in the first chapter of this thesis, has suggested numerous factors, both structural and process-related, associated with differences in failure to rescue rates. This has presented two important points. First, that changes in care process appear broadly superior to structural interventions in terms of their practicality, when considering cost, implementability, scale, and potential impact on outcomes. Second, that care processes relating to the detection and management of complications have not been subject to empirical consideration until now.

The surgical ward round represents the focal process of clinician-patient interaction in the pre- and postoperative phase of care. It is during this daily process that patients are assessed, their progress summarised, new diagnoses made, and a management plan set out. In the abrupt transition from junior trainee with minimal responsibility, to becoming the registrar charged with the responsibility of the daily round, trainees have historically relied
on their own judgment, and the Halstedian paradigm of “see one, do one, teach one.”

Despite growing recognition of the importance of WRs, there is a profound dearth of scientific evidence on the topic. At the point of commencement of this thesis, WRs lacked the tools for formalised assessment, and were rarely the subject for formal training.

One possible result of this, it might be reasoned, is the variability in the management of postoperative morbidity implied by variations in FTR rates. Additionally, process interventions such ERP in colorectal surgery suggest that the improvement and standardisation of postoperative care may result in not only better management of complications, but their prevention altogether. This may not only improve short-term outcomes, but long-term survival as well, as suggested by the meta-analysis presented in chapter 1 of this thesis.

In light of the growing burden of evidence that poor WRs result in poor outcomes, it is clear that this can no longer be considered acceptable, for trainees and patients alike. The aim of this thesis was to develop a framework for the assessment, training, and improvement of the surgical ward round. Though focusing on the domain of general (gastrointestinal) surgery, it is anticipated that these largely generic principles could be in future applied to any specialty, both medical and surgical.
7.2 Summary of findings

The surgical WR has not previously been subjected to empiric examination, nor tools for its improvement subjected to appropriate processes of development and validation. This thesis represents the first body of work to consider the means to assess, quantify, and improve upon the WR. As such, it was necessary to first broadly explore the concept of the surgical WR, both to identify potential markers of WR quality upon which to base subsequent studies, but also to ensure that any quality markers used would be valid, with support from key stakeholders in the surgical WR.

A semi-structured interview of equal numbers of patients, nurses, house officers, registrars, and consultants was conducted. Such a study design allowed both directed questioning to address pre-defined issues such as the purpose of the WR, quality markers, and potential improvement measures, but also unstructured exploration of related issues and interviewees’ opinions. By sampling a broad range of professions, people, and hospitals, a representative sample of interviewees was achieved.

The results of the interview study (chapter 2) confirmed that variations in WR quality were broadly perceived to occur, particularly by clinical staff, placing patients at risk of avoidable harm and injury on a routine basis. The lack of structured training for WRs was recognised, with potential quality markers for the assessment of WRs identified, and potential means of improvement recommended. The development of simulation-based curricula, analogous to successful training programmes which have been implemented for technical skills training in surgery, was the most frequently named intervention which staff and patients alike named as being potentially effective for this application.

The development and validation of a simulated ward was described. Though exploratory studies of simulated ward environments have been previously described, a thorough consideration and description of factors to consider in the establishment of a ward simulator.
has not been previously undertaken. A detailed framework for establishing a ward simulator, including the consideration of organisational and resource-based requirements, was described. Novel tools for the assessment of ward-based care, the SWAT and W-NOTECHS scores, were developed. SWAT, a checklist-based scoring system incorporating quality markers described in the exploratory interview study (content validity), and the adapted W-NOTECHS system for non-technical performance assessment, were scientifically validated within the simulated ward. The ability of the SWAT and W-NOTECHS to correctly and reliably distinguish between the performance levels of trainees of varying experience demonstrated construct validity of the system, with extremely positive questionnaire feedback further supporting the face validity (i.e. fidelity) of the simulator.

The next study, an observation and assessment of surgical WRs for critically unwell HDU patients in the clinical environment, resulted in several important, and novel, findings. Further evidence for the validity of SWAT and W-NOTECHS was demonstrated, with their application in the clinical setting shown to be feasible and unobtrusive, and distinguishing SWAT further as not only a first validated score for WR assessment, but one of very few scores for medical assessment of any kind to be able to demonstrate relevant association to clinical outcomes. The assessment of sequential WRs in this study was also the first example of prospective, observational data for this critical process of surgical care. Though the large body of research on the topic of failure to rescue has implied variations in postoperative care, this has been confined to retrospective and database-driven study to date. The results reported in this thesis confirm, using prospective data for the first time, the degree of variability present in routine surgical care and the surgical WR. Crucially, these also demonstrated the consequences of this variation. Poor surgical WRs, this study has shown, result not only in less thorough patient assessment, but greatly increase the risk of actual patient harm in the form of preventable complications such as pneumonia, wound infections,
or urinary sepsis, setting out a powerful argument for the need to improve WRs and reduce preventable complications.

The reduction of preventable complications, and improvement of surgical outcomes, forms the remit of many national healthcare quality organisations such as the AHRQ in the US, or CQC in the UK. However, despite reporting adverse event rates, the prescription of the means by which to reduce them is sometimes lacking. Considering the association between WR quality and preventable complications, coupled with the lack of current training, this highlights WRs as a critical area of care to be addressed if patient outcomes are to be improved. This call has been echoed in recent the recent joint statement issued by the Royal College of Physicians and Royal College of Nursing – *Ward Rounds in Medicine: Principles for Best Practice* – which, though calling for renewed focus on the improvement of WRs, lacks guidance on how this is to be achieved.

This thesis has presented the development and implementation of an evidence- and simulation-based curriculum which addresses precisely this training gap. Driven in particular by the training needs presented by the introduction of laparoscopic surgery over the past 20 years, extensive work on the development and implementation of simulation-based surgical curricula has been previously undertaken for technical skills in surgery. Research has repeatedly demonstrated the benefits to patient safety, trainee proficiency, and cost to health systems, which may be achieved through advances in educational techniques. Building upon this body of prior work, an evidence-based curriculum utilising the most up-to-date teaching and assessment methods for the conduct of surgical WRs was developed and implemented.

Integrating seamlessly into existing training schemes for junior (core surgical, CT1-2) trainees, completion of the curriculum resulted in significant improvements in both technical and non-technical skills in the conduct of WRs by trainees. Vitally important, also, is the positive
feedback received by trainees completing the session. The enthusiastic participation of trainees is crucial to such a programme’s success. By reinforcing concepts of structured and thorough WRs in the simulated ward, this may translate to improved WRs in the clinical environment, with earlier detection of complications and improvement of outcomes.

Finally, the simulation curriculum aims to improve the WR structure and patient assessment. One consequence of this may be for clinicians to detect complications earlier as a result of the more structured approach, and fewer omissions in the assessment of the patient.

However, though earlier diagnosis allows the earlier initiation of care processes for the management of complications, if this care is not appropriate then the benefit to the patient is likely to be minimal. This was demonstrated in the casenote review conducted as part of the development process for the POSCH checklists. Here, the management of common complications was compared to guidance by expert bodies such as the British Thoracic Society guidelines for management of pneumonia, or the National Institute for Health and Clinical Excellence guidance for sepsis, as well as best evidence. The results of even this limited casenote review demonstrated a striking variability in care, even in a tertiary academic centre rated as one of the top 5 hospital trusts in England. Patients whose care demonstrated poorest compliance with guidelines were also at significant risk of developing further complications, after adjustment for factors such as gender, age, and ASA grade.

Thus, checklists may play a clear role in the improvement of care of postoperative morbidity. However, the implementation of expert guidelines alone is not enough. Centres attempting to implement process change without first considering local factors and culture have encountered significant resistance. Without staff buy-in or appropriate training, checklists are vulnerable to “box-ticking” — that is, for staff to view them as an administrative exercise rather than an opportunity to change practice. In such situations, outcomes will remain unchanged even if checklists are alleged to have been completed.
For this reason, a structured developmental process was undertaken for POSCH checklists. This considered only the most common complications, to maximise the checklists’ effectiveness, and took local practice into account by consulting local guidelines, experts, and end-users. This resulted in a final iteration of checklists felt to be feasible, practical, and supported by consultants, staff, and trainees alike.

The implementation of the POSCH checklists was assessed in a simulated environment. The use of the simulated ward for this trial, as opposed to in situ clinical testing, offered a number of advantages. By avoiding the use of real patients, their safety could be assured. Furthermore, this allowed the use of standardised, reproducible clinical scenarios, with which the effect of checklists on clinician behaviour could be assessed in a controlled environment. The simulated ward also offers a highly realistic setting, but is equipped with integrated recording systems (which trial participants were not aware of) to allow better scientific assessment of performance. Finally, by trialing the checklists first, staff responses and buy-in could be assessed, with any necessary changes undertaken, before introducing them to clinical practice.

This study demonstrated a large, significant improvement in adherence to best practice, when checklists were used. Patient care was improved, errors were reduced, and overall WR performance increased. Perhaps more importantly, there was no significant negative feedback, with no difference in WR duration, and participants expressed a desire to use checklists in their practice. With the six checklists developed covering the six postoperative complications reported to make up over 90% of all morbidity, POSCH checklists may thus greatly improve patient care in such situations.
7.3 Impact of this work

Fully implemented, the work presented in this thesis has the potential to significantly impact upon training and clinical practice in the domain of postoperative patient care. The pursuit of valid, relevant, and practical means of assessment for trainees continues to pose difficult questions for educators and surgical education researchers alike. Throughout the simulation-based studies conducted in pursuit of this degree, many trainees repeatedly expressed their opinion that a simulation-based ward round would be an effective means of assessment.

With the advancement of technical skills assessment, both virtual reality and benchtop simulator models, with corresponding assessment frameworks, have been gradually incorporated into many trainee assessment and selection processes in the UK. A similar model could be pursued to incorporate simulation-based assessments for ward-based clinical practice, which would provide greater fidelity and validity than existing paper-based workplace assessments. One potential model might, for example, see trainees perform a simulated WR, with trainees’ SWAT and W-NOTECHS performance assessed against their peers and given as a decile or percentage score (summative feedback), complemented by subjective assessor comments (formative).

The work presented in this thesis has similar potential to address currently overlooked areas of clinical practice. National recommendations for best practice in WRs, for example, would be greatly strengthened if backed by specific guidance and evidence base such as presented here – just as the RCP/RCN have issued their recommendations for WR practice in medicine, so too might a guidance document for WR practice in surgery be considered.

Finally, institution of localised POSCH checklists to manage surgical (as well as in future potentially other specialties’) complications will reduce variation in care and improve patient outcomes. Ideally, such implementation would require a centralised body, such as the
National Institute for Health and Care Excellence (NICE), an NHS England special health authority whose remit includes the maintenance of evidence-based care guidelines, to maintain and continue to expand and develop checklists for postoperative surgical complications. Implementation and localisation, in turn, should be the responsibility of local NHS Trusts and their internal committees (in accordance with the principles outlined in Chapter 6). A similar structure exists, for example, for antibiotic prescribing, wherein guidelines on the indications for antibiotic treatment in respiratory tract infection\textsuperscript{378} or surgical prophylaxis\textsuperscript{379} have been published by NICE and SIGN (the Scottish Intercollegiate Guidelines Network), but the exact type and duration of antimicrobial therapy is selected by Trusts themselves, based upon local microbial prevalence and resistance patterns.
7.4 Limitations

Several limitations to this work may be considered, and should be taken into account in the interpretation of its findings.

The fact that many of the studies were simulation-based limits the generalisability of the findings. Ultimately, without clinical outcomes with which to correlate the interventions proposed, the magnitude of their impact upon actual patient care cannot be assured.

Despite this, the simulated environment offered a number of advantages over *in vivo* study which made it desirable, or even necessary, to conduct the initial exploratory studies of this thesis in it. The reproducible and standardised nature of the scenarios made the direct comparison of performance between trainees possible. This would not have been possible in a clinical setting, or would have required unfeasibly large numbers of patients to allow adjustment for patient-related and other confounding factors. Additionally, the audiovisual capabilities of the simulator suite allowed for recording of the trial participants, allowing for unobtrusive and blinded rating of the subjects by a second rater to ensure reliability of the scoring system used, and reduce the risk of rater bias.

Ideally, simulation-based pilot studies such as those presented in this thesis will be followed up in future with ones that utilise clinical endpoints instead. This was in part achieved in chapter 4, wherein the use of the SWAT score, and the demonstration of its association with the risk of preventable complications, provides significant evidence for its validity. To assess similar outcomes for the simulation-based WR curriculum, or POSCH checklists, would be the next step.

The interventions developed to improve ward-based care – the WR curriculum and POSCH checklists for management of postoperative morbidity – were assessed in WRs which included only the lead clinician as the trial subject. Though accompanied by a nurse and a house officer throughout, these were confederates. In reality, of course, the WR
incorporates all of these multidisciplinary members (and more, e.g. dieticians, pharmacists, physiotherapists) as variable factors affecting WR quality. Indeed, the inclusion of nursing staff and other disciplines is one of the primary recommendations of the RCP / RCN ward round statement. However, it is worth noting that the primacy of the physician’s role is also mentioned by the same document, and was implied in the results of the interview study conducted (chapter 2). As such, it was felt valid to control for as many factors as possible in initial studies by standardising the behaviour of the rest of the WR team, to isolate clinician behaviour as the variable of interest in the studies described. With the positive results gleaned from this thesis, future studies may follow in the path of training sessions for operative teams, in including all multidisciplinary members and stakeholders in educational interventions.

By standardising other variables within the simulated environment, this thesis specifically sought to assess variability in clinician knowledge and behaviour as the determining factor of WR quality. Other factors, such as workload (i.e. number of patients to see, or time available in which to see them), or environmental stressors (ease of access to results, noise levels), may also play a significant role, though these are likely to be secondary to the ability of the lead clinician.

Finally, the nature of this thesis has meant that all studies were conducted within a single UK tertiary academic centre, potentially affecting the applicability of findings to other centres, systems, or countries. Though much of the bias, if present, is likely to have been positive (i.e. smaller community or non-academic centres may have, in anything, greater levels of error and variability to address with these interventions), further external validation of this work is required.
7.5 Future work

Though the results presented by this thesis are encouraging, many questions in the improvement of postoperative care remain, which future work in this area will seek to address.

As described in the previous section (chapter 7.3), this thesis is not without its limitations, which future studies should aim to remedy. Further external validation of the simulation-based WR curriculum would be desirable, as would correlation with clinical outcomes. Initially, this might include qualitative or questionnaire-based assessment, e.g. qualitative assessment of trainees’ clinical WR performance by educational supervisors, but will ideally graduate to using clinical outcomes as well.

The domain of technical skills assessment is an example of how a field may continue to evolve. The Objective Surgical Assessment of Technical Skill (OSATS), for example, is one of the best validated and widely used rating scales. Recently, however, it has been in some areas supplanted by newer scales that have often included subtle, yet significant refinements of the original OSATS, such as the Global Operative Assessment of Laparoscopic Skill (GOALS) or the Operative Procedural Rating Scale (OPRS), owing to their procedure-specific nature and therefore potentially greater validity. Since the publication of the SWAT scale for surgical WR assessment, already a further (as yet unvalidated) rating scale has been proposed. Reflecting the growing interest in this area of surgical care, other proposals may follow. The comparison, revision, and refinement of these scales will be required to ensure accurate and objective measurement of WR performance.

Similarly, the simulation-based curriculum for WRs is likely to undergo iterative change in response to growing understanding of the field. For the present, perhaps the greatest challenge is to ensure its effective implementation. Too many educational interventions undergo a process of development and validation, only to fail to be disseminated into
practice. Future work must seek integration with existing surgical training curricula, and consider how to best serve trainees – determining the necessary number and frequency of sessions to improve WR performance, retention of learning, and clinical outcomes.

As has been the case for technical skills, varied curricula with different learning objectives for trainees of different ability levels should be developed for WRs and ward-based care. The described curriculum (chapter 5) is suitable for junior trainees, but likely less relevant to mid- or late-stage registrars. For these trainees, the incorporation of more complex scenarios may be appropriate, but should also include the participation of the entire WR team (i.e. nursing staff, in particular) to improve team performance on the ward.

POSCH checklists have been developed, optimised, and validated within a simulated environment. Clinical implementation, likely within the context of a pre-/post-intervention study, would be the next step. This should first incorporate initial clinical pilot testing, with further assessment of user feedback. The importance of appropriate end-user training and buy-in cannot be understated. Analogous to the implementation of the WHO Safer Surgery Checklist, this may be facilitated through identification of local champions and top-down leadership, with POSCH checklists then trialed in a large-scale clinical trial, involving either cluster-randomisation or pre- / post-assessment of multiple centres. As seen in the simulation study (chapter 6), it is anticipated that POSCH checklists may help reduce treatment delays, and improve outcomes, but further study is required to demonstrate this.

It should also be recognised that the surgical ward round, whilst certainly the most prominent, is by far not the only interaction between clinical staff and patients. Throughout a patient’s journey they are likely to experience hundreds of other clinical encounters with junior clinicians, nurses, and other allied health professionals, which are likely to play an equally important role in outcomes, but have yet to be identified. Donchin et al’s previously described observational study of patients on the intensive care unit, for example, highlights
the number of clinical events which are sometimes taken for granted, but which also place patients at risk of error and adverse events. Such a study conducted on the ward might provide novel insights to clinical care, and new leads to improve it.

To further optimise the ward-based care environment, the consideration of environmental factors which might augment or hinder the WR and clinical care in general should be considered. Within the operating theatre, studies have reported the negative effects of noise,\textsuperscript{383} interruptions,\textsuperscript{384} and poor ergonomic design\textsuperscript{385} on outcomes. It stands to reason that the same factors may play a role on the ward. The identification and optimisation of these factors may further contribute to improvements in care.

As ward simulators increase in number and use, the next steps to promote their appropriate and effective use in future should be considered. This should include the formation of multi-centre working and research groups, to guide development and use of simulated wards as well as a centralised database of validated scenarios and assessment tools.
7.6 Concluding remarks

The variability of postoperative care, and its effect on surgical patient outcomes, has been the subject of great interest in recent years. Whereas existing research has been largely descriptive in this respect, this thesis has sought instead to develop prescriptive measures by which care might be improved.

The WR has long been the traditional focus of postoperative surgical care, without being subjected to formalised assessment or training. This thesis has developed a number of assessment and training tools with which to address this gap, and recommended the simulated ward as the ideal environment in which training and assessment may take place.

For too long, ward-based care has lagged behind technical skills as a focus for educational development and skills improvement. Whereas trainees must now practice on benchtop, virtual reality, and cadaver models (to name but a few), and must demonstrate proficiency before being allowed to operate on patients, the conduct of WRs remains a case of skill acquired through practice. It is not enough to see one, in order to do one, much less teach one. The tools have now been developed with which to assess, improve, and standardise critical structures and care processes in the assessment and management of the post-operative surgical patient. Future implementation of these and integration into surgical curricula will benefit clinician training, patient care, and surgical outcomes alike.
List of References


Appendix
### 1. Demographics, introduction

“*We are conducting research looking at measuring and improving surgical care outside of the operating room. The daily ward round of course plays a very important role in this area of patient care.*

*The goal of this is to interview study is*
- first, to define the WR and general skills trainees need for this
- second, to come up with ways in which we might measure the quality of ward rounds
- finally, moving on from that, to think of ways to improve them."

*Can I just start by confirming that you are a patient / what your role is?*

(How long have you been a consultant/trainee?)
(How long have you been in hospital for?)

*And can I just confirm that you are aware that we are being recorded and that you are happy for this?*

---

### 2. Identifying the problem

*I would like to start by asking about surgical ward rounds, in general. In your experience, do you think there is variability in way in which different clinicians conduct a ward round, and do you think this impacts patient care?*

*How do you think that variability manifests itself / can you give specific examples?*

*Do you think this impacts patient care?*

---

### 3. Define WR process and sub-processes

*What is your personal definition, if I were to ask you what a WR is, and what happens during a WR?*

Prompts: “*What about...*”
- Patient assessment?
- Management plan?
- Communication with patient?
- Communication with team / nursing staff?

---

### 4. Define required skill set

*In order to do all this, what do you think are some of the skills a clinician needs to have to be able to do all this?*

Prompt:
- “*I’m thinking here of both clinical or technical skills, but also non-technical qualities, like communication skills or teamworking?*”

Prompts: “*What about...*”
- Teamworking?
| 5. Identify quality markers | “As I mentioned at the start, part of the aim of this interview is to try and identify quality markers for WRs. Imagine it’s Saturday morning and you are meeting a patient on HDU for the first time, he was operated on yesterday, and he looks a bit unwell. What are you going to do?”

Prompts: “What about individual items, are any or all of them particularly critical...”
- Checking drug chart?
- Observation chart?
- Assessing the patient?
- Reading the medical notes?
- Checking the operation note?

We’ve talked about assessing the patient. What about the management aspect of things? What does the team need to know about how to manage the patient for the rest of the day as a result of the WR?

Prompts: “What about...”
- Dietary status
- Further investigations
- Treatment / antibiotics |

| 6. How to improve practice | “As we said at the beginning, there is a great amount of variability which occurs. How do you think we might address this to improve individual clinicians’ practice?

ERP in colorectal, or goals sheets in ITU, have suggested that proformas or checklists can be effective – do you think they might play a role here?

Implementing any kind of change can be difficult without first appropriately training staff. How would you go about assessing and training clinicians for this kind of thing?” |

| 7. Final points? | Thank you for your time. Just before we finish, do you think there is anything else you would like to add on the topic? |
Appendix 2. SWAT assessment tool

SimWard Simulation Course

Surgical Ward care Assessment Tool (SWAT)

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient #</td>
<td>Patient #</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Patient verbal</td>
<td>Analgesia</td>
</tr>
<tr>
<td>Nurse verbal</td>
<td>IV fluids</td>
</tr>
<tr>
<td>Notes review</td>
<td>Antibiotics</td>
</tr>
<tr>
<td>Obs Chart</td>
<td>Dietary status</td>
</tr>
<tr>
<td>Drug chart</td>
<td>Bloods</td>
</tr>
<tr>
<td>Abdo exam</td>
<td>New imaging</td>
</tr>
<tr>
<td>Chest exam</td>
<td>Consider cultures</td>
</tr>
<tr>
<td>VTEP checked</td>
<td>Discussed with team</td>
</tr>
<tr>
<td>Wound</td>
<td>Discussed with patient</td>
</tr>
<tr>
<td>Imaging</td>
<td>Blood results</td>
</tr>
<tr>
<td>Diet status</td>
<td>Diet status</td>
</tr>
</tbody>
</table>

Notes / errors / risk events:
e.g. not NBM if for theatre,
inappropriate diagnosis or delay,
does not consider alternate antiepileptic

Each task either completed, not completed, or not applicable.

% assessment tasks complete:

% management tasks complete:
### Appendix 3. W-NOTECHS assessment tool

<table>
<thead>
<tr>
<th>Date:</th>
<th>Clinical team:</th>
</tr>
</thead>
</table>

#### W-NOTECHS

<table>
<thead>
<tr>
<th>Leadership</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team leader not clear</td>
<td>Team leader defined but does not fulfil all functions</td>
<td>Team lead clear at all times</td>
</tr>
<tr>
<td>Team lead does procedures that should be delegated</td>
<td>Team lead does procedures that should be delegated</td>
<td>Debriefs team after round</td>
</tr>
<tr>
<td>Transition between leaders unclear</td>
<td>Accepts input from team, facilitates problem solving</td>
<td>Accepts input from team, facilitates problem solving</td>
</tr>
<tr>
<td></td>
<td>Clarifies team roles as needed</td>
<td>Clarifies team roles as needed</td>
</tr>
<tr>
<td></td>
<td>Excellent time management – efficient and adheres to best practice</td>
<td>Excellent time management – efficient and adheres to best practice</td>
</tr>
</tbody>
</table>

#### Cooperation and Resource Management

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role of most team members unclear</td>
<td>Role of all members not clear or do not fulfil role</td>
</tr>
<tr>
<td>Some members idle or overworked most of the time</td>
<td>Some members idle some of the time</td>
</tr>
<tr>
<td>All members clearly identified and fulfill role</td>
<td>All members clearly identified and fulfill role</td>
</tr>
<tr>
<td>Members ask for help if cannot complete task or order</td>
<td>Members ask for help if cannot complete task or order</td>
</tr>
<tr>
<td>Shift of responsibilities to underutilised members – no idle members</td>
<td>Shift of responsibilities to underutilised members – no idle members</td>
</tr>
</tbody>
</table>

#### Communication and Interaction

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communication frequently inaudible or incoherent</td>
<td>Communication not always through leader or not relayed to scribe</td>
</tr>
<tr>
<td>Comms on many different levels simultaneously</td>
<td>Orders not always “read back” to confirm or notes not checked</td>
</tr>
<tr>
<td></td>
<td>Team lead is “hub” for all info and questions, all orders come from team leader</td>
</tr>
<tr>
<td></td>
<td>Comms clear and audible to all</td>
</tr>
<tr>
<td></td>
<td>All key findings verbalised to all</td>
</tr>
<tr>
<td></td>
<td>All orders and communications acknowledged and “read back”</td>
</tr>
</tbody>
</table>

#### Assessment and Decision Making

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient not assessed or major omissions in patient assessment</td>
<td>Assessment somewhat out of order</td>
</tr>
<tr>
<td>Multiple team members not clear on overall plan or “next step”</td>
<td>Patient thoroughly considered but some clinical information lacking or not assessed</td>
</tr>
<tr>
<td></td>
<td>Full patient assessment</td>
</tr>
<tr>
<td></td>
<td>Findings summarised</td>
</tr>
<tr>
<td></td>
<td>Team input solicited for complex problems or deviations from expected course</td>
</tr>
<tr>
<td></td>
<td>Goals and plan communication to team – all aware of plan</td>
</tr>
</tbody>
</table>

#### Global Awareness / Coping With Stress

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unforeseen events disrupt pt assessment and treatment</td>
<td>Untoward findings cause minor delay but do not preclude task completion</td>
</tr>
<tr>
<td>Team members stressed or panicky</td>
<td>Limited or no planning beyond management of immediate problem (i.e. discharge arrangements)</td>
</tr>
<tr>
<td>Not anticipatory of outside events impacting care, no contingency plan</td>
<td>Untoward findings, distractions do not disrupt evaluation and treatment</td>
</tr>
<tr>
<td></td>
<td>Team plans ahead (i.e. arranges discharge assessments)</td>
</tr>
<tr>
<td></td>
<td>Awareness of problems which impact care (e.g. theatre unavailable)</td>
</tr>
<tr>
<td></td>
<td>Awareness of team members condition (anxiety, fatigue)</td>
</tr>
<tr>
<td></td>
<td>Steps taken to minimise impact of resource / personnel deficit</td>
</tr>
</tbody>
</table>
Appendix 4. Final POSCH checklists for postoperative complications, optimised for use in Imperial College London, UK hospitals.

<table>
<thead>
<tr>
<th>POSCH Post-operative Surgical Patient Checklist</th>
<th>Obvious blood loss</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Surgical bleeding</strong></td>
<td>OR</td>
</tr>
<tr>
<td>Does the patient have two large-bore cannula (19G, green) or larger – at least one additional cannula if CVC in situ</td>
<td>Otherwise unexplained Hb drop</td>
</tr>
<tr>
<td>Check valid G&amp;S sample, or resend with bloods incl clotting</td>
<td></td>
</tr>
<tr>
<td>Review drug chart: cross off anti-thrombotic medications (e.g. Aspirin, Clopidogrel, Tinzaparin / LMWH)</td>
<td></td>
</tr>
<tr>
<td>Ensure patient NBM in case intervention required</td>
<td></td>
</tr>
<tr>
<td>If re-operation suspected or likely, liaise with theatres</td>
<td></td>
</tr>
</tbody>
</table>

**Stable patient, <1.5L loss:** ensure valid G+S, crossmatch 2-4 units
**Major bleed / unstable patient:** activate major transfusion protocol (dial 2222)

Inform senior (SpR or cons) after completing

Name

Sign and date

Imperial College Post-operative Checklist v.1.0 ©2013
**Anastomotic leak**

- Immediate fluid resuscitation and electrolyte replacement
  - Maintain urine output >0.5mL/kg/hr, CVP 8-12 cm H₂O

- Contrast imaging: CT (colorectal) or swallow (upper GI)

- ABG and bloods (FBC / U+E / LFT / CRP / Clotting / G+S)

- Ensure patient NBM in case intervention required

- Commence sliding scale if diabetic

- Broad-spectrum antibiotics
  - Tazocin (4.5g IV TDS) + Metronidazole (500mg IV TDS) + consider Gentamicin 5mg/kg IV OD

  **Remember to prescribe stat dose as well as regular dose**

- Inform senior (SpR or cons) after completing

**Imperial College Post-operative Checklist v 1.0 ©2013**

**Name**

**Sign and date**
Wound Infection

If purulent discharge or obvious collection
- Remove 2-3 clips from wound to allow drainage

Take cultures before commencing any antibiotics
- Send wound swab for MC+
- Blood cultures if pyrexial or spiking temperatures

Regular dressing changes

If cellulitis or systemic signs of sepsis consider antibiotics
- Mark extent of cellulitis and document
- Start Flucloxacillin 2g IV QDS or (if penicillin allergy) Clindamycin 1.2g IV QDS

Remember to prescribe stat dose as well as regular dose

Inform senior (SpR or cons) after completing

Name
Sign and date
**Urinary Tract Infection**

<table>
<thead>
<tr>
<th>Consider IV fluids if appropriate</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Send bloods – check renal function</td>
<td></td>
</tr>
<tr>
<td>Is catheter indicated (e.g. output monitoring or incontinence)?</td>
<td></td>
</tr>
<tr>
<td>• If still required, change catheter</td>
<td></td>
</tr>
<tr>
<td>• If no longer indicated, remove catheter</td>
<td></td>
</tr>
<tr>
<td>Take cultures before commencing antibiotics</td>
<td></td>
</tr>
<tr>
<td>• Urine (+/- blood if pyrexial)</td>
<td></td>
</tr>
</tbody>
</table>

**Remember to prescribe stat dose as well as regular dose**
- Routine UTI: Cephalexin 500mg bd (3d female / 7d male)
- Severe: Augmentin 1.2g IV TDS + Gentamicin 3-5mg/kg IV

Inform senior (SpR or cons) after completing

<table>
<thead>
<tr>
<th>Name</th>
<th>Sign and date</th>
</tr>
</thead>
</table>
**Pneumonia / Chest infection**

Suspicion of chest-related sepsis **AND**
At least TWO of leucocytosis, pyrexia, purulent sputum, drop in O₂ saturation levels

<table>
<thead>
<tr>
<th>Arterial blood gas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest radiograph</td>
</tr>
<tr>
<td>Send cultures <strong>before</strong> commencing antibiotics</td>
</tr>
<tr>
<td>• Blood</td>
</tr>
<tr>
<td>• Sputum</td>
</tr>
<tr>
<td>Maintain O₂ ≥ 94% or 88-92% if pre-existing lung disease</td>
</tr>
<tr>
<td>Chest physiotherapy and consider regular nebulisers if patient unable to cough and clear sputum</td>
</tr>
<tr>
<td>Broad-spectrum antibiotics <strong>prescribe stat dose as well as regular</strong></td>
</tr>
<tr>
<td>Tazocin 4.5g IV TDS or (penicillin allergy) Ciprofloxacin 500mg PO BD / 400mg IV BD</td>
</tr>
<tr>
<td>Inform senior (SpR or cons) after completing</td>
</tr>
<tr>
<td>Name Sign and date</td>
</tr>
</tbody>
</table>

*Imperial College Post-operative Checklist v 1.0 ©2013*
**Intra-abdominal sepsis**

Abdominal pain **AND** Systemic signs of sepsis: at least TWO of pyrexia, tachycardia, tachypnoea, leucocytosis

<table>
<thead>
<tr>
<th>Rule out other sources of sepsis</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Examine chest, wound, line sites, request urine dip</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Further investigations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• CT scan with contrast</td>
<td></td>
</tr>
<tr>
<td>• Serum lactate / FBC / U+E / LFT / CRP</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Commence sliding scale if diabetic or BM &gt; 10mmol</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Send cultures first: blood cultures</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Consider urine / wound / sputum</td>
<td></td>
</tr>
</tbody>
</table>

*Then:* Cefuroxime 1.5g + Metronidazole 500mg (stat + reg dose)

<table>
<thead>
<tr>
<th>Resuscitate with IV fluids, urine output &gt; 0.5mL/kg/hr</th>
<th></th>
</tr>
</thead>
</table>