# What influences the effectiveness of Quality Improvement in surgery?

## Learning from the evaluation of two large improvement programmes for emergency general surgery in the UK National Health Service


Timothy Stephens


A thesis submitted in fulfilment of the requirements of the degree of Doctor of Philosophy


William Harvey Research Institute

Barts and the London School of Medicine and Dentistry

Queen Mary University of London

# DECLARATION

The work in all chapters is a result of two large collaborative projects in which I played a central role. I, Timothy Stephens, confirm that the research included within this thesis is my own work. Where components of this collaborative work have been carried out with or by others that this is duly acknowledged below, and their contribution indicated for each chapter. Previously published material is also acknowledged below.

Components of the collaborative work included in this thesis that was undertaken by others:

- Chapter 2: The systematic review and Delphi study that determined the clinical components of EPOCH trial intervention was led by Professor Ravi Mahajan and others in the EPOCH trial team.
- Chapter 3: The collection of ethnographic data used within the mixed methods process evaluation was led by Professor Graham Martin and other colleagues within the EPOCH trial team.
- Chapter 5: The statistical analysis of Hospital Episodes Statistics data was undertaken by Professor David Cromwell at the Royal College of Surgeons. I was involved in the design and interpretation of these analyses.
- Chapter 6: The qualitative data collection and analyses was designed and led by me, but was undertaken by a team comprising Jonathan Bamber, Ellie Duncan and myself.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

# LIST OF PUBLICATIONS AND COLLABORATIONS

This thesis describes, analyses and synthesises the outputs from two major Quality Improvement programmes that I have co-led and evaluated. Both focused on improving care for patients requiring emergency general surgical procedures whilst concurrently progressing the field of Improvement Science. As large, national projects these outputs do not represent my work alone but rather collaborative efforts. However, my role in both programmes placed me at the forefront of each collaboration. Each paper and each chapter in this thesis details work that either I led or to which I made a substantial contribution.

Peden CJ, **Stephens T,** Martin G, Kahan BC, Thomson A, Rivett K, Wells D, Richardson G, Kerry S, Bion J, Pearse RM. Effectiveness of a national quality improvement programme to improve survival after emergency abdominal surgery (EPOCH): a stepped-wedge cluster-randomised trial. The Lancet. 2019 Jun 1;393(10187):2213-21.

**Stephens TJ,** Peden CJ, Pearse RM, Shaw SE, Abbott TE, Jones EL, Kocman D, Martin G, Epoch Trial Group. Improving care at scale: process evaluation of a multi-component Quality Improvement intervention to reduce mortality after emergency abdominal surgery (EPOCH trial). Implementation Science. 2018 Dec 1;13(1):142.

**Stephens TJ,** Peden CJ, Haines R, Grocott MP, Murray D, Cromwell D, Johnston C, Hare S, Lourtie J, Drake S, Martin GP. Hospital-level evaluation of the effect of a national Quality Improvement programme: time-series analysis of registry data. BMJ quality & safety. 2020 Aug 1;29(8):623-35.

Bamber JR, **Stephens TJ**, Cromwell DA, Duncan E, Martin GP, Quiney NF, Abercrombie JF, Beckingham IJ. Effectiveness of a Quality Improvement collaborative in reducing time to surgery for patients requiring emergency cholecystectomy. BJS open. 2019 Dec;3(6):802.

**Stephens TJ,** Bamber JR, Beckingham IJ, Duncan E, Quiney NF, Abercrombie JF, Martin G. Understanding the influences on successful Quality Improvement in emergency general surgery: learning from the RCS CholeQuIC project. Implementation Science. 2019 Dec;14(1):1-4.

# PUBLICATIONS NOT INCLUDED IN THE THESIS

Yang F, Walker S, Richardson G, **Stephens T**, Phull M, Thompson A, Pearse RM, for High EP. Cost-effectiveness of a national Quality Improvement programme to improve survival after emergency abdominal surgery: Learning from 15,856 patients. International Journal of Surgery. 2019 Dec 1;72:25-31.

**Stephens T,** Pearse RM. Learning from the EPOCH trial (Editorial). What we have learnt from a trial of an intervention to improve survival following emergency laparotomy. Anaesthesia Critical Care & Pain Medicine. 2019.

**Stephens T**, Johnston C, Hare S. Quality Improvement and emergency laparotomy care: what have we learnt from recent major QI efforts?. Clinical Medicine. 2019 Nov;19(6):454.

Aggarwal G, Peden CJ, Mohammed MA, Pullyblank A, Williams B, **Stephens T**, Kellett S, Kirkby-Bott J, Quiney N. Evaluation of the collaborative use of an evidence-based care bundle in emergency laparotomy. JAMA surgery. 2019 May 1;154(5):e190145-.

# ABSTRACT

**Background**

Robust evaluations of the effectiveness of Quality Improvement (QI) remain rare, and subsequently the evidence for the use of QI, and what may influence the success of such efforts, is weak.

**Methods**

EPOCH was a stepped-wedge, randomised trial of a QI programme, in 93 hospitals, designed to reduce mortality in patients requiring emergency abdominal surgery. CholeQuIC was a controlled evaluation of a 12-hospital project designed to reduce time to emergency laparoscopic cholecystectomy for patients with acute gallstone disease. Both studies had concurrent process evaluations.

**Results**

The EPOCH trial found no reduction in 90-day mortality associated with the improvement programme (16% mortality in both groups (Hazard ratio, QI vs usual care: 1.11 [0.96-1.28])). Hospital-level time-series analysis identified that only a small cohort of hospitals (14/93) improved half or more of the target care processes, suggesting a degree of implementation failure. Major influences included limited time and scarce resources to support clinicians leading improvement, including an onerous burden of data collection. CholeQuIC demonstrated that eight of 12 participating hospitals significantly reduced time to surgery when compared to national data from the same period (relative change in surgery ≤8 days, QI vs control group: 1.45 vs. 1.08 ([1.29-1.62]). Major influences include stakeholder

engagement, allocated time to lead plus effective project support and a willingness to test out new ideas. The QI methods used in both projects were similar, but the scale and complexity of the change required was less in CholeQuIC and more within the control of clinicians to improve.

**Conclusions**

Concurrent outcome and process evaluations are necessary to understand if and how Quality Improvement projects work. Choosing a problem amenable to clinician-led Quality Improvement, using QI multiple methods that focus on effective stakeholder engagement and protected time for clinicians to lead improvement are major influences on whether Quality Improvement is effective or not.

# PLAIN ENGLISH SUMMARY


**Background**

As the population ages, the need for surgery to treat surgical disease will continue to increase. The quality of care provided to surgical patients is known to impact upon both patients' experience of care and also their chance of survival. Concerns have been raised by doctors and professional bodies for many years about the variable quality of emergency surgery, and in particular emergency general surgery, in the UK NHS. Quality Improvement is a relatively new discipline that uses specific methods, such as collecting data on important activities and feeding the results back to doctors and nurses, to improve the quality of health care. Although increasingly popular as a method to address quality problems, the evidence that Quality Improvement really works remains weak, as does the evidence on what may influence the success of such efforts. This thesis is based on studies of two large-scale Quality Improvement projects: the EPOCH trial and the CholeQuIC Quality Improvement collaborative.

**Methods**

The EPOCH trial was a randomised trial of a quality improvement programme, in 93 hospitals, designed to reduce mortality in patients requiring emergency abdominal surgery. CholeQuIC was an evaluation of a 12-hospital project designed to reduce time to emergency gallbladder surgery. CholeQuIC used data from all other NHS hospitals from the same time period to act as control in comparison to the intervention group. Both studies had concurrent, mixed-methods process evaluations, including use of interviews and period of observation in hospitals, to understand what was happening during the delivery of the

Quality Improvement interventions. This combination of research methods provided data on both whether the interventions were effective and how they worked (and what influenced their effectiveness).

**Results**

The EPOCH trial did not demonstrate a significant reduction in mortality, which was the main outcome of interest. Hospital-level analysis, using specific methods called time-series analysis, identified that only a small number of hospitals (14/93) improved half or more of the target care processes. These processes included getting the patient into the operating theatre on time or admitting them to critical care units after surgery. This suggested a degree of implementation failure, whereby the Quality Improvement intervention failed to help doctors and nurses to make the necessary improvements to care. Major influences included limited time and scarce resources to support the doctors leading improvement and an onerous burden of data collection. CholeQuIC demonstrated that 8 of the 12 participating hospitals significantly reduced time to surgery and this remained significant when compared to national data from the same period. Major influences include effective efforts to engage colleagues and being given allocated time to lead the project, plus effective project and data support. The Quality Improvement methods used in both projects were similar but the scale and complexity of the change required was less in CholeQuIC, plus those leading the ChleQuIC had more time and were supported by the central project team

**Conclusions**

Various evaluation approaches are required to measure if, how and why Quality Improvement projects work or do not work. Findings from these evaluations indicate that it may be possible to improve care for emergency surgical patients. Effective stakeholder engagement plus protected time and project support for clinician-led Quality Improvement appear to be major influences. Our research suggests that insufficient staff time may have been a critical factor and similar programmes will only succeed if staff have the time and resources needed to deliver them.

## ACKNOWLEDGEMENTS

Many colleagues have supported me through my research training over the past six years and I am indebted to all of them. Of particular note, my supervisor Rupert Pearse has been a patient and steady source of expertise and guidance throughout my training. I will always be grateful for his support during this time. My external co-supervisors, Graham Martin and Sara Shaw, have been invaluable in supporting my development as a qualitative researcher. Finally, my frontline colleagues in hospitals across the United Kingdom, who trusted me to support and also evaluate their improvement efforts, are to be thanked. Without them, this thesis would not have been possible.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# PREFACE

### *A brief summary of my journey from nurse to Improvement Scientist*

As a critical care nurse whose career moved first toward medical simulation and human factors training and then to a more corporate quality and patient safety role, I spent many years frustrated with the systems my frontline colleagues were having to work in daily (and that I continued to work in occasionally). Nowhere was this more apparent than in my experiences in the simulation lab and how they were juxtaposed with my role investigating serious incidents. In the simulation lab, where we would train teams in managing sick patients, the participants would leave enthused and, more often than not, with a little more knowledge and insight than when they arrived. It felt good and was undoubtedly at least somewhat beneficial. However, every time I had to investigate a Serious Incident, I would see that often the root causes were not knowledge based, but human factors based, and nearly always contained one or multiple system failures. This thesis is not about the relation and / or difference between human factors and systems factors. Rather it is my attempt to understand how we can further the science of improvement, by generating learning about how we can 'do' Quality Improvement more effectively and also how we can most usefully study and continue to learn from new efforts.

### *My involvement in major Quality Improvement programmes in the NHS*

I held central roles in both the EPOCH trial and CholeQuIC Quality Improvement collaborative and worked on the projects from initial application for funding, through the design, delivery and project management, evaluation and finally dissemination stages. These were large collaborative efforts, as is the way with any scientific endeavour that is to have real impact upon practice and ultimately benefit patients. As such this work is not

solely my own, but my contributions at each stage, and my unique perspective on both projects, means that the story in the thesis is very much mine to tell.

In Chapter 1, Background, I detail the clinical setting, the problems that needed to be fixed by these two Quality Improvement projects and also a brief history of healthcare quality improvement. Chapter 2 describes how I, with colleagues, developed the complex interventions that sat at the heart of the projects. Chapters 3 and 4 detail evaluations of the EPOCH trial from different perspectives: a mixed-methods process evaluation that I co-designed and co-led, to understand how and why the intervention worked (or did not work) in reality (Chapter 3) and then a hospital level evaluation that I designed and led, using time series charts to understand at a granular level exactly what changed during the EPOCH trial Quality Improvement period (Chapter 5). Chapters 5 and 6 detail evaluations of the CholeQuIC Quality Improvement collaborative: a controlled evaluation that I co-designed to understand if the intervention worked compared to hospitals that did not participate (Chapter 5) and a mixed-methods process evaluation that I designed which was able to capitalise on knowledge of the participating sites' performance to provide learning about what may have influenced the success of efforts to improve (Chapter 6). In the discussion I provide my interpretation of these findings in the whole and what I believe are the key learning points about both doing effective Quality Improvement and also effectively studying Quality Improvement, so that we can learn how do it better in the future.

# CHAPTER 1: BACKGROUND

## *1.1 The clinical setting*

As the volume of surgical procedures performed worldwide continues to increase,[1, 2] the need for improvement in the quality and safety of surgical care has become a global health care priority.[3-5] This is of importance considering both the increasing age and complexity of the surgical population and the global mortality burden associated with surgery.[6, 7] In high income countries, such as the UK, the rate of surgical morbidity and mortality has fallen during the past 20 years, although the overall incidence has not[8], due to increases in the volume of surgery. This reflects significant improvements in perioperative care as well as surgical and anaesthetic techniques. However, specific surgical populations continue to have a high risk for morbidity and mortality when undergoing emergency surgery,[9, 10] and concerns have been raised that advancements in the quality and safety of elective surgical services have not been matched for emergency surgical procedures.[11] It is within this context that my research on improving emergency general surgery in the UK is set.

### *1.1.1 Improving the quality and safety of emergency general surgery in the UK*

One major aspect of acute health care delivery in the UK that has come under increased attention, regarding the safety and quality of the service provided, is emergency general surgery. General surgery is an umbrella term for the management of patients presenting with elective or emergency abdominal disease. It includes surgery of the gastro-intestinal tract (oesophagus, stomach, small and large bowel) and the repair of hernias. General surgery also encompasses liver, biliary tree and pancreas, breast and transplant procedures. Emergency general surgery is carried out by general surgeons, increasingly by those with a special interest in upper and lower gastrointestinal care. Up to 50% of general surgical beds

are taken up by emergency general surgery patients, highlighting the large volume of work undertaken by such services. [12] Typical surgical diseases requiring emergency general surgery procedures run along a spectrum from simple abscesses, through acute gallstone disease and infected appendices to major and complex surgery to treat infections or obstructions in the abdomen. In 2007, the Association of Surgeons of Great Britain and Ireland acknowledged that standards of care for emergency surgical admissions were often unsatisfactory. There was a failure to prioritise patients, inadequate senior input, and unsatisfactory allocation of infrastructure and manpower.[13]

This thesis focuses on improvement work in emergency abdominal surgery (known as emergency laparotomy) and emergency laparoscopic cholecystectomy (gallbladder removal) for gallstone disease. Emergency laparotomy is a common emergency procedure and is considered high risk in terms of associated morbidity and mortality.[14] Acute gallstone disease is relatively high volume in terms of cases per week seen in an individual hospital but is considered generally low risk in terms of mortality. However, there is a potential high burden in terms of reduction in quality of life if gallstone disease is left without definitive treatment, due to the risk of recurrent symptoms.[15] Further details about these conditions, and associated procedures, are presented below to provide background context for the improvement work detailed in this thesis.

### *1.1.2 Reducing the mortality associated with emergency laparotomy*

More than 1.53 million adults undergo in-patient surgery in the UK National Health Service (NHS) each year with a 30-day mortality of 1.5%.[1] However, patients undergoing emergency

laparotomy have a much greater risk of death.[14, 16] Emergency laparotomy is a collective term that describes a heterogeneous group of unplanned intra-abdominal surgical procedures that are performed for a variety of indications, including intestinal obstruction, perforation of the bowel or peritonitis plus complications of elective surgery.[14] Approximately 30,000 emergency laparotomies are performed annually in England and Wales. Data available prior to the commencement of the EPOCH trial (in 2014) indicated that mortality was high, with a 30-day mortality of between 13.3% and 19%.[16-18] A key study in 2012, using data from the Emergency Laparotomy Network, found that there were substantial variations in the way that patients requiring emergency laparotomy were cared for. For example, wide variations were found in the grade of surgeon and anaesthetist performing the operation, how long it took to get the patient into the operating theatre and whether the patient was admitted to critical care afterwards. These variations were found to be associated with differences in mortality rates, and it was hypothesised that standardising care may lead to improved outcomes.[16] These findings prompted a report by the Royal College of Surgeons of England, commissioned by the UK Department of Health, which proposed extensive improvements to quality of care for this patient group.[11] Recommendations included interventions across the pre, intra and post-operative phases such as consultant-led decision-making, cardiac output guided fluid therapy and early admission to critical care.

Concurrently this situation prompted the Healthcare Quality Improvement Partnership (HQIP) to prioritise the setup of the National Emergency Laparotomy Audit (NELA). The audit officially started in December 2012 with the aim to enable the improvement of the quality

of care for patients undergoing emergency laparotomy through the provision of high-quality comparative data from all providers of emergency laparotomy. NELA was established to enrol the patients treated in NHS hospitals within England or Wales who were aged 18 years and over and who undergo an expedited, urgent or emergency (NCEPOD definitions) abdominal procedure on the gastrointestinal tract. The operations that NELA covers include: 1) Procedures involving the stomach, small or large bowel, or rectum for conditions such as perforation, ischaemia, abdominal abscess, bleeding or obstruction; 2) Washout/evacuation of intra-peritoneal abscess (unless due to appendicitis or cholecystitis); 3) Bowel resection/repair due to incarcerated umbilical, inguinal and femoral hernias (but not hernia repair without bowel resection/repair); 3) Return to theatre for repair of substantial dehiscence of a major abdominal wound or after patients underwent non-elective gastrointestinal surgery. There are a number of abdominal procedures that are outside the scope of the Audit, including: 1) Uncomplicated appendicectomy or cholecystectomy; 2) Non-elective hernia repair without bowel resection; 3) Vascular surgery, including abdominal aortic aneurysm repair; 4) Caesarean section, obstetric laparotomies or gynaecological laparotomy; 5) Laparotomy/laparoscopy for pathology caused by blunt or penetrating trauma.

The data items in the patient dataset were chosen based on their relevance to measuring practice against clinical recommendations and national standards of care, and the need to adjust for differences in the characteristics of patients and operations between hospitals. The dataset contains data items covering various characteristics of the patient and the care they received: 1) Patient age, gender, region of residence; 2) Preoperative assessment and

imaging; 3) Preoperative patient risk factors; 4) The type of procedures performed and the seniority of the surgeon and anaesthetist that performed it; 5) Postoperative patient risk factors; 6) Postoperative care, including the use of critical care and input from Elderly Medicine specialists where appropriate.

NELA was a mandatory national audit for all acute NHS Trusts in England and Wales. Each Trust was required to nominate surgical and anaesthetic leads for the audit who would have overall responsibility for data collection and input.

This was the background context that led to the funding, by the National Institute of Health Research, of the EPOCH trial in 2013.

*1.1.3 Improving early access to emergency laparoscopic cholecystectomy*

Gallstone diseases account for approximately one-third of emergency general surgery admissions and referrals.[19] The commonest presentation is acute biliary pain (56%), followed by acute cholecystitis (36%), and gallstone pancreatitis (4%). The majority of patients presenting to hospital with biliary pain will go on to have the removal of the gallbladder (cholecystectomy) as definitive treatment. Around 20–33% of patients with acute cholecystitis or pancreatitis will re-present with gallstone-related symptoms before they have a cholecystectomy.[20-23] Current national guidance from the UK National Institute for Health and Care Excellence (NICE) is for laparoscopic cholecystectomy within seven days of diagnosis of acute cholecystitis, and within index admission for pancreatitis.[24] Professional associations and societies, including International Hepato-Pancreato-Biliary

Association, World Society of Emergency Surgery and British Society of Gastroenterologists provide similar guidance on times to cholecystectomy.

Reducing the time to surgery for people who need a cholecystectomy reduces the number of times patients are readmitted with the same diagnosis and decreases their overall length of stay. Compared with delayed cholecystectomy, emergency procedures are associated with overall fewer workdays lost and higher patient satisfaction and quality of life.[20, 23] Furthermore, patients with acute pancreatitis run the risk of a fatal episode whilst awaiting cholecystectomy which is reduced by early removal of the gallbladder and stones.[25] A traditional concern that complications leading to conversion from laparoscopic to open surgery or risks of bile duct injury are higher with emergency surgery than delayed are not supported by current evidence.[21, 26] Overall, patients having early surgery (within the first week of presentation) do better on all indicators than those in delayed surgery groups.[27, 28]

However, symptomatic gallstone patients in the UK wait longer and are more likely to be readmitted than in many other countries. Patients in France, USA and Sweden tend to undergo cholecystectomy on first admission with an average length of stay under 36 hours. [29-31] As with care delivery for emergency laparotomy, within the UK there is wide variation between NHS hospitals in the management of these patients, and wide variation in cholecystectomy rates despite the seven-day standard given in NICE guidance.[23, 24] It was in this context that the Royal College of Surgeons of England (RCS) decided to choose this issue as the focus for its first ever Quality Improvement Collaborative (CholeQuIC).

With both these identified quality problems in emergency general surgery service provision, it was the *variation* in care for the same or similar patients and procedures that drove the impetus for using a Quality Improvement approach. The foundation of industrial Quality Improvement techniques, largely drawing on the work of W Edwards Deming and Walter Shewhart, is based on the notion that *uncontrolled variation is the enemy of quality*.[32] Quality Improvement is largely focused on reducing uncontrolled or unnecessary variation so that patients all receive care in line with what is consider best practice. In this next section I provide a brief summary of Quality Improvement and the history of its application in health care and specifically within the UK NHS.

## *1.2 Quality in health care and Quality Improvement*

### *1.2.1 What is quality in health care?*

Avedis Donabedian stated that quality care is "care which is expected to maximise an inclusive measure of patient welfare, after one has taken account of the balance of expected gains and losses that attend the process of care in all its parts".[33] The use of the term 'patient welfare' is notable because it combines the concepts of patient outcome and patient experience, which have come to be defined separately in subsequent definitions of healthcare quality (see below). There is also an interesting recognition that the process of receiving health care may inherently bring both positive and negative consequences with it; by reducing variation and delivering care in line with best practice (as per Deming and Shewhart), we have an opportunity to minimise the negative and maximise the positive. The second definition of quality of health care, which has remained one of the most used and aimed for globally is from the Institute of Medicine in the US, which in 2000 defined quality care as care that is effective, safe, patient-centred, efficient, equitable and timely.[34] In the UK NHS the definition of quality, as enshrined in the NHS Outcomes Framework, and with a set of measurable indicators identified, is largely a simplification and amalgam of the above, with the statement that 'the NHS is organising itself around a single definition of quality: care that is effective, safe and provides as positive an experience as possible'.[35]

Using the Institute of Medicines definition, that are some quality problems that are immediately obvious for patients requiring either emergency laparotomy or laparoscopic cholecystectomy. Overall, from the data presented in the sections above, it would be fair to say provision of these procedures within the NHS was not wholly effective, due to the high

mortality associated with emergency laparotomy and the high readmission rate for biliary patients, not wholly safe due to the variations in standards of care delivered for both patient groups nor efficient or timely, due to systemic delays identified in the delivery of both procedures. Lack of patient centredness seems a particular problem for biliary patients, due to the long waits for surgery with associated pain, reduced function and re-admissions but is almost certainly a feature of the inpatient stay for those requiring emergency laparotomy too due to poorly organised and delayed care delivery. Certainly, it was felt in our community at that time that patients certainly were not receiving care 'that provides as positive experience as possible'. The issue of equity is harder to assess, but at the very least, delays in definitive care for biliary patients will disproportionally effect those on low income, and especially those on zero-hour contracts, due to working days lost due to gallstone related problems.

Considering the Donabedian definition of quality raises a key issue for both procedures, regarding the risks and benefits of having or not having emergency surgery. In the case of emergency laparotomy, which is a major operation, there is a sub-set of older, frail patients for whom the benefit of the procedure may not necessarily outweigh the risk of the presenting condition. In such situations, quality care may require a frank shared decision-making conversation and a non-surgical palliative option is offered to the patient to consider. In this case the perspective of the quality dimension of 'effectiveness' is important; pro-active palliation is highly ineffective at reducing mortality but can be highly effective at providing a good death.

For emergency laparoscopic cholecystectomy the procedure risk is lower, although a general anaesthetic is still required which poses its own risk for the older, frail and multi-morbid patient. In such cases an emergency procedure may not be in the patients' best interest, but again a shared-decision making conversation where the patient can be presented with the potential risks and benefits would be warranted.

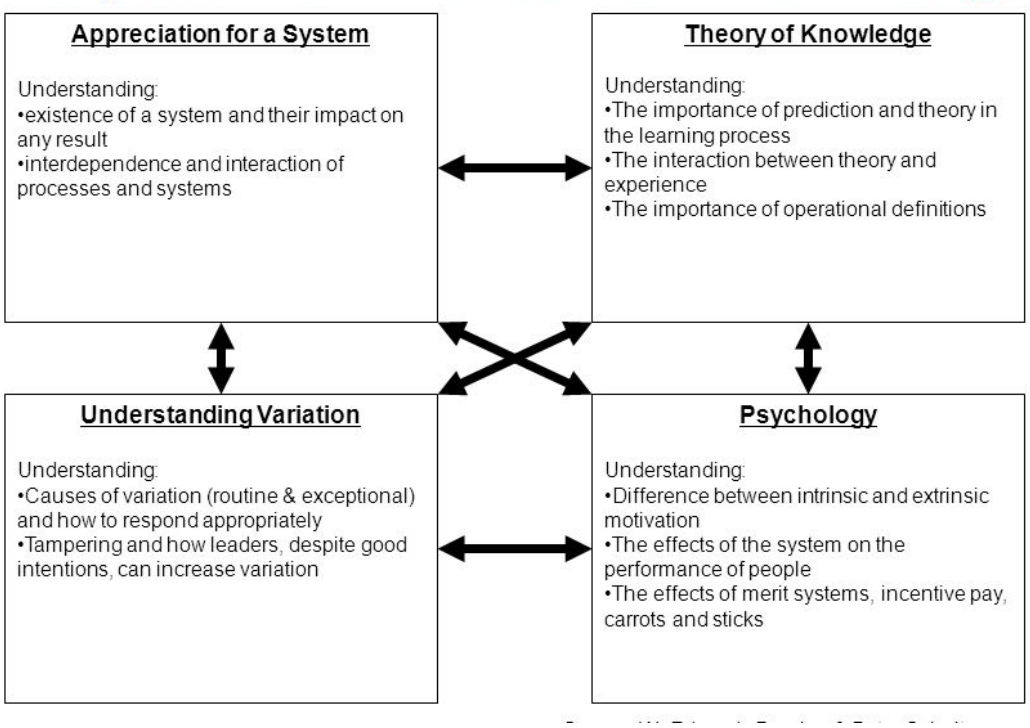
In both cases therefore, whatever definition of quality is used, the assessment of quality needs to be undertaken from the perspectives of both the clinician AND the patient and not just the latter. There will be substantial overlap, as patients and clinicians concur on obvious quality issues that may lead to risks and danger, but likely divergence too as some of the experiential aspects of quality will be judged differently by those providing and those receiving care. [36]

*1.2.2 What is Quality Improvement in health care?*

Quality Improvement can be understood as purposeful efforts to improve quality in line with the definitions as above, for example better patient outcomes or better system performance, using structured methods.[37] Quality Improvement efforts generally can be considered to have a hard core (targeting a change in a clinical intervention or care process that will positively impact upon patient care, experience or outcome) and then a soft periphery which is the (sometimes multiple) structured methods used to make the target change happen.[38] The theory and practice of these structured methods for Quality Improvement are broadly based on a number of principles first defined by W Edwards Deming (see Figure 1.1).[39, 40] Broadly applying these principles in today's health care

includes training staff in the nature of systems, use of statistical and quantitative data over time to understand variation,[41] inclusiveness such that all workers have an opportunity to contribute and act on ideas, and a focus on the needs and experience of the people served by a system. [34, 42] They also include employment of learning in action, by trialling out ideas and tests of change and a focus on teamwork and co-operation amongst all stakeholders delivering a service.[41] Below is a brief critical review of the main structured methods used for improving the quality of health care services.

***Figure 1.1: Original Principles of Quality Improvement, from W Edwards Deming***




*1.2.3 A brief summary of QI methods*

In this summary I describe the main QI methods used with the NHS. This is not intended to be exhaustive but rather a summary of the methods most commonly promoted and used before the inception of the improvement projects detailed in this thesis. I divide these methods into 2 categories: 1) those intended to organise and structure the delivery of

clinical interventions or processes to facilitate improvements in the quality of care and 2) those intended to invoke behaviour or process change to facilitate improvement.

1) Methods to structure the delivery of interventions and processes:

*Care pathways*

Care pathways are a mechanism through which the delivery of clinical interventions or processes can be organised and structured. They can be thought of schematics that define and operationalise a set of interventions or processes for a patient group. Based upon a Cochrane review process, the key criteria that define a care pathway are: 1) it aims to standardise care for a specific problem, procedure or episode of care in a specific patient population; 2) it channels the translation of guidelines and / or evidence into action within local structures; 3) it details the steps in a course of treatment or care in a pathway, algorithm, protocol or other 'inventory of actions' ; 4) there are time frames or criteria based progression i.e. the patient moves 'along' a pathway and 5) it is designed to be used by all the multidisciplinary team members involved in care of the target population for the specific episode of care.[43, 44]There is some debate regarding how many of these criteria need to be present to define something as a care pathway. Care pathways have become increasingly used in many healthcare settings, including surgery. Within general surgery their use has particularly been promoted through the widespread adoption of Enhanced Recovered After Surgery (ERAS) pathways for colorectal cancer surgery.[45] Care pathways have been associated with reduced in-hospital complications and improved documentation without negatively impacting on length of stay and hospital costs and have been found to effective in supporting the timely implementation of clinical interventions, although as

discussed this generally requires an associated package of improvement or implementation methods for this to occur.[46, 47]

*Care bundles*

A care bundle is a different mechanism through which the delivery of clinical interventions or processes can be organised and structured. The concept was developed by the Institute of Healthcare Improvement in the USA and was first used with a 'ventilator care bundle' for critical care patients.[48] Although the overall purpose of a care bundle has overlap with a care pathway there are some key differences. Bundles generally have a small number of components (usually three to five) and crucially these all need to have strong evidence of effectiveness. The idea behind the utility of a care bundle is that it supports the consistent and reliable delivery of the target evidence-based practices, so all components of the bundle need to be delivered to achieve the optimal effect for patients. [48, 49] Within surgery, but not specifically general surgery, care bundles have been designed to reduce surgical site infections.[50] A systematic review of care bundles found that most included studies had a weak design, generally observational, with a moderate to high risks of bias and so the current evidence for the effectiveness of bundles remains weak.[51]

For any clinical practice or set of practices targeted by QI, whether organised in a care pathway or bundle or not, defining the target practice does not, in general, lead to improvement. Rather the prevailing wisdom in QI is that one or more of the following methods is required to facilitate improvement.

2) Methods for behaviour or process change to facilitate improvement

*Audit and feedback*

In an audit and feedback process, an individual's or team's performance is measured and then compared with agreed professional standards or targets. The results of this comparison are then fed back to the individual or team with the aim of encouraging greater adherence to the desired professional standards.[52, 53] Most audit and feedback has been focused at the individual clinician level and the most recent Cochrane Systematic Review of the evidence found this methodology can lead to small but potentially important improvements in professional practice.[52] Effectiveness seems to depend on baseline performance, with poor baseline performance more amenable to change, and also on how the feedback is provided. Studies of team-based audit and feedback were absent from the 2012 Cochrane Review but there is a small but emerging body of research suggesting that a team-based approach can be effective, although the mechanisms of effect may be different from individual-level feedback.[54, 55]

*The Model for Improvement and Plan-Do-Study-Act (PDSA) cycles*

The Model for Improvement focuses on using data to understand current performance, setting clear, measurable goals and then developing potential solutions to achieve these goals.[41] The PDSA cycle promotes rapid cycle testing of these potential solutions, re-evaluating performance on a regular basis and adjusting solutions iteratively based on that review. The four stages mirror the scientific experimental method of formulating a hypothesis, collecting data to test this hypothesis, analysing and interpreting the results and

making inferences to iterate the hypothesis.[56] Unlike audit and feedback, which has been extensively researched, there has been little empirical evaluation of this approach. The research that does exist suggest, on the PDSA cycle specifically, suggests clinical staff often find PDSAs difficult to carry out in the methodical fashion intended, often with consequently disappointing results.[56, 57]

*Quality Improvement Collaboratives*

A Quality Improvement Collaborative is an organised, multifaceted approach that includes teams from multiple health care sites coming together to learn, apply and share improvement methods, ideas and data on service performance for a given health care topic.[58, 59] Although sites may be helped to use methods such as audit and feedback or the Model for Improvement to effect change locally, the added value of the collaborative approach is thought to be the creation of a new co-operative space for clinicians (and sometimes patients too) to share, learn and potentially challenge one another outside their normal working environment.[59, 60] A recent systematic review found that 53 of the 64 (83%) studies that met the standards for inclusion found measured improvements in at least one target process. Collaboratives reporting success generally addressed relatively straightforward aspects of care, had a strong evidence base and noted a clear evidence-practice gap in an accepted clinical pathway or guideline.[59] Notably, only one of the included collaboratives focused on perioperative care.[61] The same review also excluded many studies due to weak study designs, highlighting the ongoing challenges of effectively studying and evaluating Quality improvement efforts.

*Lean methodology*

Lean methodology was originally developed in the automotive industry to make factory floor processes more efficient. The specific goal of Lean methodology in health care, which differentiates from other Quality Improvement methods, is to approach processes through the lens of reducing waste, redundancy and inefficiencies within processes and systems. It is included as an improvement method, rather than a managerial or cost-saving activity, because it is underpinned by the logic that reducing inefficiency can directly increase the time staff members have for patient care, so that patient outcomes and / or patient experience is enhanced.[62] Common steps in Lean methodologies are to generate shared understanding of processes and system, generally using process maps; engaging stakeholders in the re-organisation and re-design for effectiveness and / or efficiency; data collection to improve error detection and so increase awareness and process reliability. Multiple Lean projects have been evaluated with many reporting success. However, three systematic reviews, two on Lean in health care generally and one focused specifically on surgery, found that most included studies had a weak design, generally observational, with a moderate to high risks of bias.[62-64] As such the current evidence for the effectiveness of Lean remains weak.

### *1.2.4 The influence of context on Quality Improvement*

Context characterises the overall environment in which Quality Improvement activities take place and includes prevailing national policies, local organisational structure and the culture both of an organisation overall and of the specific departments within an organisation.[65]

Contextual factors are distinct from the clinical and Quality Improvement interventions within an improvement project but are highly influential on the success or otherwise of such projects.[66] Using one of our target procedures (emergency laparotomy) as an example, recent work by the NELA group has identified that a substantial amount of the observed variation in mortality nationally was explained by differences in hospital structures and characteristics, such as the number of operating theatres available or the existence of an emergency surgical unit.[67] Although he never specifically used the term 'context', the work of Donabedian on the relationship between structures and processes in a given health care setting and the outcomes that can be achieved, may be considered to remain relevant.[68] He suggested that whilst some patient outcomes are fixed based upon physiology, comorbidities and diseases process, good structure (for example an adequate number of critical care beds for the needs of a hospital) increases the likelihood of good process (for example the quality of care in the ICU), and good process increases the likelihood of a good outcome. He was one of the first to suggest combining organisational, behavioural and health sciences to better understand the impact of the health care environment and systems / processes in that environment on health outcomes,[68] beyond structural characteristics of a hospital or other setting for improvement is the social context. This can be thought of as operating at three levels: macro, meso and micro.[65, 69] The macro level social context could be considered what is happening nationally regarding a quality problem, how much attention it is receiving, policy changes etc. At the meso level, contextual consideration would be normally focused on the organisation in which improvement is happening, for instance the overall culture of the organisation, how supportive of improvement (or this particular improvement) the leadership team is etc. At

the micro level, the social dynamic and relationship of the main stakeholders in the system under focus e.g. the surgeons, anaesthetists, nurses, managers and patients across the various departments (emergency department, wards, operating theatres, critical care), where care for these patients was delivered. Donabedian did not include this social context in his structure, process and outcome model, although he was very much alive to the importance of the social and human side of health care delivery. However the importance of understanding context as an aspect of planning, doing and evaluating Quality Improvement has developed substantially over time,[70, 71] and is now fundamental to the recommendation for studies of Quality Improvement to include qualitative research components to understand context.[72]

### *1.3 The state of improvement science – before the EPOCH trial (2014)*

Prior to the EPOCH trial there were good examples of discrete Quality Improvement interventions being associated with improved patient outcomes,[73, 74] but there were others that yielded mixed or disappointing results.[75-78] This seemed especially true for complex interventions requiring co-ordinated change across a health care system. Moreover, the lack of rigorous evaluations of many Quality Improvement efforts hampered the development of the improvement science,[79] and the degree of optimism bias or 'magical thinking' about the positive effects of Quality Improvement intervention remains a concern held by many contemporary critics of Quality Improvement.[80, 81] Despite this, the direction in health care policy, in the NHS and beyond, continued toward ever more widespread use of Quality Improvement to drive large-scale change. Below, I describe four major influential Quality

Improvement programmes within the NHS, and their evaluations, to provide some further context regarding the state of improvement science before the EPOCH trial.

### *1.3.1 A review of previous influential large-scale QI projects in the UK NHS*

#### *The Productive Ward*

The Productive Ward programme was developed by the NHS Institute for Innovation and Improvement (NHSI), launched in England in 2007 and ran initially over five years. The programme was underpinned by Lean methodology, to simplify processes and reduce waste in nursing and ward based processes such as altering patient handover time or reorganising storage facilities. The central concept of the programme was that, as nurses improved ward processes (Figure 1.2), this would release more time to care for patients by reducing time wasted by inefficiency.[82, 83]

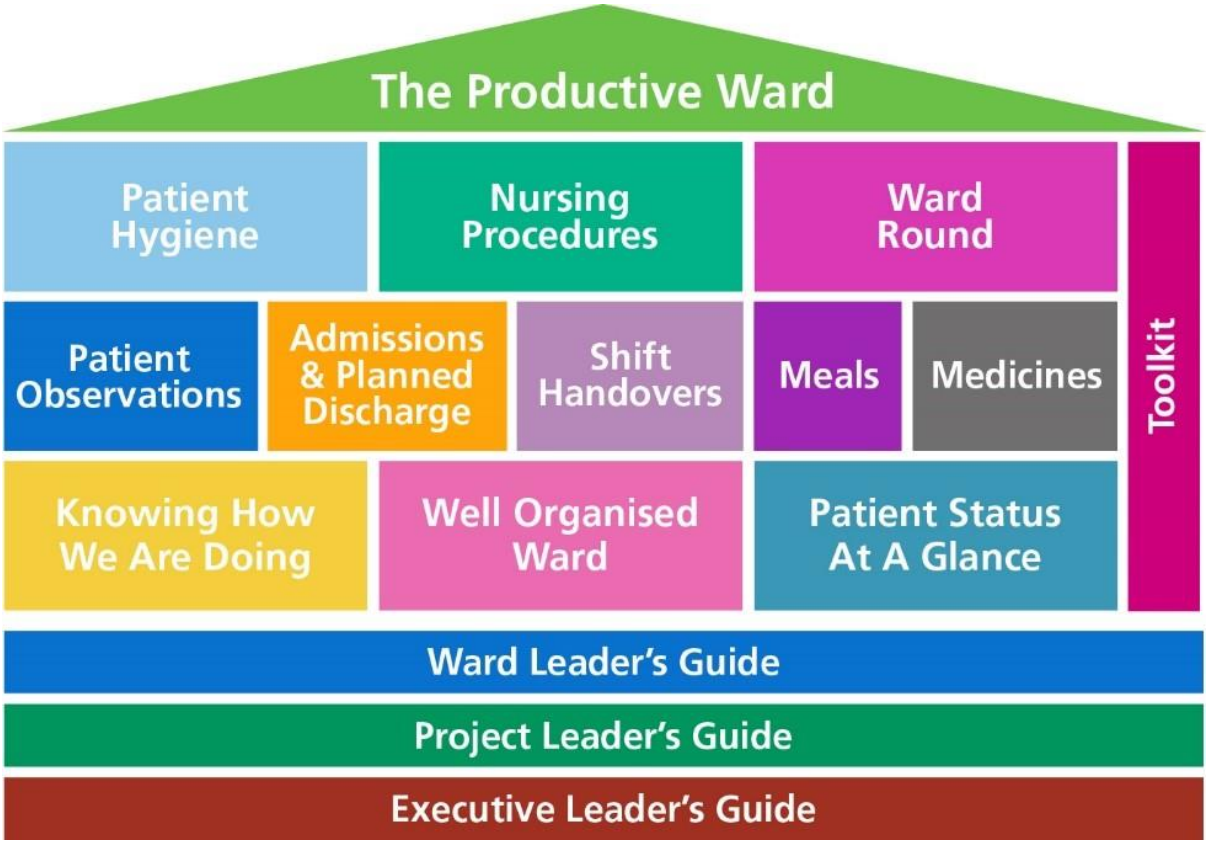



***Figure 1.2: Ward process and available resources for the Productive Ward programme***

A formal, theory based evaluation, using the Diffusion of Innovation framework,[84] was undertaken to examine the learning from and impact of the Productive Ward programme.[85, 86] The study was primarily qualitative in design and used interviews with national and regional stakeholders, a national web-survey of frontline staff, and case studies of implementation within five acute NHS Trusts. The study design meant that only the perceived impact of the programme, rather than specific impacts on measured outcomes, were reported. These reported impacts included: staff having more time to provide care to patients, improved teamworking, better organised working environments, reduction in patient falls and lower staff stress levels. A subsequent systematic review of studies evaluating the programme using a range of quantitative methods and metrics found some

evidence of reduction in time spent inefficiently but also reported that most studies used designs that could lead to a high risk of bias.[87] As such, although the perception of the Productive Ward programme was largely positive, and it may have been a useful first introduction of any kind of Quality Improvement thinking for many NHS staff, it was hard to confidently associate any firm benefits with the programme.

*Safer Patients Initiative 1 & 2*

From 2004 to 2008, the Health Foundation in the UK, and the Institute for Healthcare Improvement (IHI) in the US, partnered to fund and deliver the Safer Patients Initiative (SPI). This was a two-phase programme that began with four pilot NHS hospital sites (SPI1) and culminated in a 20-site programme (SPI2).[75, 76] The Safer Patients Initiative was the first major Quality Improvement programme addressing patient safety in the UK. It was designed to help hospital teams to improve the reliability of multiple acute care processes through the implementation of a comprehensive package of evidence-based clinical practices. Teams were supported in using various improvement methods including PDSA cycles, audit and feedback of process data to colleagues and facilitating linking and sharing between teams through a Quality Improvement Collaborative (Figure 1.3).

***Figure 1.3: Safer Patients Initiative processes and improvement methods[75]***

**Change package elements by work area**:
• Perioperative care: DVT prophylaxis, beta-blocker use, focus upon surgical site infections, communication ('SBAR'; safety briefings)
• Medicines management: medicines reconciliation, focus upon high-risk medications (anticoagulants).
• General ward care: early warning systems and outreach/rapid response team, communication ('SBAR'; safety huddles), infection prevention and control, hand hygiene
• Critical care: ventilator bundle, central line bundle, multi-disciplinary ward rounds, infection prevention and control, daily goal sheets
• Leadership: leadership walk-rounds, strategic prioritization of quality and safety issues

**Programme tools and methodology:**
• Continuous Quality Improvement approach and philosophy: semi-autonomous local Quality Improvement teams
• PDSA cycles and small tests of change
• Incremental spread to successively larger work systems
• Process measurement and analysis of run charts to determine effects of process changes
• Expert faculty support from experienced clinical improvement leaders (site visits, conference calls, presentation seminars and online email support)
• Large-scale learning sessions for multi-disciplinary improvement teams from each site (with educational and support components)
• Online extranet for uploading and comparing process data generated by each site, with monthly faculty feedback
• Collaborative learning community for networking and sharing best practices

DVT, deep vein thrombosis; 'SBAR', Situation Background Assessment Recommendation; PDSA, Plan-Do-Study-Act.

Both phases of the Safer Patients Initiative incorporated concurrent internal and external evaluations. The internal evaluation, using an uncontrolled study design, identified multiple improvements across target processes including central venous catheter infection rates reduced to zero in 17 hospitals, ventilator associated pneumonia rates reduced to zero by 10 hospitals and Methicillin Resistant Staphylococcus Aureus infection rates (a major NHS problem at the time of the project) halved by seven hospitals. These data were used to hail the Safer Patients Initiative as a success, and were said to represent thousands of patients who were prevented from suffering avoidable harm.[88] The external evaluation, which used

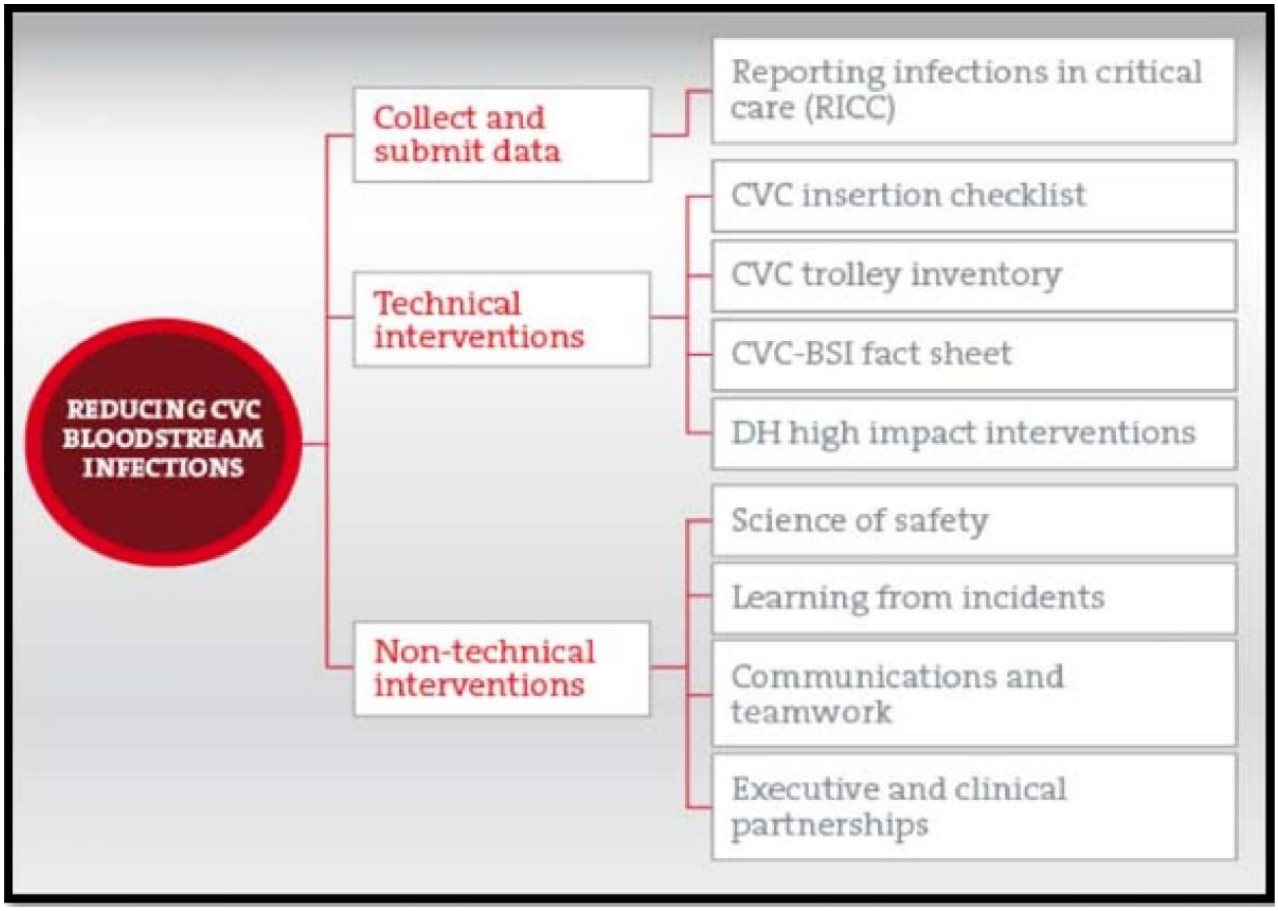
matched contemporaneous controls, found improvements in both intervention and control sites with no significant additional effect of the Safer Patients Initiative identified.[75, 76] The main explanation for the absence of an additional effect of the programme was a 'rising tide' phenomenon, whereby process improvements, and a more widespread recognition of patient safety issues, were driven by common forces affecting all NHS hospitals. Despite no significant effect of the Safer Patients Initiative programme interventions being identified in the controlled evaluation, a generalised optimism in improving patient care through Quality Improvement methods was established on the back of the Safer Patients Initiative.[89, 90]

*Matching Michigan*

Matching Michigan was run from April 2009 to the end of March 2011 by NHS England. The aim was to replicate the success of the Michigan-Keystone project, a large-scale improvement program in the US, which reported dramatic reductions in the rate of central venous catheter bloodstream infections.[74] The original Keystone project was seen as a demonstration that large scale improvements in patient safety (more than 100 Michigan ICUs were involved) could be successfully achieved using a Quality Improvement approach. The cohort study design in Keystone meant that no causal relationship could be established but there was strong optimism that the intervention had 'worked'.[60, 74] Matching Michigan involved three components: technical interventions, focused on consistent use of evidence-based measures for reducing risks of infections; non-technical (behavioural) interventions, designed to intervene in local culture and systems; the establishment of a standardised national reporting system (see Figure 1.4). Matching Michigan was evaluated using a prospective, interventional, non-randomised, stepped, four-cluster, design which included

continuous feedback of results to participating ICUs, and a concurrent embedded ethnographic evaluation.[91, 92]

***Figure 1.4: Matching Michigan interventions***




Matching Michigan reported a final central venous catheter bloodstream infections rate of 1.48 per 1,000 CVC patient days across 215 adult ICUs, thus 'matching' the rate of 1.4 seen by the end of the Keystone project. However, the stepped wedge study design identified a less certain signal toward success. Clusters of ICUs that were waiting to join the program were reducing their infection rates in parallel to ones already in the program, and infections acquired outside ICUs (not a target of Matching Michigan) were reducing at the same rate as infections acquired inside ICUs (which were the target). This was seen as clear evidence

of an overall secular trend toward improvement, essentially another 'rising tide' effect, which meant that the reductions in infection could not confidently be attributed solely to the program interventions.

*Stroke 90:10*

The Stroke 90:10 Quality Improvement Collaborative was funded by the Health Foundation and ran from July 2008 to December 2010 across hospitals in the North West of England.[93] It was based on the Quality Improvement Collaborative approach[94] and was designed to improve key care processes for stroke patients in through the implementation of two evidence-based care bundles (early hours care and rehabilitation care). The Quality Improvement was evaluated in a 30-hospital, cluster randomised trial, using data from an ongoing national audit, to compare the hospitals participating in the Stroke 90:10 Quality Improvement Collaborative with the non-participating controls plus a concurrent embedded ethnographic process evaluation.[93, 95] Improvement was identified in both the intervention and control groups, suggesting an underlying secular trend toward improvement over time. However, the improvement in the intervention group was greater, suggesting a modest benefit to QIC participation.

All the above-mentioned projects had some form of concurrent mixed-method or qualitative process evaluations in recognition of the complexity of the Quality Improvement interventions being studied. The Medical Research Council in the UK recommends concurrent process evaluations for all effectiveness evaluations of complex interventions.[72, 96] A complex intervention is defined as one that is designed to change the behaviour of one

or more target groups.[96] The more components an intervention needs to achieve this behaviour change, the greater the number of people involved, and the more levels at which the intervention operates, the more complex it is. Furthermore, the greater the degree of freedom of both intervention components and the people involved to act independently or inter-dependently to influence outcomes, the more complexity increases. This often makes it hard to define the 'active ingredients' and to be sure which component or combinations of components is more important or more influential on outcome. A simple intervention should have predictable outcomes; a complicated intervention will often have a predictable outcome, but there will be a greater number of 'known unknowns' that require a deeper analysis to understand outcomes, and complex interventions have emergent outcomes that often require detailed multifaceted evaluations to understand what combination of influences contributed to the outcome achieved.[97-99]

### *The EPOCH trial and CholeQuIC Quality Improvement collaborative*

Above I have presented the background context to the two major pieces of Quality Improvement research that I worked on and that form the basis of this thesis. In both I played a central role in all aspects of the studies, including the funding application, design, set-up, data collection and analysis stages. At the inception of both the EPOCH trial and the CholeQuIC collaborative, there was a clearly defined and recognised problem of clinical quality to be addressed. Concurrently, there was an ongoing need for further rigorous evaluations of Quality Improvement, with low risk of bias, to better understand if Quality Improvement could be effective in improving health outcomes accompanied by mixed-

methods evaluations to understand what influences the effectiveness, or lack thereof, for such interventions.

The Enhanced Perioperative Care for High Risk patients (EPOCH) trial was a stepped wedge, cluster randomised trial in 93 hospitals across the UK. The trial was designed to evaluate whether a Quality Improvement programme to support implementation of a care pathway for patients requiring emergency abdominal surgery, could reduce 90-day mortality. The novel trial design capitalised on the new founded National Emergency Laparotomy Audit (NELA). Trial data were collected through the NELA database (www.nela.org.uk), and then linked using unique patient identifiers to Hospital Episode Statistics and the Office for National Statistics in England and Wales, and the Information Services Division of NHS Scotland, to provide data describing mortality and hospital re-admissions. The primary outcome measure was all-cause mortality within 90 days following surgery. Secondary outcomes were all-cause mortality within 180 days following surgery, duration of hospital stay after surgery and hospital re-admission within 180 days of surgery. We selected ten predefined process measures (key components of the care pathway) for inclusion in the main report: 1) consultant-led decision to operate, 2) consultant review of patient before surgery, 3) pre-operative documentation of risk, 4) time from decision to operate to entry into operating theatre, 5) patient entered operating theatre within time-frame specified by their urgency (less than 2 hours, 2-6 hours, 6-18 hours, or more than 18 hours), 6) consultant surgeon present in operating theatre, 7) consultant anaesthetist present in operating theatre, 8) cardiac output guided fluid therapy used during surgery, 9) serum lactate measured at end of surgery and 10) critical care admission immediately after

surgery. EPOCH also had concurrent a mixed-methods process evaluation running alongside to understand local context and how the QI programme within the trial was delivered in practice (Chapter 3).

The main EPOCH trial analysis found no effect on the interventions on any of the trial outcomes measures: 90-day risk adjusted mortality, length of hospital stay or hospital readmission. Analysis of ten trial process measures suggested little improvement had occurred as a result of the intervention across the entire cohort. These results did not differ significantly between hospitals activated earlier in the stepped-wedge design compared with those activated later.

Due to my central role in EPOCH, I as invited by the Royal College of Surgeons of England to build a team to design, deliver and evaluate their first even Quality Improvement Collaborative. Based upon a membership survey, this was focused on improving care for gallstone patients. The CholeQuIC QI collaborative was a 12-hospital project designed to support clinicians to safely modify local care pathways to reduce the time to surgery for patients with acute gallstone disease. CholeQuIC had a concurrent, mixed-methods, controlled evaluation running alongside to evaluate effectiveness and to understand how teams succeed and/or failed with their improvement efforts.

This thesis presents my substantial contribution to both these major pieces of research.

# RESEARCH QUESTION AND OBJECTIVES

### *Overarching research question*

What influences the effectiveness of Quality Improvement in emergency general surgery in the NHS?

### *Research objectives*

1. To design, using the best available evidence, QI programmes focused on improving care for patients requiring emergency general surgery (Chapter 2 & Chapters 6)
2. To investigate the effectiveness of these QI programmes using three different methods (randomised trial and interrupted time series for EPOCH and controlled evaluation and time series for CholeQuIC)
    I. Strategic Objective A (Chapter 3): to evaluate, in a stepped-wedge, cluster randomised trial, the effect of a Quality Improvement programme designed to support implementation of the EPOCH care pathway on 90-day survival following emergency abdominal surgery in NHS hospitals.

    II. Strategic Objective B (Chapter 5): to evaluate, using time-series analysis at the individual hospital level, whether participation in the EPOCH trial QI programme led to implementation of the EPOCH care pathway; to assess the relationship between care-pathway implementation and use of the implementation strategies; to describe the number of improvements in care processes overall; to describe which care processes were most commonly improved.

III. Strategic Objective C (Chapter 6): to evaluate, using a controlled evaluation approach, whether participation in a Quality Improvement Collaborative (CholeQuIC) reduced time to surgery to within eight days from admission for patients requiring emergency cholecystectomy

3. To investigate the influences on the effectiveness of the Quality Improvment programmes using concurrent mixed-methods process evaluations

I. Strategic Objective D (Chapter 4): To understand, using a mixed-methods process evaluation of the EPOCH trial, how hospital clusters and sites were recruited, delivery of the intervention at the cluster level, response to the intervention at the cluster level, delivery of the intervention at the site level and the response to the intervention by individuals targeted (in this case, the EPOCH Quality Improvement leads)

II. Strategic Objective E (Chapter 7): To understand, using a mixed-methods process evaluation of the CholeQuIC Collaborative, how the collaborative was delivered by the faculty and received, understood and enacted by the participants locally, and what influenced teams' abilities to improve care for patients requiring emergency cholecystectomy

# CHAPTER 2: Developing the components of a complex intervention: The EPOCH trial Quality Improvement interventions

## *2.1 Introduction*

The use of Quality Improvement approaches to reduce variations in healthcare processes and improve the standards of healthcare delivery has been increasingly encouraged over the two decades. QI interventions can be considered complex interventions, often with numerous active ingredients intended to influence the behaviour or clinical practice of a range of professionals and/or patient groups.[96] There is published guidance on accurately describing complex interventions such as those used in Quality Improvement — the Standards for Quality Improvement Reporting Excellence (SQUIRE) and the Template for Intervention Description and Replication (TIDieR) checklist,[66, 100] – but such detailed reporting has not always been the norm in Quality improvement literature. In this chapter I describe the design, rationale and delivery plan for the EPOCH intervention in line with this guidance. Chapters 3, 4 and 5 describe how the intervention was delivered in reality, how it worked in practice and how this differed from the intended design.

## *2.2 Trial intervention*

The trial intervention was a Quality Improvement programme designed to improve care processes and outcomes for patients undergoing emergency laparotomy. Recruited hospitals were grouped into 15 clusters of six to eight geographically co-located hospitals. Each recruited hospital was asked to nominate three senior clinicians (consultants) to act as Quality Improvement leads (referred to hereafter as 'QI leads') from key clinical areas (surgery, anaesthesia and critical care) and to confirm executive board support from each

hospital. The aim of the EPOCH Quality Improvement programme was to enable the nominated QI leads and their teams to effectively improve the care pathway for patients undergoing emergency laparotomy.

The EPOCH trial intervention operated at two main levels. At the cluster level, we (Carol Peden and myself) developed a Quality Improvement programme to train and support QI leads and their colleagues in the delivery of the hospital level intervention in each of the 93 trial sites. Quality Improvement interventions can be seen as having a hard core, the clinical processes or practices that are the focus of improvement, and a soft periphery, the improvement methods that will enable change to occur.[38] In the EPOCH trial, the hard core of the hospital level intervention was a set of recommended clinical processes, organised within a care pathway for patients undergoing emergency laparotomy. The Quality Improvement intervention (the soft periphery of the hospital level intervention) was designed to enable the QI leads and their teams to effectively implement the care pathway for patients undergoing emergency laparotomy. Below I detail the development and content of these interventions.

*2.2.1 The cluster level intervention: The EPOCH QI programme*

We developed a Quality Improvement programme to change the practice and culture of care for patients undergoing emergency abdominal surgery. QI leads from each stakeholder discipline (surgery, anaesthesia, and critical care) were tasked with leading hospital-wide improvement to implement the care pathway with the support and guidance of the national EPOCH Quality Improvement team. The key aims of the programme were: 1) to reframe the

high mortality rates for these patients as a 'social problem', requiring re-organisation of existing care processes rather than technical innovation; 2) to support QI leads to engage their frontline colleagues and executive leaders in the change process; 3) to train local QI leads and their colleagues in basic improvement skills based around the Model for Improvement;[41] and 4) to supporting teams to analyse and feed back key process measure data to their colleagues to drive change. The EPOCH Quality Improvement team provided a one-day activation and training meeting for each geographical cluster shortly before or during the first week of activation. The purpose of this meeting was to develop the knowledge, skills and attitudes that the QI leads required to achieve change. Nominated QI leads were informed 12 weeks before the date of activation to the Quality Improvement intervention. Five weeks before activation QI leads were sent a 'pre-activation' checklist, which included the requirement to review five sets of notes from recent patients to establish current performance and identify gaps in care delivery. A notes review tool was provided, and each hospital presented their findings at the initial cluster meeting. A training package was designed for hospital QI leads and their colleagues, the main content of which was delivered at the initial cluster activation and training meeting, and employed a mixture of didactic, workshop and discussion sessions. Publicity resources, such as pens, posters, lanyards and mugs were distributed to each team on the day, to be shared with colleagues to raise awareness about participation in the EPOCH trial.

A Virtual Learning Environment (VLE) housed all training resources and acted as a repository for all the tools and documents required to enact the EPOCH Quality Improvement strategies. This was created to support QI leads who had attended the training and wanted

further Quality Improvement resources, as well as ensuring that QI leads and other team members who could not attend the training meeting could view all the necessary presentations and resources. In particular, the site housed a tool developed to allow the creation of time-series charts, using local NELA data, to allow QI leads to monitor key care processes during the improvement period. It also incorporated an interactive 'route-map', providing evidence sheets for each of the clinical recommendations within the EPOCH pathway (Figure 2.1). All hospital QI leads were automatically registered for the VLE five weeks prior to activation and could request that additional colleagues and team members be registered.

Once a cluster was activated, telephone and email support for the intervention was available. Separate email contact, including a regular newsletter, was maintained with all hospitals (both activated and those in-waiting) by the trial manager. Each hospital was offered a small amount of funding (£3700) for QI leads to spend on relevant activities. Half-day follow-up meetings were added soon after commencement of the study, to offer teams formal opportunities to share successes and challenges as they progressed, supported by advice from the programme leads. All clusters were offered a follow-up meeting. There were also two national meetings to facilitate shared learning during the trial period. QI leads were only eligible to attend these if their hospital had been activated to the trial intervention.

### *2.2.2 The hospital level interventions*

*Evidence-based care pathway; 37 recommended processes of care*

The EPOCH trial care pathway was developed through an evidence-based Delphi consensus process to update existing guidelines published by the Royal College of Surgeons of England.[11] The purpose of the pathway was to define the gold standard of care for this patient group. Evidence for the component interventions was assessed using the GRADE criteria.[101] The 37 component interventions are detailed in Figure 2.1 and a graphical display, designed for the Quality Improvement programme to show how the patient may move along the care pathway, is displayed in Figure 2.2. The development of the pathway was undertaken by the whole EPOCH core investigator group with the Delphi consensus led by Professor Ravi Majahan. As a member of the core group I contributed to this process and reviewed the iterations of the consensus with the other core EPOCH trial group members. As such Carol Peden and myself, who were responsible for the develop of the QI intervention to support implementation of the pathway, had some scope to shape the form and content of the care pathway under development but it was not under our direct control.

*How was the EPOCH trial intervention designed to improve quality?*

The development of the care pathway was guided by the explicit notion that it would increase the quality of care provided to this patient group, up to what could be considered a gold standard based upon professional guidance (and thus a healthcare professionals perspective on quality). Considering the Institute of Medicines six domains of quality the pathway was designed to:

1) Improve effectiveness, by standardising the delivery of a range of evidence-based interventions (thought by experts or proven in randomised controlled trials) to reduce morbidity and mortality for this patient group.
2) Improve safety, primarily through mandating consultant-led care at all key stages which was considered a safety issue considering the acuity of this patient group (i.e. junior doctors should not be assessing and operating on these patients unsupervised).
3) Improve timeliness, by providing clear time frames within which key interventions (Computed tomography imaging, surgery and critical care admission) should be delivered to prevent delays.
4) Improve equity by standardising the care for what was considered a 'forgotten group' of patients, (this term was specifically used in the RCS /Department of Health guidance that the EPOCH intervention was based upon. Patients with conditions requiring emergency abdominal procedures, and especially those known to have poor outcomes are generally older multi-morbid and sometime frail adults). This standardisation would also potentially address regional or geographical inequities in care delivery.

The pathway did not explicitly address efficiency within its design, but it could be inferred by organising the necessary evidence-based care within the structure of a pathway, if implemented care would become more efficient, by reducing delays (waste) and improving outcomes such as length of hospital stay (resource management). The sixth IoM domain of quality is patient centred-ness, defined as, 'Providing care that is

respectful of and responsive to individual patient preferences, needs, and values and ensuring that patient values guide all clinical decisions. This was not considered in the design of the care pathway, the perspective on quality was solely that of the clinician. The EPOCH trial did have Patient and Public Involvement, but these patient researchers were not involved in the design of the intervention.

**Figure 2.1. The EPOCH care pathway components**

**Before surgery**

1. Consultant led decision making
2. Computed tomography imaging within two hours of decision to perform test
3. Early goal directed therapy for patients with severe sepsis/septic shock
4. Analgesia within one hour of first medical assessment
5. Antibiotic therapy within one hour of first medical assessment
6. Correction of coagulopathy
7. Maintain normothermia
8. Active glucose management
9. Documented mortality risk estimate
10. Provide patient and relatives with oral and written information about treatment

**During surgery**

11. Surgery within six hours of decision to operate
12. Consultant delivered surgery and anaesthesia
13. WHO safe surgery checklist
14. Early antibiotic therapy (unless inappropriate)
15. Fluid therapy guided by cardiac output monitoring
16. Low tidal volume protective ventilation
17. Maintain normothermia
18. Active glucose management
19. Prescribe post-operative analgesia
20. Prescribe post-operative nausea & vomiting prophylaxis
21. Prescribe post-operative venous thromboembolism prophylaxis
22. End of surgery risk evaluation
23. Measure arterial blood gases and serum lactate
24. Confirm full reversal of neuromuscular blockade
25. Document core temperature
26. Re-evaluate mortality risk estimate

**After surgery**

27. Admission to critical care within six hours of surgery
28. Analgesia: early review by acute pain team
29. Continued antibiotic therapy where indicated with microbiology review
30. Prophylaxis for post-operative nausea & vomiting
31. Venous thromboembolism prophylaxis
32. Maintain normothermia
33. Active glucose management
34. Daily haematology & biochemistry until mortality risk is low (senior opinion)
35. Nutrition: early dietician review with consideration of benefits of enteral feeding
36. Chest physiotherapy review on day one after surgery
37. Critical Care Outreach review on standard ward with use of Early Warning Scores

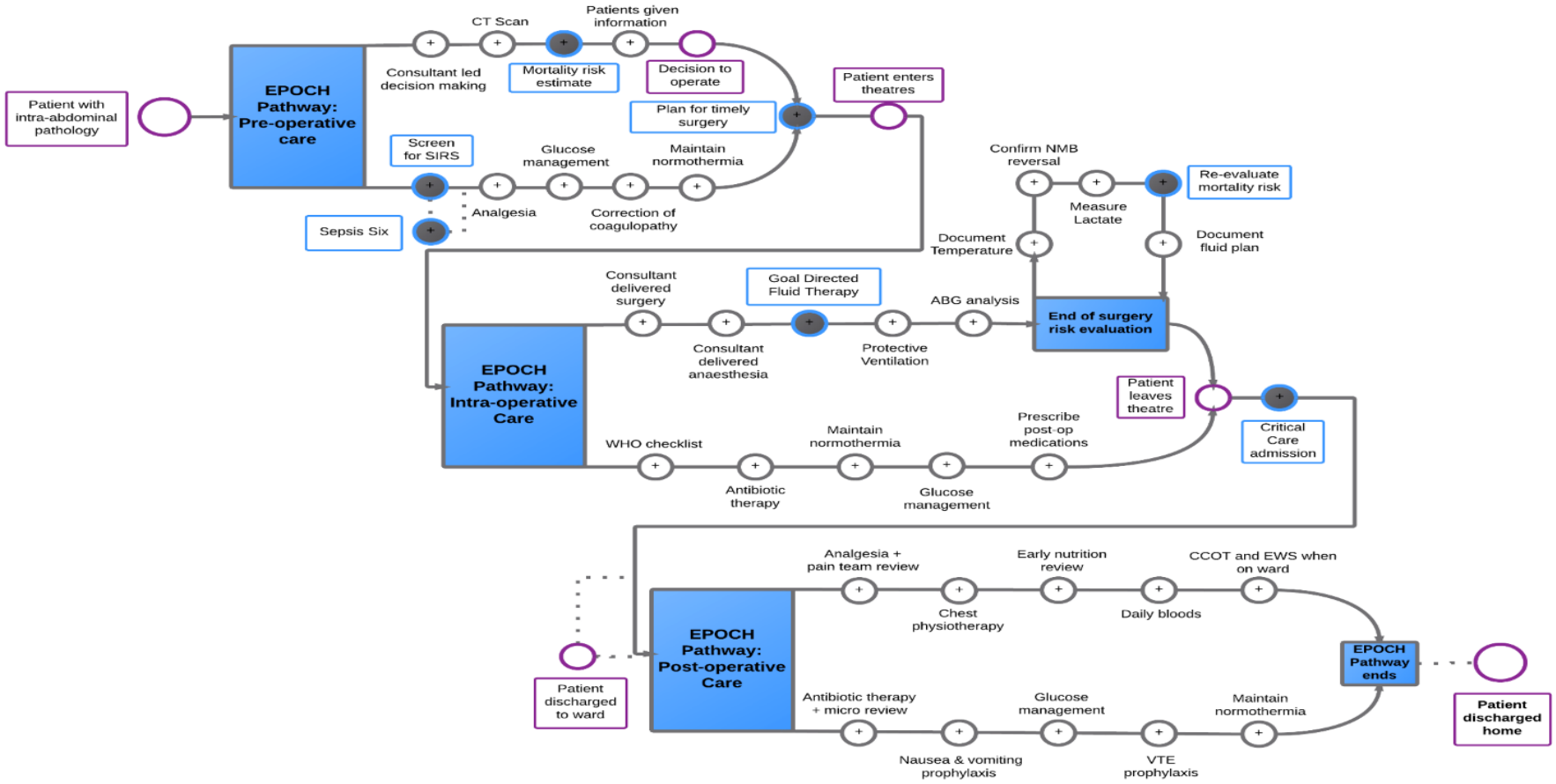



**Figure 2.2. The EPOCH care pathway**

Legend: SIRS, Systemic Inflammatory Response Syndrome; Sepsis Six, a protocolised treatment for sepsis; CT, Computer-aided Tomography; WHO, World Health Organisation; ABG, Arterial Blood Gas; NMB, Neuro-muscular Blockade; CCOT, Critical Care Outreach Team; NEWS, National Early Warning Score; VTE, Venous Thrombo-embolism

### *2.3 The QI intervention, comprising six core improvement strategies*

Carol Peden and I developed the programme theory for the Quality Improvement intervention, defining 'the how' and 'the why' of the QI intervention; see Figure 2.3 and Table 2.1). The EPOCH programme theory was based on current evidence and learning from a range of other Quality Improvement programmes.[60, 91, 102, 103] Six strategies were developed to facilitate the translation of the programme theory into practice by local QI leads at their hospitals. These were intended as a minimum set of activities for QI leads and colleagues to undertake. The strategies were:

1) QI leads hold a stakeholder meeting after activation
2) Each hospital forms an inter-professional improvement team
3) QI leads analyse their own data (NELA data +/- case note reviews and local audit data) and feed back to colleagues regularly
4) QI leads and team members use time-series charts ('run-charts') to inform progress
5) QI leads and team members segment the patient pathway to assist implementation planning
6) QI leads and team members use PDSA cycles to support process change

Table 2.1 details the relationship between the EPOCH programme theory, the QI resources available and the QI strategies proposed.

**Figure 2.3. Summary of programme theory for the EPOCH QI intervention**

***If***

- Relevant data are reviewed and fed back to teams regularly, &
- Key professionals come together to form an improvement team, &
- QI leads and colleagues learn basic quality improvement approaches, &
- Relevant stakeholders are made aware of the project and improvement goals

***Then***

- A shared view of performance and improvement gaps can be created , &
- Professionals can work as a team to define and achieve local improvement goals, &
- Basic quality improvement approaches can be employed to achieve the improvement goals, &
- Stakeholders will be more engaged in the need for change and aware of how improvement will occur

***So that***

- Improvements in care delivery in line with the recommended care pathway can be achieved

***So that***

- Mortality after emergency laparotomy can be reduced.

**Table 2.1: EPOCH Programme Theory: main interventions, desired outcomes and evidence for inclusion**

| Desired outcomes | QI strategies | QI Programme activities and resources | Evidence for inclusion within programme theory |
|---|---|---|---|
| Motivation for change created amongst stakeholders and improvement goals clearly understood | QI leads hold a stakeholder meeting after activation (**QI strategy 1**) | 1. Pre-activation checklist (providing guidance for planning of stakeholder meeting)<br>2. Evidence for QI and need for change provided<br>3. Presentation on achieving engagement | • Improvement projects require attention to the social context in which improvements are to be made which in turn requires relevant stakeholders to be informed and engaged (e.g. evidence from both Michigan Keystone and Enhanced Recovery programmes)[60, 104]<br>• Data feedback can create cognitive dissonance if it is at variance from self-assessed or perceived performance, which in turn can lead to motivation for change.[53] |
| Inter-professional collaboration (IPC) fostered | Each hospital to form an inter-professional improvement team (**QI strategy 2**) | 4. Team approach promoted<br>5. QI leads encouraged to invite colleague to EPOCH meetings<br>6. EPOCH Virtual Learning Environment open to all local QI team members | • There is sound theoretical and empirical evidence for the specific role of clinically-led Quality Improvement teams in successful QI.[65, 105] |
| Shared view of current performance created ('situational awareness') | QI leads analyse their own data (NELA data +/- case note reviews and local audit data) and feed this back to colleagues regularly (**QI strategy 3**) | 7. Case-note review tool<br>8. Training on data for improvement<br>9. Training on how to access and analyse NELA data<br>10. Excel workbook programmed to create run charts from NELA | • Creating situational awareness regarding clinical performance is seen as fundamental to the Model for Improvement,[41] and is the foundation of Feedback Intervention Theory.[53]<br>• Recent empirical data points to data feedback as central to success of several key QI projects.[60, 103, 106, 107]<br>• Cochrane reviews on data feedback indicate a positive |

| Desired outcomes | QI strategies | QI Programme activities and resources | Evidence for inclusion within programme theory |
|---|---|---|---|
| | | data<br>11. Secure data sharing site created on VLE | impact on Quality Improvement if feedback is appropriate and timely and when a path to improvement is proposed.[52, 53] |
| Frontline teams develop and use basic QI skills to effect change | QI leads and other team members:<br>Use time-series charts ('run-charts')<br>(**QI strategy 4**)<br>Segment the patient pathway<br>(**QI strategy 5**)<br>Use the Plan-Do-Study-Act (PDSA) cycles<br>(**QI strategy 6**) | 12. Introduction to QI skills training provided<br>13. Links to further reading and training resources for QI<br>14. Telephone and email support | • Application of improvement science approaches such as the Model for Improvement require at least some basic skill acquisition, and evidence points to a deficit in this area putting significant strain on the ability of an improvement project to achieve its potential.[108, 109]<br>• Time-series charts ('run-charts') are a simple and robust method of analysing and presenting (for data feedback) changes to care processes.[110]<br>• Segmentation of the proposed patient pathway involves introducing interventions within the pathway in an iterative fashion. Pathway segmentation makes the clinical element of this intervention less complex, more compatible with current systems and may makes process changes more trial-able and lower risk.[84]<br>• The Institute for Healthcare Improvement (IHI) Model for Improvement, incl. the PDSA cycle, is an internationally accepted approach to Quality Improvement.[41, 56] |

## *2.3 Discussion and critique of the EPOCH intervention*

In this section I will discuss and critique the development of the interventions 'hard core' (developed by the wider EPOCH investigator team) and its 'soft periphery' (developed my Carol Peden and myself).

*The EPOCH care pathway ('hard core')*

The hard core of the hospital level intervention, the care pathway of recommendations, was an extensive, almost exhaustive inventory of interventions and was based largely and the Delphi consensus process we undertook.[11] Regarding the clinical interventions, the EPOCH pathway drew upon current Department of Health / Royal College of Surgeons of England guidance and was developed through a Delphi consensus process following a systematic literature review with evidence ranging from expert opinion to high-quality randomised trials. Expert opinion has always played a role in evidence-based guidelines. Some interventions (e.g. consultant delivered care) are based on weak evidence but are nevertheless evidence based and strongly recommended.[111] As such the objective value of the individual pathway components is had to critique. The pathway was also designed to adhere to the Cochrane review group definitions of a care pathway.[43, 46] However, on reflection the pathway was perhaps too extensive and too comprehensive, making it hard to determine which interventions may have the highest impact on patient outcomes. There was also a challenge in terms of data collection as only ten of the pathway components were included as process measures in the NELA dataset. Thus, additional efforts would have to be made to collect data to monitor improvement in any component not within the NELA

data-set. This may have been a major flaw in the pathway design and failed to optimally capitalise on the synergy between NELA and the EPOCH trial.

*The QI intervention ('soft periphery')*

Carol Peden and I hypothesised that the complex, evidence-based intervention that we had designed would support implementation of the EPOCH Care Pathway and ultimately reduce mortality for this patient group. Like many complex interventions, it was designed to operate on multiple levels and to be adapted to fit local contexts, required behaviour change by a range of different professionals and needed varying levels of skill to achieve this, in both the recipient and the delivery teams.[96] The intervention drew on current evidence available at the time about 'what works' in Quality Improvement, was conceived with the non-QI expert clinician in mind and designed to be as easy as possible to use to fit within the prevailing UK NHS paradigm of clinician-led improvement. To that end the cluster level intervention was relatively parsimonious, requiring minimal contact time and providing all resources online and by remote email / phone support. As such, we viewed the EPOCH programme and interventions overall as enabling clinicians to implement existing (albeit updated) guidance from the Royal College of Surgeons and The Department of Health. At the time of the EPOCH trial, this had not been put into practice nationally despite widespread appreciation of the high mortality for this patient group and support for change.

The strengths of our intervention development process included the clinician-led programme and relevance of the project with regards to patient care, both of which have

been shown to enhance buy-in amongst clinical teams.[103] The multi-modal training resources (face to face, online and by phone and email) also meant that the training to support use of the intervention was highly accessible. Both the clinical and Quality Improvement interventions were also based upon best available evidence and the rationale underpinning the Quality Improvement intervention – the programme theory – focused and guided the intervention design. However, on reflection, the process of developing the intervention and its final form had several weaknesses.

Firstly, we did not investigate prior to developing the intervention the reasons why the RCS and Department of Health guidance on the pathway of care for these patients had not been widely adopted. At the inception of the EPOCH trial we had solid data on poor patient outcomes and the variability of care delivery across the nation but very little empirical data on the reasons underpinning poor care. As such the root causes of this non-adoption of the pathway, particularly in the context of guidance that was ostensibly widely supported, were not established and therefore not specifically addressed within the EPOCH programme theory. Considering the desired outcomes and the QI intervention and the strategies designed to achieve these (Table 2.1), strategies 1-3 are probably to be considered pre-requisites of any improvement activity in terms of preparation of the local context to be more receptive to and supportive of such efforts. Beyond this however the improvement strategies were what could be defined as 'standard' QI methods at that time (2013-4) with the focus on data feedback and PDSA cycles.

Second, whilst our programme theory (Figure 2.3) outlined how we envisaged the QI strategies would work we did not, beyond this, engage with or incorporate formal theory into our intervention design. On reflection this seems like a missed opportunity. Diffusion or innovation theory, the Consolidated Framework for Implementation Research (CFIR) or the Normalisation Process Theory could all have ben usefully employed to develop an intervention that went beyond standard QI methodology, which is largely atheoretical in its approach.[84, 112, 113] In particular using Diffusion of Innovation theory or the CFIR could have alerted us to multiple intervention and context level factors that we may have wished to consider addressing in the improvement intervention. The MRC guidance on developing and evaluating complex interventions does recommend the use of theory at stage 1 (development stage), [96] and the EPOCH intervention could be considered a (highly) complex intervention. In this sense, flaws at this stage in the trial may have meant that problems and challenges in delivering the intervention were 'designed' into our QI programme. I would consider this a key learning point for my future work in complex intervention design.

Third, the MRC framework also recommends a piloting stage (stage 2) and we did not do this fully during the development of the trial intervention. The Quality Improvement interventions drew on learning and evidence from similar, smaller-scale work, [106] and in the case of the pathway were based upon the best available evidence, but the primary limitation was the lack of pilot trials of this particular set of interventions, in this particular context. Had we tested the feasibility of using the recommended Quality Improvement intervention to implement the extensive care pathway we may have identified multiple

areas to address before moving on to testing it in a more pragmatic cluster randomised trial setting.

# CHAPTER 3: Improving care at scale: Process evaluation of a multi-component quality improvement intervention to reduce mortality after emergency abdominal surgery.

## *3.1 Introduction*

Quality Improvement interventions such as that delivered within the EPOCH trial, are complex due to their interacting components, and the multiple organisational and social levels at which they operate.[72] Delivering a complex intervention into a complex system such as the perioperative care pathway in a hospital is challenging, with many possible barriers to achieving the intended outcomes. Even within a trial setting, such complexity may mean that the target group is not actually exposed to the intervention as planned.[100] Therefore, in addition to the main trial, we conducted a concurrent ethnographic evaluation in six trial sites and a post-hoc process evaluation of the trial overall. There is published guidance on complex intervention reporting (SQUIRE guidelines[66] and TIDieR checklist[100]), but such detailed reporting is not common in the quality improvement literature.[114] In Chapter 2 the development of the interventions and underpinning rationale was detailed. In this chapter, I focus on the process evaluation data to describe how one of the largest trials of a Quality Improvement intervention to date was delivered and received across 93 hospitals that offer emergency abdominal surgery within the UK NHS, and provide detailed analysis to facilitate a greater understanding of the main trial results.

## *3.2 Methods*

We undertook a mixed-methods process evaluation based upon recommended guidance for evaluation of cluster trials[115], and structured using the following framework: how clusters and sites were recruited; delivery of the intervention at the cluster level; response to the intervention at the cluster level; delivery of the intervention at the site level and; the response to the intervention by individuals targeted (in this case, the EPOCH QI leads).

### *3.2.1 Data sources and data collection*

Table 3.1 details the evaluation foci and the data sources used to investigate each. Following commencement of the trial, the variability of engagement with the Quality Improvement programme prompted a wider-scale, post-hoc, process evaluation with the aim of capturing data across the trial cohort. For the post-hoc component of the process evaluation, I collected a range of Quality Improvement programme activity data (see Table 2.1) and sent an exit questionnaire to all QI leads. The 37-item, online questionnaire, administered at the end of the trial, was designed to allow description of activities undertaken as well as their overall experience of leading the improvement projects. The questionnaire comprised categorical, yes/no and free text questions, with opportunities to elaborate on any answers as free text. The questionnaire was piloted multiple times in line with best practice with two rounds of testing using research team members, for readability and usability and a final round of testing using eight QI leads.[116, 117] Changes from this final round were very minor, and therefore responses from this sample were included in the analysis. Only one response was required per hospital, but QI leads were asked to complete the questionnaire with colleagues. I collected Quality Improvement programme data, including data on

participation in programme activities such as meetings and use of the trial Virtual Learning Environment (Table.1).

A pre-planned ethnographic evaluation was undertaken in six trial sites by researchers who I collaborated with, but who were outside the main EPOCH team (GM, DK, see Publications). A maximum variation sample of sites were chosen, with criteria focused on size of the hospital, surgical volume and discipline of the QI lead. Periods of observation were scheduled and interviews with clinicians were held at several points during the trial to monitor progress and reflect on what had been achieved and what had impeded progress. All interviews were audio recorded, and field notes were recorded in a diary at the time of observation, or immediately afterwards. Further details of the ethnographic methods are reported elsewhere.[118]

*3.2.2 Data analysis*

The programme activity and questionnaire data were analysed and reported using descriptive statistics (frequency percentage for categorical data or median range for continuous data). Answers to three free text questions within the questionnaire, designed to stimulate reflection on participation in the Quality Improvement programme and on leading Quality Improvement locally, were analysed using deductive and inductive content analysis.[119] Data were initially managed in Microsoft Excel and coded manually. Another researcher (TA, see Publications) and myself independently generated codes and categories emerging from these data inductively. These were compared and refined through rounds of discussion and sense-making. A set of overarching sub-themes was agreed and used these

as a framework for further, more deductive, coding. Finally, these sub-themes were grouped into high-level themes for each question [29,30]. Themes were discussed with the EPOCH ethnographic team in order to enhance validity and to support the analysis and emerging conclusions; this occurred after analysis of the ethnographic data had been completed but prior to findings being reported to the main trial team. Data analysis of the ethnographic data was initially undertaken by independent researchers (GM and DK, see Publications). These primary analyses are detailed elsewhere.[118] I undertook secondary analysis of these data and was responsible for the integration of the process evaluation datasets. My analysis was predominately deductive, whilst remaining open to new concepts emerging.[119] The deductive analysis framework was structured using the concepts from EPOCH Programme Theory (Figure 2.3) and the literature underpinning this (for example, the role of context on QI projects (Table 2.1)). The logic of using this approach was that we considered the programme theory, and the six Quality Improvement strategies that were based on it, to be the primary mechanism by which the EPOCH pathway would be implemented. As such, how EPOCH site teams interacted with and used those strategies was considered the primary focus of interest, which led to a more deductive approach being chosen.

Data from different sources, as outlined in Table 3.1, were analysed separately and then integrated to meet the evaluation aims. Data analysis from the questionnaire provided a cohort-wide picture of response to the programme and of intervention delivery at site level, with ethnographic data analysis adding granular detail and understanding. Integration was achieved through discussion in data meetings among the investigators responsible for

analysis of the different components (core EPOCH trial team and ethnographic team), identifying points of confluence and apparent contradiction between the data, and particularly focusing on the ways in which insights derived from the ethnographic work might explain or add detail to findings from the survey.

**Table 3.1. Data collected for process evaluation**

Legend: QuIP, Quality Improvement Programme; VLE, Virtual Learning Environment; NELA, National Emergency Laparotomy Audit

| **Aspect of process evaluation** | **Data collection method** | **Data collected and data type** |
|---|---|---|
| Recruitment of sites | Review of trial administrative records | Recruitment strategy incl. inclusion and exclusion criteria<br>Reasons given for non-participation (text in trial administrative documents) |
| Delivery to the clusters | Collation of registers from QuIP meetings (30 meetings)<br>Collation of VLE usage logs | Names, roles and hospital of each of the attendees at the QuIP meetings (2 meetings/cluster)<br>The level of usage of the Virtual Learning Environment (VLE) per hospital, determined by the number of visits / views logged by any staff member from each hospital |
| Response of the clusters | Online exit questionnaire. | Free text responses regarding the positive and negative aspects of the programme |
| Delivery at the site level – QI intervention | Online exit questionnaire | Whether a stakeholder meeting was held (QI strategy 1)<br>Whether a QI team was formed and professional composition of any such team (QI strategy 2)<br>Whether and how data feedback occurred (QI strategy 3)<br>Whether run-charts were used (QI strategy 4)<br>Whether the patient pathway was segmented (QI strategy 5)<br>Whether the PDSA approach was used (QI strategy 6) |
| Response of sites / individuals | Online exit questionnaire. | Free text responses to 2 reflective questions:<br>If you were to be involved in EPOCH again, a) 'what would you continue doing' and b) 'what would you do differently'? |

## *3.3 Results*

Programme activity data, as defined in Table 3.1, were available for all 93 hospitals. 83% (77/93) of QI leads completed the exit questionnaire. All but four responses (73/77) included input from clinicians from the disciplines of anaesthesia or critical care. In comparison 17/77 (22%) of responses included surgical input and 6/77 (8%) included nurse input. The evaluation results are structured using the following framework: delivery of the intervention at the cluster level; response to the intervention at the cluster level; delivery of the intervention at the site level and; the response to the intervention by individuals targeted (the EPOCH QI leads).

### *3.3.1 Recruitment of clusters*

Hospitals were recruited following an open call and promotion through existing critical care and perioperative medicine research networks. All NHS hospitals in the UK (except Northern Ireland) were eligible to take part if emergency general surgery was performed on site and if there was no previous or on-going improvement work focused on emergency laparotomy. There were documented records of 14 hospitals expressing interest but subsequently not participating due to existing improvement work in that site. For other hospitals who expressed an interest but subsequently did not join the trial, the most common reason given was that there was insufficient support from colleagues for the trial interventions.

### *3.3.2 Delivery of the intervention at the cluster level*

A total of 15 face-to-face Quality Improvement educational meetings, planned to coincide with cluster activation, and 15 follow-up meetings (one for each geographical cluster) were held as part of the Quality Improvement programme. Figure 3.1 summarises the EPOCH Quality Improvement programme ‘as planned’ and ‘as delivered’; the major change to the

plan was the addition of follow-up cluster meetings at 12-16 weeks post-activation to the intervention. Aside from local QI leads (surgeons, anaesthetists and critical care physicians), research nurses, theatre nurses and trainees in surgery and anaesthesia were the most common groups to participate in the educational meetings. The number of participants from each hospital at the follow-up cluster meeting was substantially fewer than at the first meeting. Figure 3.2 displays the numbers of QI leads attending the meetings from each hospital. The median number of participants (both QI leads and other invited colleagues) at the educational meetings and follow-up meetings were three per hospital (range 0-19) and one per hospital (range 0-8) respectively.

The web-based resources were housed within a Virtual Learning Environment which contained a total of 66 pages or resources, to be viewed online or downloaded, at the commencement of the programme, increasing to 84 pages or resources by the end of the study. The site could only be accessed by registered EPOCH trial local QI team members. In total, 16,120 'hits' (visits to the site, page view and resource views or downloads) were logged over the course of the trial period. The median number of Virtual Learning Environment hits per hospital was 136 (min 11, max 519; IQR = 123). The number of users per hospital ranged from one to seven with a median of three users but site teams were small (see below) and on-line resources could be downloaded and / or printed by one user for colleagues to view. As such the total hits per hospital is a more useful metric of Virtual Learning Environment usage. Given the number of pages and resources available (84 by the study end), these data suggest likely appropriate usage by much of the cohort but with some variability and a substantial minority of low users.

*3.3.3 Response to the intervention at cluster level*

Themes derived from responses to a free text question in the exit questionnaire about the improvement programme are described in Table 3.2. Themes emerged pointing to the utility of the meetings, both for learning and for networking, the overall support offered and energy for change generated by the programme team and the helpfulness of the run-chart tool. Conversely themes emerged from comments regarding the need for greater clarity about, and fewer components within, the intervention and more meetings and input together with more time in the intervention period. QI leads were also asked to rate the support they received from the Quality Improvement programme team on a five-point scale (very good to very poor). Of the 75 who responded to this question, 36/75 (48%) rated support as very good, 30/75 (40%) as good and 9/75 (12%) as average.

Findings from the ethnographic evaluation mirror the themes described in Table 3.2, indicating that participants had a positive perception of the EPOCH cluster activation meetings, as well as the 12-week follow-up meetings. Participants felt that the EPOCH Quality Improvement team demonstrated the relevance of the project and felt energised by the meetings. They also reflected positively on the practical nature of the meetings, the opportunity to share ideas and learn from others, and the utility of the web-based resources and tools to analyse NELA data. Analysis of the ethnographic data indicated that buy-in from QI leads was often already high and many had achieved local improvements relevant to EPOCH's mission long before the activation meetings. Nonetheless, even for those individuals, the activation meeting was an important place for learning and sharing experiences. It was important for local enthusiasts both to see that they 'were not alone' in struggling to improve peri-operative care and to learn how other sites managed to change

aspects of care. However, themes derived from the questionnaire data indicate that satisfaction with the Quality Improvement tools was more mixed, in particular the run charts to support data analysis and visualisation and the guidance on how to improve care in line with care pathway.

**Table 3.2. Common themes identified from feedback**

| *"What was most helpful about the QI programme" (from 56 free text responses)* | *"What could have been better about the QI programme" (from 36 free text responses)* |
|---|---|
| QI training (at the meetings) and online resources (n=14) | More clarity about the intervention and how to implement it (n=10) |
| Networking with colleagues from other hospitals (facilitated by meetings) (n=11) | More meetings, and more input from the central team (n=8) |
| Good communication and support (n=12) | Better support / better run chart tool (n=7) |
| The Excel tool to generate run-charts from National Emergency Laparotomy Audit (NELA) data (n=11) | A longer intervention period for those activated late (due to the stepped wedge trial design) (n=7) |
| Enthusiasm and motivation generated by the EPOCH team and project overall (n=8) | Fewer components in the clinical pathway (n=4) |

**Figure 3.1. The quality improvement programme as planned and as delivered**

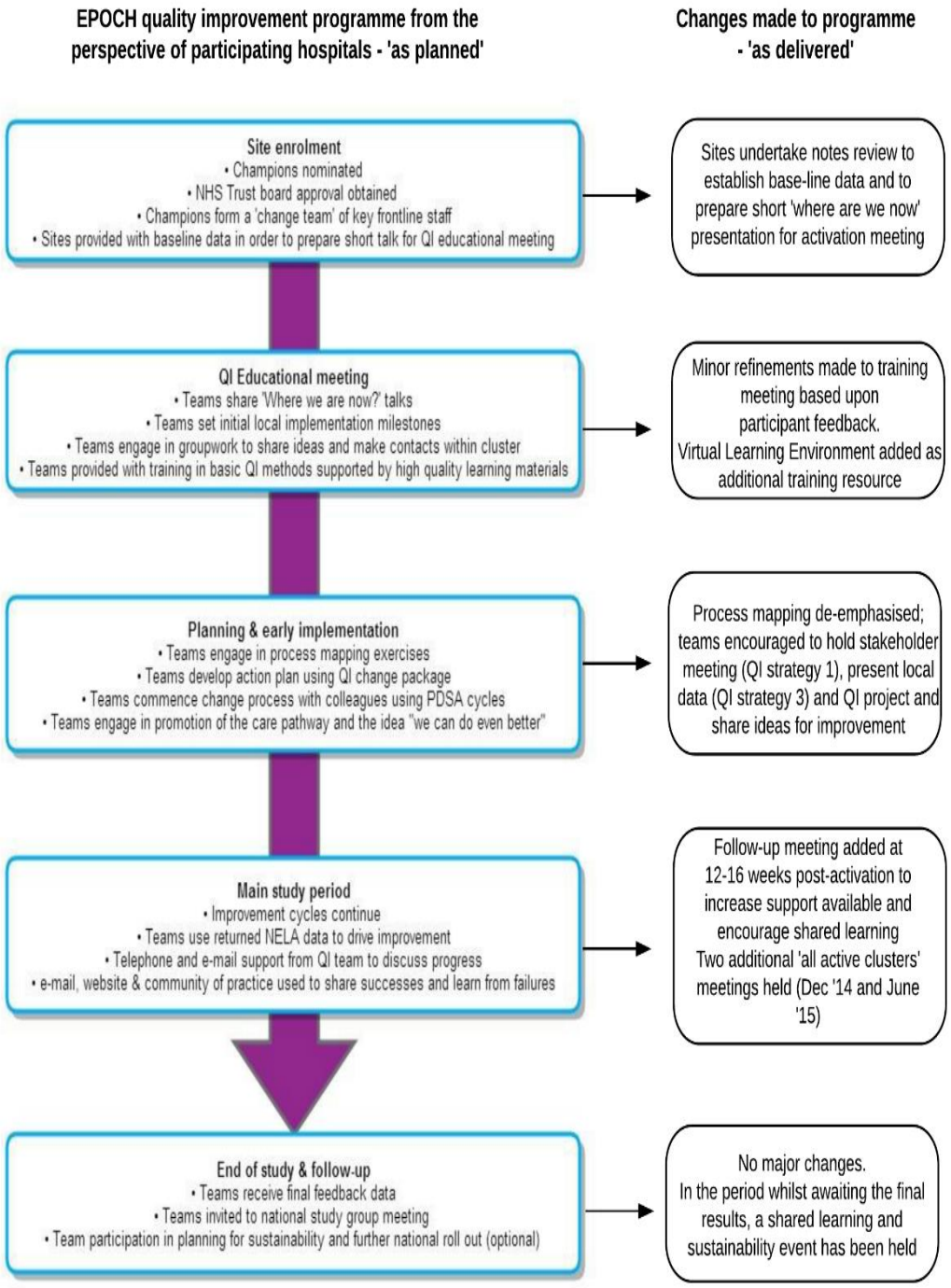


**Figure 3.2. QI lead attendance at QI meetings**

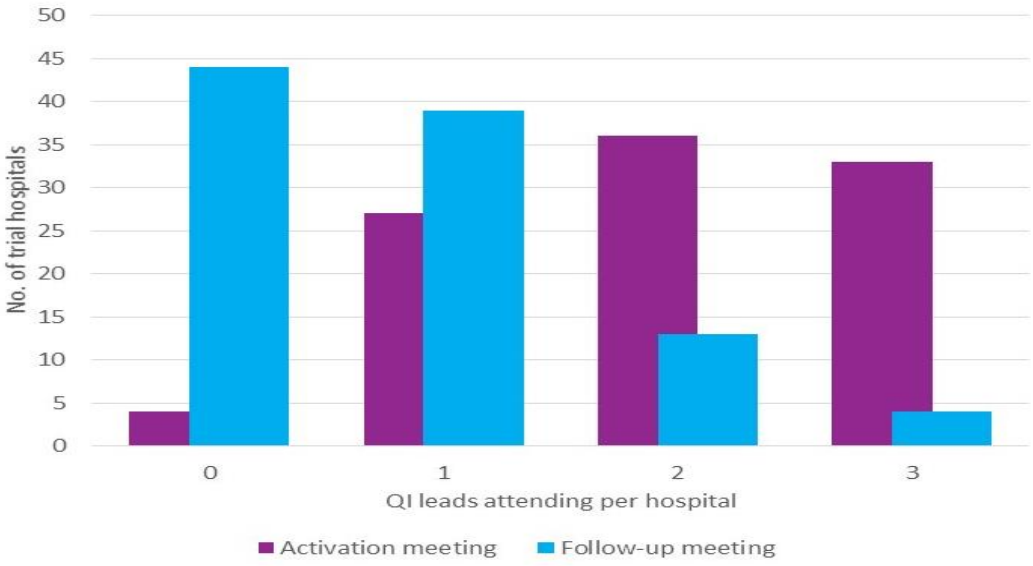


**Table 3.3 Reported usage of each Quality Improvement strategy**

| Question related to QI strategy usage | Response (n = variable) |
|---|---|
| **PDSA approach**<br>Did you or your colleagues use the PDSA cycle approach during your QI activities? | 61% (45/74) Yes, sometimes<br>5% (4/74) Yes, often<br>34% (25/74) No |
| | |
| **QI team formation**<br>At your site, was a formal team created to work on QI activities related to EPOCH?<br>Definition of QI Team:<br>*A group of individuals that work together on the QI project. The team is defined by their shared goals and mutual accountability for the QI* | 60% (46/77) Yes<br>27% (21/77) No<br>13% (10/77) Other (comments indicated informal teams often existed) |
| | |
| **Data collection and analysis**<br>After starting EPOCH did you or your colleagues download and analyse your local NELA data?<br><br>If yes, how frequently did you do this?<br><br>If yes, did you use run charts?<br><br>Were systems set up to collect NELA data prospectively? | 79% (61/77) Yes<br>21% (16/77) No<br><br>43% (26/61) : Analysing data monthly or bi-monthly<br>57% (35/61) Analysing data less frequently<br><br>92% (56/61) Used run charts to analyse data<br><br>51% (38/74) Yes<br>49% (36/74) No |
| | |
| **Stakeholder meeting**<br>Did you hold a stakeholder meeting as one of your QI activities? E.g. a meeting for all professionals involved in patient care | 55% (41/75) Yes<br>45% (34/75) No |
| | |
| **Pathway segmentation**<br>Please indicate statement most closely fits your hospitals improvement or implementation activity during EPOCH | 22% (17/77) We introduced a single pathway of care (across Pre, Intra and Post-operative phases)<br>32% (25/77) We introduced separate pathways or care bundles for the perioperative phases<br>40% (31/77) We focused on introducing individual / separate interventions<br>5% (4/77) Other |

*3.3.4 Delivery to individuals at local site (QI intervention fidelity)*

The clinical intervention was a 37-component care pathway (see Figure 2.1). Questionnaire data showed that, regarding the care pathway, only 11 care processes were the focus of improvement efforts in more than 50% of responding hospitals; the remaining pathway components had more variable uptake (Figure 4.4 and Segmentation section below). The Quality Improvement intervention comprised six strategies (Figure 2.2 and Table 2.1), as described by the intervention programme theory (Figure 2.3). The programme theory was developed with the assumption that these strategies would enable QI leads to make the clinical improvements recommended by the EPOCH care pathway. Using the QI strategies recommended was therefore the main mechanism through which the intervention would lead to improvement in care processes and there was an expectation that QI leads would undertake these activities, as a minimum, at their sites. Questionnaire data showed that 10/77 (13%) of QI leads responding said that all six strategies had been used, 23/77 (30%) indicated five had been used, 21/77 (27%) indicated four had been used, 8/77 (10%) used three strategies, 10/77 (13%) used two and 5/77 (6%) just one. No QI lead reported zero quality improvement strategy usage. Below, questionnaire and ethnographic data are combined to elaborate on the usage of each of these strategies and the effects of these on care pathway implementation.

*Use of the Plan-Do-Study-Act (PDSA) cycle*

At activation meetings, the use of PDSA cycles was presented to participating teams explicitly as a model for experimentation and the planning of change, with instructions and supporting tools for putting it into practice. The data in Table 4.3 indicates this approach was used, but perhaps not in the regular, methodical manner recommended. The

ethnographic findings also indicated that no site applied the formal PDSA methodology 'by the book'. However, this did not mean sites failed to engage in creative experimentation. Instead, sites adopted a less formal planning approach, which included the general tenets of trying out small tests, reviewing and making further change, but typically excluding the setting of numerical goals against which to measure progress:

"*The only thing is we are not being particularly good at is the PDSA cycle but then again […] Well I suppose we are. We are just not doing it formally [...] I have carried on and done it in a way that works and makes sense to me*." (Intensive Care Consultant, Site 6)

*Using a team approach*

At the activation meetings, QI leads were strongly advised to recruit a formal team of 'willing' inter-professional colleagues to work with them on local improvement activities. The data in Table 3.3 indicate that just under two-thirds of sites had a formal team to work on this major project. All sites had committed to an inter-professional team approach by formally nominating representatives from surgery, anaesthesia, and critical care; for those who managed to recruit others to their team, the benefits were apparent:

"*The really important thing was that we had a group, from our point of view, I've got an engaged surgeon who I work with, and I've got some good junior guys, and we've got plenty of people who've actually just taken the ball and run with it […] So possibly we should be*

*involving others but the small team we have at the moment has been quite productive and we seem to be hitting most of the QI targets with the team we have got.*" (Intensive Care Consultant, Site 5)

However, only three of the six sites included in the ethnographic work maintained surgical leadership throughout the intervention period; in two sites surgical involvement in the QI team decreased after activation, and in the other site surgical involvement did not become apparent until later on. Unsurprisingly, in these sites, lack of a surgical QI lead was seen as a disadvantage to wider surgical involvement with the improvement work (see also 'Engagement' below):

"*It started as an anaesthetic project basically but it is really a surgery thing. […] Looking back I wish we took advantage of [having an engaged surgical lead] right at the beginning. I think we would have got more involvement with the surgeons which is obvious because they are the thing that runs right through it all.*" (Research Nurse, Site 1)

*Use of data feedback and run-charts*

At the activation meetings, use of NELA data as a driver for engaging colleagues and monitoring improvement was promoted and tools designed for the EPOCH project were provided to analyse the data. Table 3.3 shows that most, but not all, teams analysed their NELA data occasionally, but far fewer were doing this on a regular (monthly/bi-monthly) basis. Many sites reported challenges in simply collecting the data; only half of questionnaire respondents indicated that systems had been set up by the end of the EPOCH

study to collect NELA audit data prospectively. For the other half of respondents, it was reported data collection usually involved the NELA lead (often also an EPOCH QI lead) collecting and entering data retrospectively:

"*We need to look at the recent outcome of the NELA. But we haven't, because we were concentrating on NELA [data collection] and less on the EPOCH care pathway, we haven't been able to monitor that unfortunately*." (Research Nurse, Site 1)

The ethnographic findings indicated that all six sites tried hard to collect and use data in their improvement efforts. However, this was undertaken more consistently in three of the six sites. During the implementation process, the EPOCH teams that seemed more successful with data collection were also those that appeared to have achieved stronger engagement with colleagues (see section below also). This perhaps reflects the challenges of collecting the large NELA data set before any analysis, or improvement activities based upon it, could occur:

"*Well there is a nominal person in charge [of the NELA audit] but in terms of actual, the whole thing is devolved back round to the anaesthetic department. Well we try and get everything done, as far as possible, doing it in the operating theatre to engage the surgeons, as part of that process. Even if they only do data entry on one page, or even if we only discuss it, and one of us will do the data entry*". (Intensive Care Consultant, Site 3)

*Engagement*

At five weeks before activation to the intervention, sites were contacted and asked to start planning a stakeholder meeting, to coincide with activation, to engage relevant colleagues with the aims of the trial intervention and the required improvements. Just over half the respondents indicated they had held such a meeting (see Table 3.3). Of the 71 QI leads who responded to a question about senior support during the trial, only 15 (21%) described active executive board support for the Quality Improvement work related to EPOCH (e.g. funding staff time to support the project or making the project a board-level quality and safety priority). The ethnographic study allowed observation of the ongoing engagement activities that occurred beyond the initial EPOCH meetings. When local teams drew on wider connections, this appeared to work to their advantage, pulling in contacts in management, other disciplines such as radiology, and clinicians and administrators with responsibilities relevant to the pathway, for example sepsis identification and treatment. The ability to engage colleagues successfully, and encourage active involvement in improvement efforts, seemed to depend to a large extent on existing relationships.

*"I think, you know, we're fairly cohesive, we have a cohesive department, and we're not perfect, but we do. We don't have any personality clashes that get in the way of this at the moment...We've had no problem with the surgical engagement and have had no problem with the anaesthetic engagement either.*" (Intensive Care Consultant, Site 5)

Even in sites where engagement per se was not seen to be a problem, the simple factor of the time required to have the required discussions with colleagues was raised as an issue,

*"I think a longer period of time would have helped because most of these changes are by default, sort of long-term changes, but also there is a lot of discussion involved with them all and getting a lot of people to agree and of course each of those conversations, despite the fact that you think it is going to be quick, ends up going back to someone else and then a week passes and another week passes and before you know it a month and a half has gone and you have finally got to the conversation you wanted to have in the first place."*

(Consultant Anaesthetist, Site 6)

*Segmenting the pathway and decisions about the clinical pathway components*

At the activation meeting, QI leads were advised to consider segmenting the proposed pathway to make the workload of implementation more manageable. Advice was offered regarding selecting which elements of the pathway to work on first and how to plan a stepwise implementation of the pathway that would work in their local context. However, by the end of the intervention period only a fifth of questionnaire respondents (17/77), said that they had attempted full pathway intervention. Of the potential 37 pathway components, there were 11 interventions which more than 50% of respondents said had been the focus of improvement efforts (Figure 3.3). There were: Consultant-led decision-making; Pre-operative formal risk assessment; Screening for Sepsis; Timely surgery; Consultant-delivered surgery; Consultant-delivered anaesthesia; Goal Directed Fluid Therapy intra-operatively; Intra-operative arterial blood gas and serum lactate analysis; Mortality re-assessment at the end of surgery; Re-check the serum lactate at the end of surgery; Plan for critical care admission. Except for 'screening for sepsis' and 'Intra-operative

serum lactate analysis' these processes were the main EPOCH trial process measures that were captured by NELA. Eight of these 11 processes were also those captured by NELA and were the same as the main EPOCH trial process measures.

The ethnographic analysis suggests that agreement on the need for a pathway for this patient group was strong amongst QI leads and colleagues. Implementation challenges were predicted, however, which shaped decisions about the initial focus for improvement. These decisions were made as pragmatic choices, based on a tension between what was felt to be most important to improve versus what was manageable within work constraints,

"*…the surgeons and the anaesthetists and* [the Principle Investigator]*, they picked what they thought would be their top ten [from the EPOCH pathway] that we would want to institute because we thought if we tried to introduce all 30 in one go, the resistance that we would be up against would be quite difficult[…] so we picked what we thought were the most important ones."* (Junior Doctor, Site 4)

The idea of a stepwise approach resonated with teams, with the hope that initial success would pave the way for further pathway components to be addressed:

"*The ideal that we are aiming for would be to have all of the 37 (pathway) points done consistently for everybody…although the way that I think we have approached it is to cater for the ones that are perhaps easier to understand and implement…then on the back of those introduce the rest of them."* (Junior Doctor, Site 1)

Some other decisions came down to components of the pathway being seen as having more marginal benefits by some QI leads,

"*I think there were some bits that we talked about before about the inter-operative delivery so things like how you ventilate people and things like that that we didn't necessarily want to have the argument about […] we might cross that bridge later but that wasn't one of our first aims."* (Consultant Anaesthetist, Site 6)

As mentioned above, where teams did not include all clinical leads in equal leadership roles, decisions about which processes to improve often depended on which discipline was most active in the EPOCH team.

This stepwise, segmentation approach was not universally adopted however:

"[*The] endpoint is reduced mortality and reduced morbidity for emergency laparotomy patients. My view would be, look, we really don't know, just do the whole bloomin' lot and then see what happens*," (Consultant Surgeon, Site 2)

In this site their main implementation tool was thus an extensive checklist which brought the EPOCH pathway together. But by the end of the trial, they were still discussing the need to "implement the checklist"; progress had not been as rapid as they had hoped.

*3.3.5 Response of QI leads: Reflections on the change process*

QI leads reflected on: *'what would you continue doing?'* and *'what would you do differently if you were to do EPOCH again?' Overall,* 96% (74/77) of respondents left a total of 299 comments. Eighteen themes were generated for each question (36 in total) and these were further grouped into nine high-level themes (Figure 3.4). Two clear themes, from responses to both questions, were related to the importance of effective engagement and involvement of colleagues (themes two and six) and to data collection and feedback systems (themes one and seven). Other reflections on what QI leads 'would continue doing', related to quality improvement methodology (themes three and five) and the utility of specific, recommended clinical interventions, particularly mortality risk estimate scoring (theme four). When considering 'what they would do differently' QI leads also highlighted some of the 'real life' challenges of delivering quality improvement at the frontline; developing leadership and project management skills (theme nine) and the need to obtain strong senior support (theme eight).

*3.3.6 Context*

Limited resources, both human and financial, and organisational upheaval were often mentioned, in particular in Ethnographic Site 3, although it is likely that this experience was shared by a significant sub-set of the 93 hospitals in the trial. Across almost half the trial sites, a lack of organisational support for data collection was noted. The challenges this posed for QI leads must not be underestimated, with the burden of collecting data (for NELA and ostensibly for use as part of the EPOCH improvement work) may have overwhelmed many. As mentioned above, teams often wanted to do more but struggled to find time:

"*Again, it's finding the time to do all this stuff…the trust hasn't given anyone any time for this, so people are doing it, you know, because they want to. So, you know, it would help if it had funded time for it, but you know that's never going to happen in the NHS […] not at the moment*". (Intensive Care Consultant, Site 5)

**Figure 3.3. Clinical process change attempted during intervention period**

**Figure 3.4. Themes emerging from inductive content analysis of QI leads responses**

| WHAT QI LEADS WOULD CONTINUE DOING | | WHAT QI LEADS WOULD DO DIFFERENTLY | |
|---|---|---|---|
| **High-level themes** | **Sub-themes (Number of supporting comments)** | **High-level themes** | **Sub-themes (Number of supporting comments)** |
| 1. Keep working on data collection and feedback | Providing feedback on performance, incl. data feedback (30) | 6. Engage and involve people more effectively | Wider engagement of stakeholders (17) |
| | Use run-charts (19) | | More surgical engagement / involvement in project (15) |
| | Good data collection process / data collection support (14) | | More interprofessional involvement (10) |
| | Using data to create situational awareness (4) | | Better engagement / involvement of trainees (6) |
| 2. Keep working on engagement, involvement and collaboration | Engage / involve all relevant stakeholders (22) | | Form a larger QI team (5) |
| | Interprofessional involvement (9) | | Involve more people (3) |
| | Form a QI team (8) | 7. Get data collection and feedback right | Improve data collection / More data support (17) |
| | Engage / involve trainees (4) | | More data feedback (8) |
| | Identify enthusiastic colleagues (4) | | More data analysis (4) |
| | Collaborate with other hospitals (2) | 8. Obtain stronger senior support for the project | Stronger senior leadership / board level support (16) |
| | Obtain senior support for the project (2) | | More protected time for the project (7) |
| 3. Using a "systems thinking" approach to improvement | Hardwire changes into system (9) | 9. Work on own leadership / project management skills | Manage the QI team more effectively (10) |
| | Building risk scoring into care pathway (8) | | Get started sooner (6) |
| | Use a checklist / care bundle approach (2) | | Be more forceful (3) |
| 4. Specific clinical interventions | Clinical interventions (9) | | Focus on motivation / behaviour change (2) |
| | Risk stratification (6) | | Use an iterative approach (2) |
| 5. Use an iterative approach to change | Take an incremental / stepped approach to improvement (6) | | More collaboration with other hospitals (2) |
| | Persist with implementation (2) | | Better planning of improvements / system changes (2) |

### *3.4 Discussion*

The principal finding of this process evaluation was that the Quality Improvement programme delivered the Quality Improvement skills training and resources as intended and the programme was generally well received by QI leads. Local adaptation to both the Quality Improvement and clinical interventions was actively encouraged, but the extent of variability and adaptation in the implementation process was greater than anticipated, particularly in relation to the clinical processes. There were only 11 clinical processes which more than half of teams attempted to improve from the clinical pathway (the hard core of the intervention) and only half of the trial cohort reported using five or all six of the QI strategies (the soft periphery of the intervention) designed to enable pathway implementation. The main trial results showed no effect of the intervention on patient outcomes or care processes (Chapter 3). Our experience during the QI programme (meeting teams, reviewing their data) suggests that some hospitals were able to make modest, and sometimes substantial, improvements in care processes, but the main trial analysis was not designed to provide this level of granularity. However, no clear signal toward substantial improvement of care processes was seen across the whole cohort as a result of the EPOCH QI Programme.

When testing clinical interventions within a clinical trial, it is important to make the distinction between the design of the intervention and the operational elements required for effective delivery.[120] This process evaluation adds to the main trial findings by providing insight into the challenges at both the design (or programme) level, and the hospital (operational) level. At the design level, adaptability is often essential in ensuring that quality

improvement interventions can fit within different contexts, and this was built into the EPOCH intervention. However, fidelity to key parts of an intervention are also important to maximise likelihood of success.[121] In this case, as reflected on in Chapter 2 it may have been that an intervention design that focussed on a smaller number of strategies, all with available data via NELA, would have achieved greater fidelity and, therefore, greater impact on patient outcomes. This may be especially relevant given that data from both the ethnography (see Chapter 5) and the exit questionnaire suggest that, at the operational level, QI leads faced many local challenges including lack of engagement of colleagues and hospital executives.

Data was also an operational challenge for many. NELA had only commenced four months before the start of the trial; 20 months after the launch of NELA, at the end of this study, only half of hospitals reported having prospective data collection systems in place. It is likely therefore that many QI leads were focused on collecting and inputting data to the detriment of other improvement activity. If future QI programmes are to capitalise on concurrent national audits or other ongoing data collection, the timings need to be considered to allow embedding of data collection processes before the start of the improvement work which may take considerably longer than anticipated. Moreover, our findings suggest that the target interventions and the collected data need to be aligned. In EPOCH we had a 37-component care pathway, of which only ten had an associated NELA process measure. These ten, having been chosen by both the EPOCH trial team and the NELA team, were likely the highest impact interventions. From a QI point of view, they were also the only interventions for which the QI leads had readily available data. A question, for which there is

currently no empirical data from this evaluation, is whether EPOCH QI leads focus predominately on the interventions with process data because they were most important or because of the data? Ideally, as was the case here it is likely that these interventions were the most impactful, yet aligning data collection more carefully to the intervention may be important point to consider for future QI programmes.

A key theme from the reflections of QI Leads was that they would have liked to have had better mechanisms, not only for data collection, but also for data feedback. Whilst data is central to any quality improvement project, it is the use of this data through feedback, often combined with other improvement strategies, that is likely to achieve more robust results.[52, 60, 103] Even if audit and feedback is the only improvement strategy to be used, we now know there are ways in which feedback occurs that may make it more effective. These include: feedback may be more effective when baseline performance is low, the source is a supervisor or colleague, it is provided more than once, it is delivered in both verbal and written formats, and when it includes both explicit targets and an action plan.[52] However, at the time of the EPOCH trial, audit and feedback was an accepted improvement intervention that was proven to be relatively effective, yet the theoretical mechanisms by which it worked had not been well articulated. The key issue here however is that audit and feedback is designed to invoke *behaviour* change in healthcare professionals (if delivered optimally). Yet not all quality problems are solely amenable to behaviour change but require a process or system level approach. A good example might be getting the patient to the operating theatre in a timely fashion. Audit and feedback can help surgeons to drive the process to try and achieve this but there will likely be multiple system level factors that

enable or hinder the achievement of this. Giving the surgeons more feedback on the poor performance of that measures will, by itself, only get an improvement programme so far. Taking the example further, it is likely that problems with theatre access may also have structural elements (lack of enough staff, theatres etc) beyond even the scope of process improvement. Audit and feedback in this situation will be even less effective. Ultimately, audit and feedback activities alone cannot solve built in and often longstanding system problems but can be a useful tool to emphasise priorities for change, inform focused actions, and evaluate progress.

There are other plausible explanations for our failure to change the primary outcome metrics. It is possible that our programme theory was incorrect, and there was only a weak causal link between the interventions and ultimate outcomes. This seems unlikely given the evidence base for the clinical and Quality Improvement interventions. Another conclusion that might reasonably be drawn from our evaluation is that the EPOCH trial intervention was too ambitious. Even where QI leads developed the capabilities to enable change (e.g. through use of the QI strategies) they were asked to lead that change in addition to their regular clinical commitments and may not have had the capacity, in terms of time, resources and other personnel, to do so. The social aspects of improvement are as likely to be as important as more technical aspects, such as data analysis and feedback, but QI leads used the socially orientated QI strategies less than those related to data. Building and maintaining effective social relationships is time-consuming and challenging and the uptake of 'non-technical' and 'socio-adaptive' interventions can be low amongst health professionals.[91] However, a key reflection of QI leads was that they would have liked to spend more time

engaging and involving colleagues. More emphasis and training in socio-adaptive interventions should be built into future programmes together with a recognition that dedicated time is required to support frontline staff in prioritising such interventions.[108] Some leads reflected on their difficulties in engaging with senior or executive-level colleagues and only a fifth of respondents indicated they received active support from their board. Effective quality improvement requires a reciprocal relationship between the employee and the organisation, and lack of organisational support is likely to have been an important barrier to improvement.[122] This is an important lesson; if the goodwill and motivation of frontline staff is to be mobilised for improvement work, then adequate time and support in the workplace plus training is required to give these professionals the best chance of success. This has ramifications for those designing future programmes, senior management and national-level policy makers.

In relation to the delivery of the programme, the time available to coach teams was limited in comparison with other reported Quality Improvement interventions, such as the Institute for Healthcare Improvement Breakthrough Series Collaborative model.[78] Our training programme was designed as a parsimonious intervention, with face-to-face meetings limited, so that it might be adapted and replicated widely if proven successful. A higher-intensity programme might have led to greater intervention fidelity, although recent evidence suggests that this may not always be the case.[78, 123] EPOCH may have suffered from the lack of a pilot trial and future similar interventions should be rigorously piloted first,[96] or use a cluster trial design that allows for iterative intervention development within the trial period to enable ongoing intervention optimisation.[124]

A major strength of this evaluation is that it provides a full, detailed description of how a large-scale trial of a complex intervention was designed, delivered and received, at over half the hospitals in the NHS. The evaluation has provided insights into likely reasons why ultimately the trial was unsuccessful and learning for future studies of this nature, in line with what is considered good practice for complete interventional research reporting.[66, 100] The evaluation was conducted by researchers both inside and outside the main trial team, offering both detailed, nuanced knowledge of the trial and an external perspective; all data collection and analysis was completed before the trial results were known.

This study also has several limitations. The process evaluation relied in part on self-reported data, often collected from a single representative of each hospital. A response rate of 83% suggests that our data were largely representative of the entire EPOCH trial cohort. However, because non-responders may have had different experiences with the EPOCH programme, it is possible that some relevant factors may be missing. Self-reported data may be subject to both recall and/or social desirability bias. To minimise recall bias, we started collecting data within a month of the completion of the trial. While the magnitude of potential social desirability bias cannot easily be quantified, many respondents reported both positive and negative experiences and many reported not using several of the quality improvement strategies.

### *3.5 Conclusions*

Programmes designed to support clinician-led improvement may need to focus on both developing the necessary QI capabilities, whilst also advocating (or even mandating) clear organisational support for these professionals to lead change. Additional capacity, including job-planned time to engage stakeholders plus data support and / or adequate data collection mechanisms, are likely prerequisites for the successful delivery of complex interventions, such as implementing a care pathway for emergency surgery.

# Chapter 4: Hospital level evaluation of the effect of a national quality improvement programme: Time-series analysis of registry data

### *4.1 Introduction*

As described in Chapter 2, the EPOCH trial was performed to test whether a national Quality Improvement (QI) programme to implement a care pathway could reduce 90-day mortality following emergency abdominal surgery. The main analyses were designed to evaluate the impact of the QI programme across a large cohort of 93 NHS hospitals, leveraging the large sample size to adequately power the trial. A different perspective is to view the EPOCH trial QI programme as an enabling factor in 93 separate hospital-level QI projects. The impact of local context on the effectiveness of QI efforts is increasingly understood, especially in relation to complex intervention delivery.[70, 71, 125] As described in Chapter 3, there was an wide observed variation in the approaches taken to implement the care-pathway, including differing ways of engaging colleagues and decisions regarding which parts of the pathway to implement first, as well variations in the challenges faced. Given the level of heterogeneity across participating hospitals, an analysis designed to understand changes in care processes at the individual hospital level was also needed.

In this chapter, I explore how a form of simple time-series chart (the run-chart) might enable detailed hospital level analysis of process change over time when system improvements are attempted at a national level.[110, 126, 127] The primary objectives of this study were: 1) to evaluate, at the individual hospital level, whether participation in the EPOCH trial QI programme led to implementation of the EPOCH care-pathway and 2) to assess the

relationship between care-pathway implementation and use of the implementation strategies. Our secondary objectives were: 1) to describe the number of improvements in care processes overall and 2) to describe which care processes were most commonly improved (or potentially degraded).

## *4.2 Methods*

This was a prospectively designed time series analysis of registry data provided by hospitals participating in the EPOCH trial, a stepped-wedge cluster randomised trial across 93 UK National Health Service (NHS) hospitals. The registry was the National Emergency Laparotomy Audit (NELA), funded separately by the UK Healthcare Quality Improvement Programme, which started collecting individual patient data on 1st December 2013.

### *4.2.1 Main trial results*

The main EPOCH trial analysis found no effect on the interventions upon any of the trial outcomes measures; 90-day risk adjusted mortality, length of hospital stay or hospital readmission. Analysis of trial process measures (see below) suggested little improvement had occurred as a result of the intervention across the entire cohort. These results did not differ significantly between hospitals activated earlier in the stepped-wedge design compared with those activated later.

### *4.2.2 The EPOCH care pathway and implementation strategies*

Details of the 37-component evidence-based care pathway and Quality Improvement programme are provided in Chapters 2 & 3. Six specific implementation strategies were developed to facilitate care-pathway implementation (see Table 2.1).

*4.2.3 Data collection*

Data were collected through the NELA database (www.nela.org.uk). Inclusion and exclusion criteria for these analyses were the same as for the main trial (see Appendix). We pre-defined a longer data collection period than the main trial, so that data from the 1st January 2014 to 31st March 2016 (six months following the end of the EPOCH trial) were analysed. The rationale for this is that the shift rule requires at least six data points (i.e. six months of data, see below) for change to be demonstrated. There is also evidence that the effects of Quality Improvement may take longer than expected to show.[128, 129] Therefore, a six-month wash-out period was included, to provide clusters activated later in the trial adequate opportunity to demonstrate improvement. Data from our process evaluation questionnaire were used to to quantify recommended implementation strategy use in each hospital (see Table 2.1 and Figure 3.3). 77/93 (83%) of QI leads completed the exit questionnaire. For this study, we used binary responses related to implementation strategy usage, e.g. “we did / did not form a QI team”.

*4.2.4 Process measures*

Process measures in this study are the same as those in the main trial, but now analysed at the individual hospital level rather than in a pooled analysis. The ten key care processes of the EPOCH care-pathway for which process measure data was available via the NELA dataset were: 1) Consultant-led decision making; 2) Consultant review of patient before

surgery; 3) Pre-op mortality risk documented; 4) Time from decision to operate to entrance to the operating theatre; 5) Entry to operating theatre within NCEPOD target timeframe; 6) Consultant delivered surgery; 7) Consultant delivered anaesthesia; 8) Cardiac output monitoring to guide fluid therapy; 9) Measurement of serum lactate intra-operatively; and 10) Admission to critical care post-operatively.

*4.2.5 Data analysis*

Process measure data were analysed for each hospital. Data for each calendar month were pooled and the mean calculated and plotted onto run charts, using a pre-programmed Excel Workbook designed specifically for the EPOCH trial (see Figure 4.1 for a worked example). A baseline median was constructed with the first ten data points (January 2014–October 2014), or from January 2014 up to and including the month of trial cluster activation, whichever provided the longer baseline period. To increase the likelihood that any signals identified in the run charts were associated with the EPOCH trial, and not pre-existing improvement efforts (such as involvement in NELA), each hospital's baseline median was assessed for signals using the run chart rules. In particular the 'runs rule' was used to identify potential improvements in patient care processes before the improvement intervention started (for the reference chart for this, see[110] ). In line with recommended practice, if no signals were seen, the baseline median was fixed and extended forward creating the centre-line for all data points on the chart, to facilitate analysis of signals over time.[110, 127, 130] Where too few runs were seen, the median was not fixed and extended but instead continued with all data points in the chart contributing to this. The patterns of data points on the charts were visually inspected for signals compatible with accepted run-chart

rules which are probability-based, predefined data patterns with a probability of less than 5% occurring by chance.[110] The two run-chart rules used in this analysis are: 1) a shift, identified as a signal with six or more data points on one side of the median line and 2) too few runs, identified by counting the number of runs (groups of data points falling either above or below the median line), and then referring to the published guidance for the upper and lower limits.[110] The trend rule was not included due to evidence of lack of utility.[130] When a signal was identified in a care-process, a new median monthly delivery rate was calculated based upon the data contributing to the signal.

To address objective 1, the care-pathway was considered to be implemented if the 10 measured processes improved to the extent that all had a median monthly delivery rate of >80% following activation to the intervention (or a sustained median of <6 hours to surgery for process measure four, as above). Eighty percent was chosen as it is considered a minimum level of process reliability and is used by NELA to define an acceptable standard of care.[14, 131] Care processes already reliably delivered to more than 80% of patients were also included.

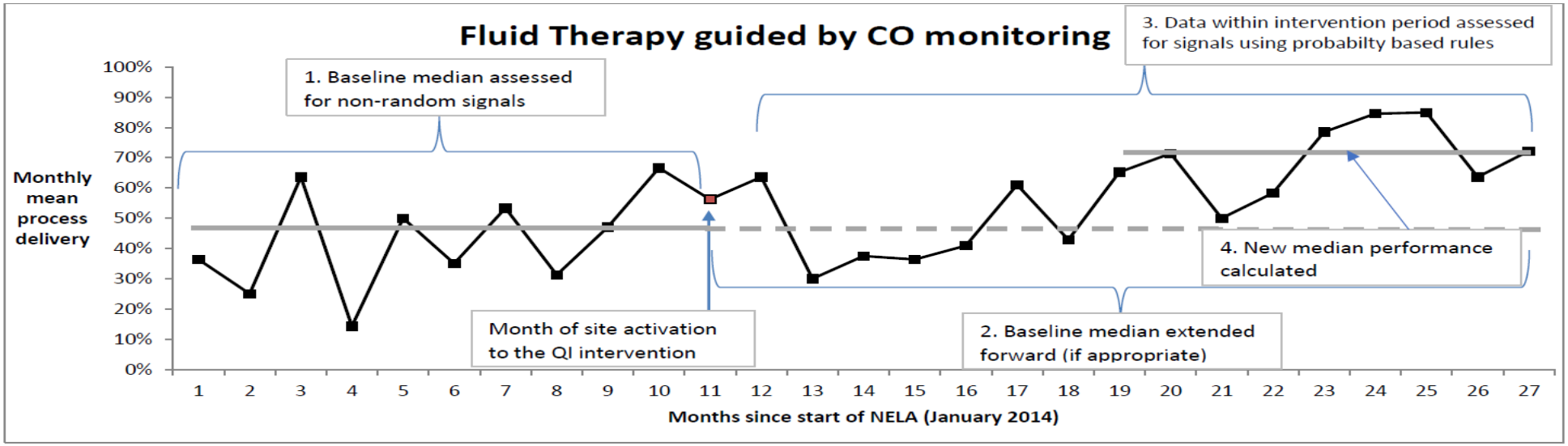
To address objective 2, care-process improvement was defined as any signal toward improvement identified on a hospital's run-chart, regardless of the magnitude of the improvement (unless followed by a subsequent signal toward process degradation). For each hospital we also calculated the proportion of patients before and after activation to the intervention who received each of the target care-processes and the median time from decision to operate to entry into the operating theatre (see care-process 4 above) pre and

post activation. These were then aggregated for all trial hospitals included in the run chart analysis to show the overall effect size of process changes.

The relationship between care-pathway implementation and implementation strategy fidelity is reported using descriptive statistics and the relationship analysed using a scatterplot and R2 calculation. The linear regression analysis was based upon the assumption that there would be a linear correlation between use of more implementation strategies and implementation of more (or indeed) all the measured care-pathway processes. Essentially this would either support or refute the central assumption of the EPOCH programme theory. However, the assumption of linear correlation in the context of a complex intervention being delivered in a complex system is potentially problematic. This is especially true given the heterogenous nature of the study cohort and the perspective taken in this paper of viewing the EPOCH trial as 93 individual QI projects. To explore this relationship further, there was a post-hoc analysis comparing the most and least improved trial sites and some key variables: fidelity to the six implementation strategies (fidelity defined as using five or six strategies vs. using fewer than five strategies), individual implementation strategy usage, NELA data collection method and care-process improvement between the least improved (fewer than two care processes improved; n=28) and the most improved (more than six care-processes improved n=14) hospitals. I used Fishers Exact Test, with 2x2 contingency tables to compare groupings and a one-sided p-value, with significance set at $p<0.05$.

We undertook a validation exercise, with an independent reviewer (RH) analysing a random selection of 200 of the total 800 charts (25%). The reviewer repeated the analysis of each chart. Results for the 200 charts were compared with the original analyses, any inconsistencies discussed and the final result agreed upon. An error rate of more than 5% was decided as the threshold for whether a further validation exercise would be necessary. We also undertook two post-hoc sensitivity analyses on the charts from the hospitals that improved more than half the process measures (14/80 hospitals improved more than six care processes) to test the different results obtained by using stricter analysis rules. These rules use thresholds for identifying signals (runs and shifts) based upon the total number of data points on the chart, rather than a fixed rule, which may provide more accurate findings.[130] In this group of most improved hospitals we also undertook analysis using a run chart centre line (median) based on all chart data, rather than the fix-and-extend method.

**Figure 4.1. Run chart analysis process and worked example**




**Explanation of steps in run-chart analysis: 1). Baseline median from month 1 - 11. There are 10 useful data points (excluding 1 point on the median) and 8 runs (groups of data points on either side of the median line). Referring to published guidance, this indicates the baseline performance is exhibiting normal (random) variation. 2). As the baseline period has no signals, baseline median is projected forward as the centre-line for the chart, against which to assess new data 3). Intervention period runs from month 12 - 27. Two signals of interest: i) number of runs, in this case 5, which when referring to published guidance would indicate a non-random signal and ii) shift, ≥6 data points on one side of the median line, which would also indicate a non-random signal. Therefore, a signal demonstrating an improved process is shown from month 19 onwards. 4) A new median performance is calculated from these data points (71% new median process delivery).**

**Figure 4.2. Inclusion of hospitals and patients in the run-chart analysis**

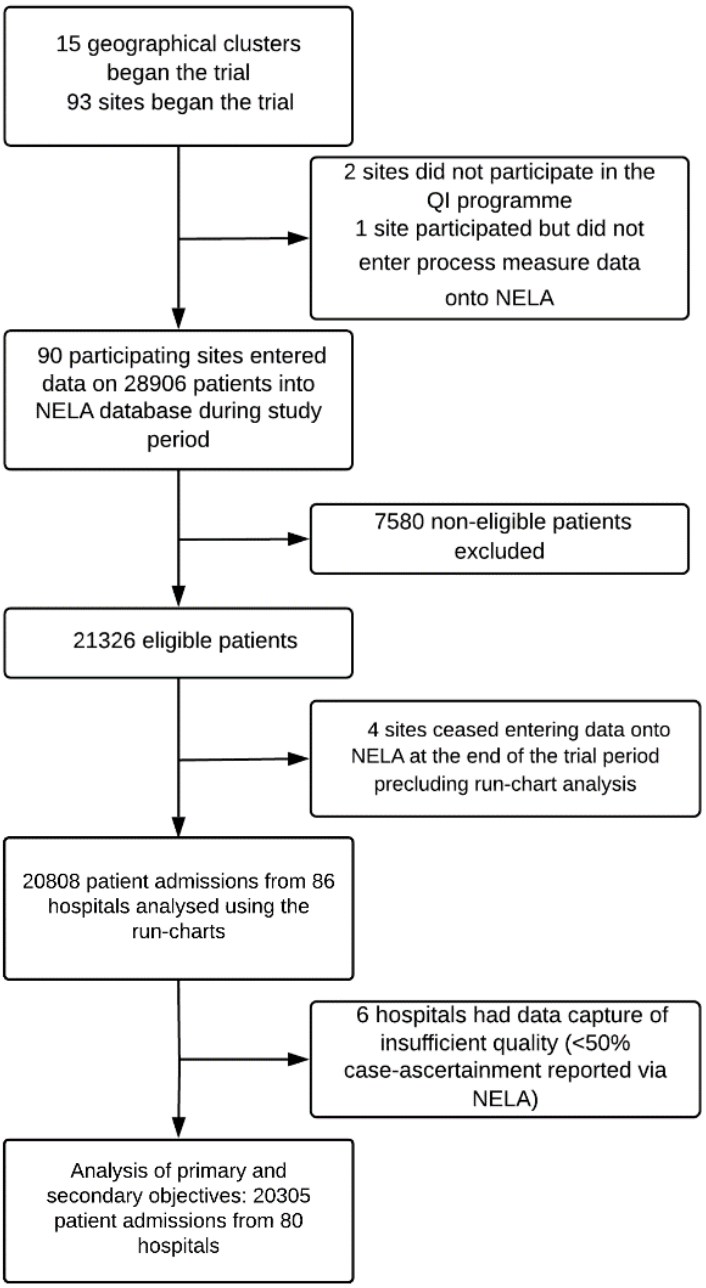


**Table 4.1: Key characteristics of data set.**

| **Key Hospital Characteristics**[132] | |
|---|---|
| **No. of operating theatres / 100 hospital beds (median (IQR))** | **2.5 (2.1 – 3.0)** |
| **No. of surgical critical care beds / 100 beds (median (IQR))** | **2.7 (2.1 – 3.5)** |
| **Emergency laparotomy volume (Years 2014/15) (median (IQR))** | **271 (204 – 371)** |
| **No. of secondary / tertiary referral hospitals** | **Secondary 58/80**<br>**Tertiary 22/80** |
| **Key Patient Characteristics** | |
| **Age (mean (SD))** | **68 (13)** |
| **Sex – Female** | **11101 (53%)** |
| **P-Possum Score (median (IQR))** | **7.6 (2.9 – 22.7)** |

### *4.3 Results*

Of the 93 hospitals enrolled in the EPOCH study, 86 hospitals had data available for analysis. However, six hospitals had data capture of insufficient quality (less than 50% case-ascertainment reported via NELA for either both Years 1 and 2 of the audit or for the year in which the hospital was activated to the QI intervention) to enable month-by-month analysis using run charts. Therefore, 80 hospitals were included in analyses resulting in the generation of 800 run charts for the 10 measures of interest, based upon analysis of data from 20,305 patient admissions (Figure 3). Table 4.1 displays key hospital characteristics of interest. In the validation exercise, six errors were identified, giving an inter-observer agreement of more than 95%. Of the six errors, three were errors where charts were marked as having signals toward improvement that were not there, and two were errors where signals toward improvement were missed. One was an error where a degraded care-process was missed. In all cases, signals were marginal and overall, these errors did not substantially change our main findings or conclusions.

**Table 4.2: Process measure improvement per hospital temporally associated with participation in the EPOCH trial QI programme**

| | Data signals (shifts and runs) identified on run chart analysis | | | |
|---|---|---|---|---|
| **Care Processes** | **Number (%) of hospitals with care process improvement observed n=80** | **Number (%) of hospitals with median baseline care process delivery ≥80% n=80** | **Number (%) of hospitals with degraded care-process after activation to EPOCH n=80** | **Difference post intervention vs. pre intervention (median, IQR, range)** |
| **1. Consultant led decision making** | **14 (17.5%)** | **71 (88.8%)** | **6 (7.5%)** | **0.44 (2.53 - 3.35, -16.19 - 20.33)** |
| **2. Consultant review of patient before surgery** | **17 (21.25%)** | **14 (17.5%)** | **4 (5%)** | **2.4109 (-3.67 - 6.37, -18.19 - 17.73)** |
| **3. Pre-op risk assessment documented** | **57 (71.25%)** | **6 (7.5%)** | **3 (3.8%)** | **13.66 (3.25 - 23.48, - 21.75 - 52.15 )** |
| **4. Time from decision to operate (DTO) to entrance to the operating theatre** | **14 (17.5%)** | **2 (2.5%)** | **5 (6.3%)** | ***Time in hrs*** **-0.500 (-1.30 - 0.37, -8.25 - 3.4083 )** |
| **5. Time to Theatre within NCEPOD timeframe** | **22 (27.5%)** | **32 (40%)** | **2 (2.5%)** | **8.391 (1.65 - 12.18, -7.81 - 25.65)** |
| **6. Consultant delivered surgery** | **24 (30.0%)** | **70 (87.5%)** | **2 (2.5%)** | **1.913 (-1.96 - 6.52, -13.86 - 18.66)** |
| **7. Consultant delivered anaesthesia** | **29 (36.25%)** | **57 (71.3%)** | **4 (5%)** | **3.8416 (-0.74 - 8.68, -22.948 - 30.30)** |
| **8. Cardiac output monitoring to guide fluid therapy** | **32 (40.0%)** | **3 (3.8%)** | **11 (13.8%)** | **4.766 (-1.10- 13.25, -29.21 - 50.72)** |
| **9. Measurement of serum lactate intra-operatively** | **42 (52.5%)** | **3 (3.8%)** | **3 (3.8%)** | **9.270 (2.14 - 17.52, -28.09 - 39.86)** |
| **10. Admission to critical care post-operatively** | **28 (35.0%)** | **15 (18.8%)** | **3 (3.8%)** | **2.222 (-3.62 - 7.33, -17.69 - 26.88)** |

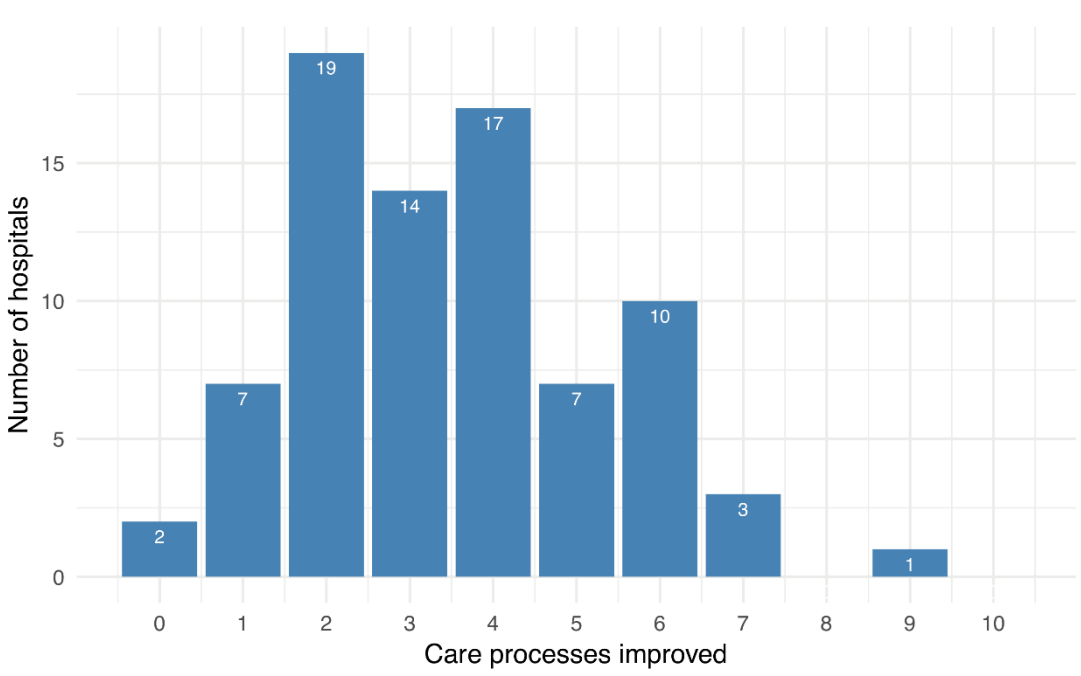
No hospital achieved implementation of the care-pathway according to our definition (all ten measured care processes improved to median delivery rate above 80%). Regarding objective 2 (describing all improvement, not just achievement of more than 80% reliability), 21/80 hospitals improved more than five of the ten measured processes and 14/80 improved more than six. Figure 4.3 displays the overall number of improved care processes per hospital. Pre-operative risk assessment (57/80 [71%]), intra-operative lactate measurement (42/80 [53%]) and cardiac-output guided fluid therapy (32/80 [40%]) were the most frequently improved care processes (Table 4.2). Consultant-led decision-making (14/80 [18%]), consultant review before surgery (17/80 [21%]) and time from decision to operate to surgery (14/80 [18%]) were the least likely care processes to improve (Table 4.2).

Questionnaire data describing implementation strategy use showed that 10/77 (13%) of QI leads responding said that all six strategies had been used, 23/77 (30%) indicated five had been used, 21/77 (27%) indicated four had been used, 8/77 (10%) used three strategies, 10/77 (13%) used two and 5/77 (6%) just one. No QI lead reported zero implementation strategy usage. Table 4.3 shows the reported usage of each QI strategy. As no hospital achieved care-pathway implementation, we undertook analysis of the relationship between implementation strategy usage and the number of care-processes improved. The cohort was divided into tertiles of implementation strategy usage (1-2 strategies, 3-4 strategies and 5-6 strategies) and we defined successful hospitals as those with six or more improved care processes (i.e. more than half of care processes improved). Of the hospitals that used one or two strategies, none improved six or more care processes, whilst among those that used three or four strategies, 4/25 (16%) hospitals improved six or more care processes, and in

those that used five or six strategies 9/30 (30%) of hospitals improved more than six care processes.

However, using a linear analysis model across the whole trial cohort, there was no correlation between implementation strategies used and the number of care-processes improved at individual hospitals (R2 = 0.084, Figure 4.4). Figure 4.5a-c presents the post-hoc analysis findings, comparing least and most improved hospitals by implementation strategy usage and NELA data collection method. Prospective NELA data collection, by all members of the care team (i.e. presenting a lower time-burden for QI leads), was positively associated with greater care-process improvement (p=0.039). Details of further evaluation of the relationship between care-process improvement and implementation strategy usage are reported in the post-hoc analysis section below.

**Figure 4.3. Number of care processes improved by each hospital during EPOCH trial (n= 80 hospitals)**



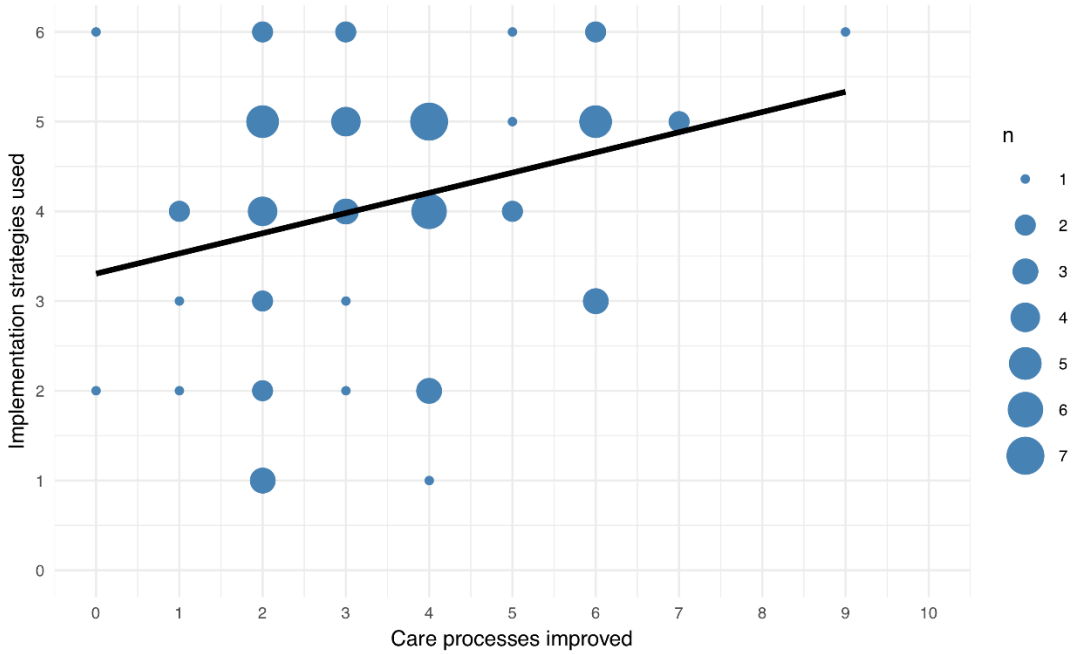



**Figure 4.4 Scatterplot of implementation strategy use and care-process improvement data**

During the analyses we identified the care processes in each hospital that were already reliably delivered, as defined by a baseline median of more than 80% delivery of a process measure. Consultant-led decision-making was the care process most reliably delivered before EPOCH, with 71/80 hospitals achieving a median of more than 80% for this measure. Of these hospitals 11 (15%) further improved upon this performance during the EPOCH intervention period. Consultant delivered surgery was often already reliably delivered, with

70/80 hospitals already achieving a median of more than 80% for this measure. Nevertheless, 19 of these hospitals (27%) managed to further improve this care process. Consultant delivered anaesthesia was the next most reliably delivered care process at baseline (57/80 hospitals) and 16 of these hospitals (28%) demonstrated further improvement in consultant delivered anaesthesia. Conversely, only 2/80 (2.5%) hospitals had a median time from decision to operate to surgery of less than six hours before the EPOCH trial started; this was also the most challenging care process to improve, although 17.5% (14/80) of hospitals did demonstrate an improvement on the run chart analysis.

Process degradation was also observed during run-chart analysis. Overall, 43/800 (5.4%) care processes across 28 hospitals were degraded after participation in the EPOCH Quality Improvement programme (i.e. a signal toward worse performance associated with activation to the EPOCH intervention). Despite being the third most frequently improved care process, use of cardiac output monitoring to guide fluid therapy was the most commonly degraded process (10/80 hospitals).

In the sensitivity analyses, using stricter run chart rules would have identified 78/140 care processes as improved in this group, rather than 90/140 using the standard rules, resulting in a group of ten hospitals, rather than 14, that improved more than six care processes. Regarding different approaches to the chart median, 6 / 140 (4%) of charts in the sensitivity analysis used a median based upon all data points in the original analysis (due to signals in the baseline period). Across the group of most improved hospitals, using a chart centre line based on all data points would have identified 57/140 (41%) care processes as improved,

resulting in a group of three hospitals, rather than 14, that improved more than six care processes (see Table 4.3).

**FIGURE 4.5: Panel 4a : Difference in implementation strategy use between least and most improved hospitals, Panel 4b: Fidelity to implementation intervention, comparing least and most improved hospitals by strategy usage, Panel 4c: Comparison of least and most improved hospitals by NELA data collection process**

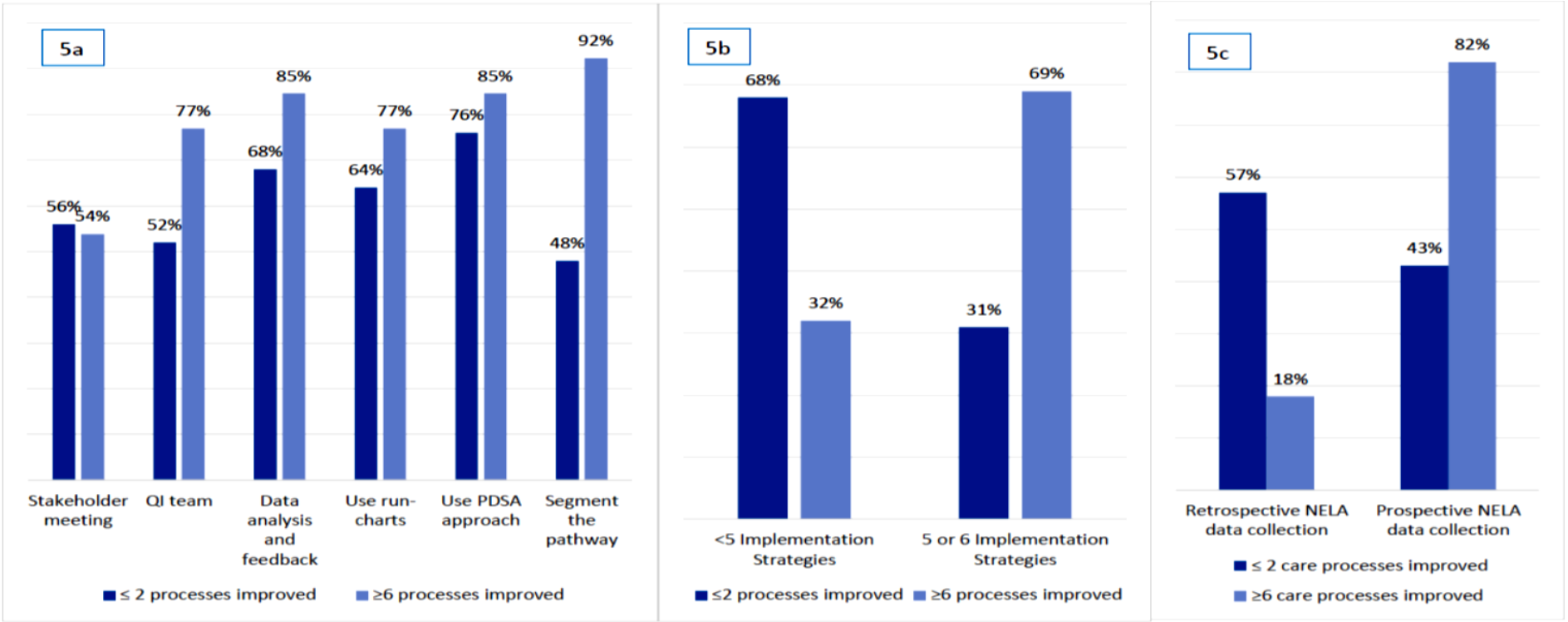


**Table 4.3. Results of sensitivity analysis, comparing 1) standard and strict run chart analysis rules and 2) a baseline median using the fixed and extend method (standard rules) and a median based on all chart data**

| Most improved hospital group | Number of processes improved (fixed baseline & standard rules) | Number of processes improved (strict rules) | Number of processes improved (median of all data points) |
|---|---|---|---|
| Hospital S.1 | 6 | 6 | 6 |
| Hospital S.2 | 6 | 6 | 1 |
| Hospital S.3 | 6 | 6 | 5 |
| Hospital S.4 | 6 | 6 | 6 |
| Hospital S.5 | 6 | 6 | 5 |
| Hospital S.6 | 6 | 5 | 3 |
| Hospital S.7 | 6 | 4 | 5 |
| Hospital S.8 | 6 | 6 | 3 |
| Hospital S.9 | 6 | 4 | 5 |
| Hospital S.10 | 6 | 5 | 2 |
| Hospital S.11 | 7 | 5 | 5 |
| Hospital S.12 | 7 | 6 | 6 |
| Hospital S.13 | 7 | 6 | 1 |
| Hospital S.14 | 9 | 7 | 4 |
| **Total** | **90/140** | **78/140** | **57/140** |

*Further post-hoc analysis of the relationship between care-process improvement and implementation strategy use.*

We used Fishers Exact Test, with 2x2 contingency tables to compare groupings and a one-sided p-value, with significance set at $p<0.05$. Only one of the six implementation strategies, segmentation of the care pathway, showed a significant difference between the two groups ($p=0.013$) (Figure 4.5a). Given the lack of hierarchy amongst five of the six QI strategies we

also compared the least and most improved hospitals by the number of QI strategies used (Figure 4.5b). When comparing the two groups and implementation strategy fidelity (i.e. using five or all six intervention strategies vs. not), there was a significant association between using more improvement strategies and more process improvement (p=0.032). Regarding other factors we used data from our process evaluation to compare the least and most improved hospitals by how NELA data was collected locally (Figure 4.5c), specifically if the data was collected prospectively, by all members of the care team, or retrospectively, by a small number of clinicians who were, in general, also likely to be the QI leads. Prospective data collection (i.e. presenting a lower time-burden for QI leads), was positively associated with greater care process improvement (p=0.039).

### *4.4 Discussion*

The main finding of this analysis was that no hospital in the EPOCH trial reliably implemented the care pathway within six months of the end of the intervention period. However, areas of improvement were identified. In total, 279 (of a possible 800) care processes were improved by hospitals through participation in the EPOCH trial and a small group of hospitals (17.5%, 14/80) were successful in improving more than six care processes. Effect sizes overall were marginal, but with substantial variance for each process across trial hospitals. Patient outcomes associated with the trial intervention were not the focus of this evaluation, but it seems logical that if only a small proportion used all the recommended implementation strategies and only a sub-set of these hospitals were able to improve more than half the target care-pathway processes, then the causal mechanism that it was hypothesised would lead to outcome improvement was largely absent in the EPOCH cohort. This confirms the findings of the main, patient level, trial analysis. This supports the

use of individual hospital level time-series analysis, both during a programme to monitor progress and support hospitals facing challenges, and as part of the evaluation strategy to provide granular understanding of cohort-level analyses. We used a prospectively defined run chart methodology, but in a sensitivity analysis our findings were found to be sensitive to the use of alternative methods of run chart construction.

This study contributes to the growing literature on methods to better understand improvement and implementation research results in the face of complexity.[60, 91, 129, 133] [24,28-30]. In particular, hospitals participating in multi-site cohorts may well achieve differing results; understanding this local level granularity enables a clearer understanding of what happened during a large-scale intervention and what led to, or hindered, overall success. In line with evidence that a multi-faceted approach to change is more effective,[134] the hospital teams in our study that achieved greater care-process improvement also reported using more of the implementation strategies recommended by the Quality Improvement programme. Whilst the relationship was absent in the linear model, this approach may be poorly suited to the complexity of this issue, especially across a large and heterogeneous cohort. However, it does identify a flaw in the assumptions within the EPOCH programme theory, which was essentially *Cetus Paribas* if you follow the implementation strategies you will be able to implement the care pathway'. But of course, the messy and complex world of the NHS means that 'all else being equal' rarely holds true and the process evaluation in Chapter 3 describes the variations in contextual barriers and enablers across the cohort in detail.

Analysis by groupings, and when comparing the least and most improved hospitals (Figure 4.5), suggests that greater improvement was possible (but not guaranteed) with use of more of the recommended implementation strategies. Although these analyses do pose a high risk of bias, and correlation is not causation, they do allow an exploration of the variation and heterogeneity of the cohort more effectively. Accepting all things are not equal, what can learn about potentially important differences between the most and least successful. These analyses support the hypothesis that the Quality Improvement intervention could be effective, but only if used in full and deployed within a supportive context i.e. they were only one of several important influences on effective improvement. Our process evaluation of the intervention describes in detail the contextual factors, both enablers and barriers, faced by hospitals as they attempted improvement (Chapter 3).[135] Major barriers included limited time and scarce resources to support clinicians leading improvement and, relatedly, an onerous burden of data collection which limited capacity to subsequently use these data for improvement. Related to this specific factor, our post-hoc analysis also indicated that in hospitals where systems to collect data prospectively existed, minimising the data burden on NELA and EPOCH QI leads, the number of improved care processes increased. Lack of interest amongst colleagues and seniors was also reported as a problem in many hospitals. Almost universally, contextual enablers were the opposite of these and future improvement programmes will need to fully address these factors to be successful, including allocating job-planned time for frontline improvement leaders and additional funding for support functions such as data collection and analysis.

There are other possible reasons why there was not greater care pathway implementation or care process change. First, our definition of reliable pathway implementation may have been too stringent. The standards set by NELA only require consultant presence in the operating room and admission to critical care for patients with a greater than 5% risk of mortality whereas the aspirational improvement goal of the EPOCH trial was for all patients requiring emergency abdominal surgery to be put onto the recommended care-pathway. Whilst 80% is an accepted threshold for defining minimum reliability,[131] it is possible that hospitals were guided by the more pragmatic standards as set by NELA, thus reducing the chances of pathway implementation as defined in this paper.

Second, three key care processes (consultant-led decision making, consultant-delivered surgery and consultant-delivered anaesthesia) were already being reliably delivered (to more than 80% of patients) in most hospitals at the start of the intervention period, which may have limited the head-room for further improvement of these particular care processes in some hospitals. Also, the value of one key process, cardiac output monitoring, was under debate in the UK during the time of the study,[136] and this may have meant some teams chose not to focus on it or, as our data shows, to move away from delivering this process completely.

Third, system level care processes, such as reducing the time to get patients into the operating room, were harder to improve than processes that individual clinicians were able to improve by themselves, such as assessing mortality risk. Nevertheless, nearly 30% of hospitals improved their performance on getting patients to the operating room in a time-

frame appropriate for their operative urgency. Considering the complexity of this system-level process, contingent on the actions of multiple stakeholders and on the other pressures faced by operating room suites in the UK, this feels like a substantial achievement. This mirrors findings from previous QI work regarding the degree of difficulty in attempting to improve systems-level processes compared to more discrete, individual professional or small-team delivered processes,[93] and supports the need to consider different, and potentially more intensive, strategies to improve system-level care processes. This may be of particular importance for this patient group given recent evidence demonstrating the positive impact upon mortality of system level changes such as single pathways of care in emergency general surgery and dedicated emergency surgery units.[67]

At the trial level, without these further analyses, the degree to which each hospital had implemented the care pathway as intended or improved would have remained unclear, as each hospital's signal was obscured within the results of a large and heterogeneous cohort. The use of run-charts to evaluate Quality Improvement programmes at scale remains rare, with some notable exceptions,[137] yet they are ideally suited to this level of granular data analysis. The main strength of this analysis is that we tested this approach experientially, alongside our main trial analyses, using the same data set, and found it was largely congruent with, but added value to, our previous understanding of what happened during the EPOCH trial. Human error, inherent in the visual inspection of run-charts (present even when using automated data analysis programmes), was mitigated by undertaking a validation exercise to provide assurance of reliability. Different approaches to constructing

and analysing run charts in sensitivity analyses were tested and found that the approach used may have a substantial impact on findings.

This analysis also had some limitations. First, performance of hospitals in six of the 15 trial clusters was analysed using run-charts that had a baseline median constructed of ten data points, which is the minimum acceptable number to use the probability-based run-chart rules. (This was due to the trial and the data collection process, via NELA, starting nearly contemporaneously, thus limiting baseline data in early clusters.) Second, analyses require decisions about the desired trade-off between sensitivity and the risk of false-positive signals being identified.[110, 126] Both of our sensitivity analyses, using stricter analysis rules and comparing different methods for creating the chart median, reduced the number of process improvements observed. In the latter analysis, this reduction was substantial. If stricter run chart rules, or a chart median based upon all data points, had been applied across all analyses, the level of care-process improvement identified would have been smaller than found than in our pre-planned analysis. Third, variations in the denominator for the monthly plotted percentages sometimes interfered with signals in the data (e.g. in a month with a small denominator, a few process failures may create a data point, breaking a signal that would otherwise indicate a move toward improvement). This, combined with the time-bound nature of the analyses, may have led to some real-world improvements not being identified using the run-charts. Our analysis may therefore ultimately have provided an overly conservative estimate of the volume of improvement associated with the EPOCH intervention.

This problem may have been mitigated by using both run charts and Statistical Process Control (SPC) charts in a head-to-head comparison. Although it would have produced further valuable reflections on various types of time-series chart for evaluation, it was beyond the scope of these pre-planned analyses to do this. Fourth, some of the analysis of the relationship between care process improvement and fidelity to implementation strategies was undertaken post-hoc, as the lack of care-pathway implementation meant we could not complete our pre-planned primary objective. Finally, whilst our process and ethnographic evaluation identified several potential enabling strategies and influences, we were not able to quantify these to explore their relationship directly with process improvement in these analyses. There may therefore be some important missing strategies that were not included in the original programme theory and were also not evaluated in this study.

### *4.6 Conclusion*

While EPOCH Quality Improvement intervention did not achieve reliable care-pathway implementation in any trial hospital, it was associated with individual improvement of care processes across the cohort and substantial improvement in a minority of hospitals. Individual hospital performance analysis using time series charts can help granular analysis of data from large, heterogeneous cohorts. This approach allowed us to fully understand changes in the delivery of patient care in response to the EPOCH trial intervention, but findings may be sensitive to the chosen run-chart design. In a longer-term project this understanding may have allowed us to adapt the intervention to be more successful.

# CHAPTER 5: Evaluation of the CholeQuIC Programme to reduce waiting time for patients requiring urgent cholecystectomy

## *5.1 Introduction*

Gallstone-related diseases account for approximately one-third of emergency general surgery admissions and referrals.[19] The majority of patients presenting to hospital with biliary pain will go on to have the removal of the gallbladder (cholecystectomy) as definitive treatment. Around a quarter of patients with acute cholecystitis or pancreatitis will re-present to hospital with gallstone-related symptoms at least once before they have a cholecystectomy.[20-23] Current national guidance from the UK National Institute for Health and Care Excellence (NICE) is for laparoscopic cholecystectomy within seven days of diagnosis of acute cholecystitis, and within index admission for pancreatitis,[24] and relevant professional associations and societies provide similar guidance on times to cholecystectomy.

Reducing the time to surgery for people who need a cholecystectomy reduces the number of times patients are readmitted with the same diagnosis and decreases their overall length of stay. Compared with delayed cholecystectomy, emergency procedures are associated with overall fewer workdays lost and higher patient satisfaction and quality of life.[20, 23] Furthermore, patients with acute pancreatitis run the risk of a fatal episode whilst awaiting cholecystectomy which is reduced by early removal of the gallbladder and stones.[25] A traditional concern that complications leading to conversion from laparoscopic to open surgery or risks of bile duct injury are higher with emergency surgery than delayed surgery

are not supported by current evidence.[21, 26] Overall, patients having early surgery (within the first week of presentation) do better on all indicators than those in delayed surgery groups.[27, 28] However, symptomatic gallstone patients in the UK wait longer and are more likely to be readmitted than in many other countries. Patients in France, USA and Sweden tend to undergo cholecystectomy on first admission with an average length of stay under 36 hours.[29-31] As with care delivery for emergency laparotomy, within the UK there is wide variation between NHS hospitals in the management of these patients, and wide variation in cholecystectomy rates despite the seven-day standard given in National Institute for Health and Care Excellence (NICE) guidance.[23, 24]

In an attempt to improve emergency cholecystectomy rates, the Royal College of Surgeons of England funded a 13-hospital improvement collaborative with an associated prospective, mixed-methods evaluation. The goal of the Cholecystectomy Quality Improvement Collaborative (CholeQuIC) was to improve emergency cholecystectomy rates. The evaluation was designed to identify whether improvements could be achieved by hospitals as a result of participation in CholeQuIC, what influenced change, and what lessons can be drawn for future improvement efforts. This chapter presents the quantitative outcome evaluation findings, using routinely collected administrative hospital data to answer the question: 'Did participation in a Quality Improvement collaborative (CholeQuIC) reduce time to surgery to within eight days from admission for patients requiring emergency cholecystectomy?' The findings from the concurrent process evaluation are presented in Chapter 7.

## *5.2 Methods*

### *5.2.1 Study setting*

CholeQuIC ran from October 2016 to January 2018 and adopted a Quality Improvement Collaborative approach,[94] incorporating evidence of what works in Quality Improvement, including specific learning from the EPOCH trial, including ensuring a very simple and streamlined data collection process, a simple and achievable QI goal addressing a recognised quality problem and adequate QI support for the scale of the quality problem being addressed (see Figure 5.1).[58, 60, 91, 103, 135] Recruitment to the collaborative was through open application, followed by a selection process; 13 of the 29 hospitals that applied were selected to participate in spring 2016. Criteria for selection were: the ability of staff at each hospital to commit sufficient surgical and support time to the programme, that the site had sufficient room for improvement (defined as sites that had emergency cholecystectomy rates that were around or below the national average) and did not have ongoing, related improvement projects. Further we sought to ensure that the selected cohort of hospitals was suitably illustrative of hospitals across England and Wales, ensuring a range of hospital size, surgical volume, and sites that provided specialist hepato-biliary services (3/13) The original Expression of Interest form, which includes the publicly available selection criteria, is in the Appendix.

**Figure 5.1: Developing an evidence-based QI Collaborative: The CholeQuIC Process**

*The right problem*

- Choosing a problem with common agreement that it needs fixing in this context (defined by stakeholders) and motivation from participants to solve
- Clearly defining and articulating the problem

*Measuring and Monitoring*

- Data collection, to understand the local demand, the size of the challenge and patient flow through their actual pathway
- Data analysis and feedback to monitor progress and motivate colleagues

*Support and collaboration*

- Sharing of ideas and outcomes with the collaborative; learning from other attempts and adapting local processes accordingly
- Expert clinical and QI support, training and coaching

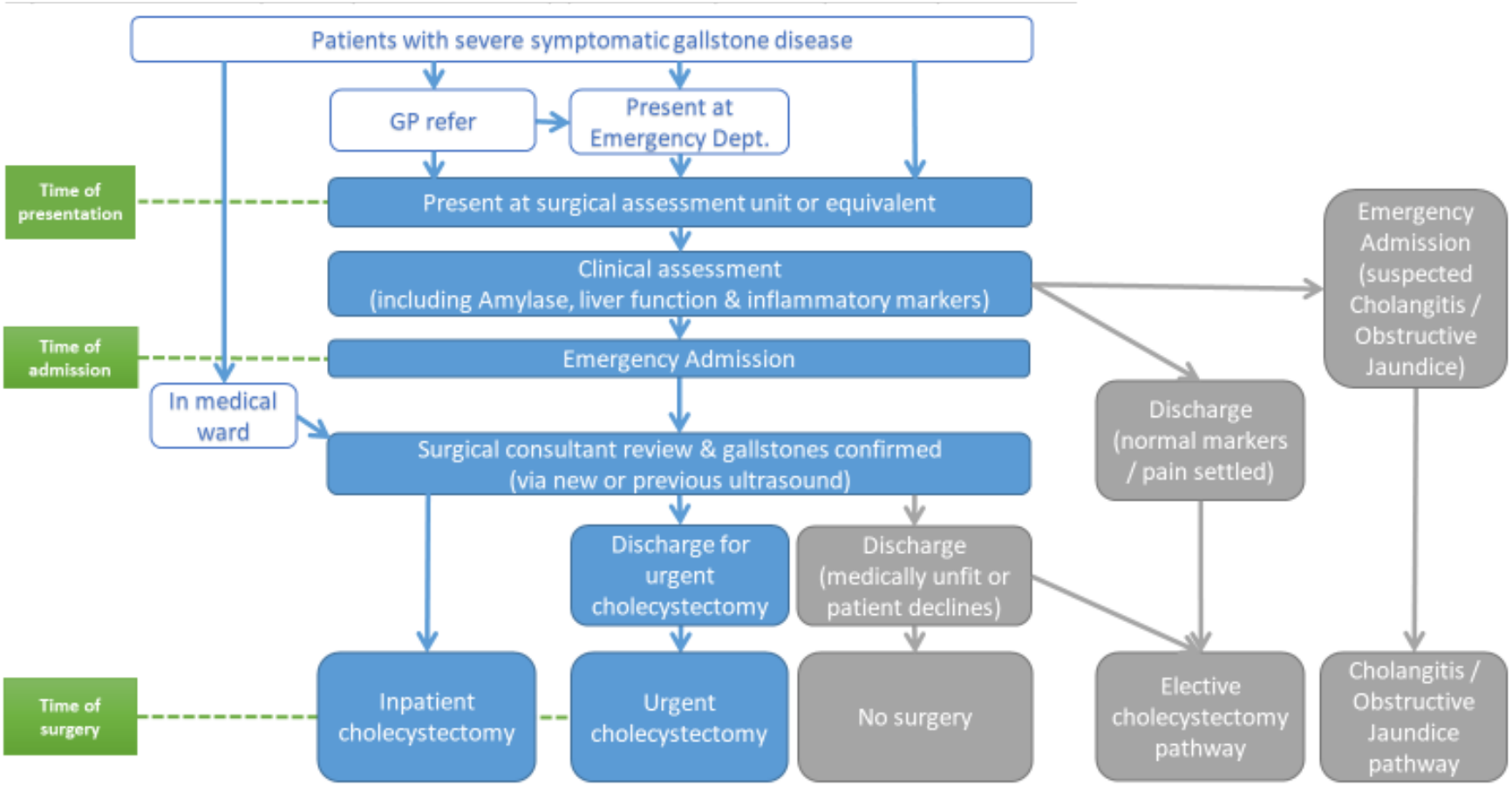
*The right solutions*

- Generating context specific solutions / new processes (supported by best evidence of any previous solutions)

Our aim was to demonstrate that time to emergency cholecystectomy can be reduced for eligible patients with acute biliary pain, cholecystitis or gallstone pancreatitis, by using QI methodologies to enable clinicians to drive change within their own hospital trusts. Surgery within eight days of presentation was chosen to match current National Institute of Clinical Excellence guidelines for acute cholecystitis (surgery within seven days of diagnosis), plus one additional day from presentation to allow time for diagnosis.[24] Eligible patients were those who agreed to have their operation (cholecystectomy) on an emergency basis and were deemed to be clinically fit for surgery as assessed by local clinical teams. This patient population was chosen as it covered the majority of symptomatic gallstone patients that follow a similar clinical pathway. Figure 5.2 shows the clinical pathway for patients requiring emergency cholecystectomy. There was no evidence to suggest that a single universally

applicable organisational approach would achieve this goal. Rather successful change would require concerted short-term resources to implement behavioural, process and system improvements.

**FIGURE 5.2: CholeQuIC pathway for acute biliary pain, cholecystitis and gallstone pancreatitis**

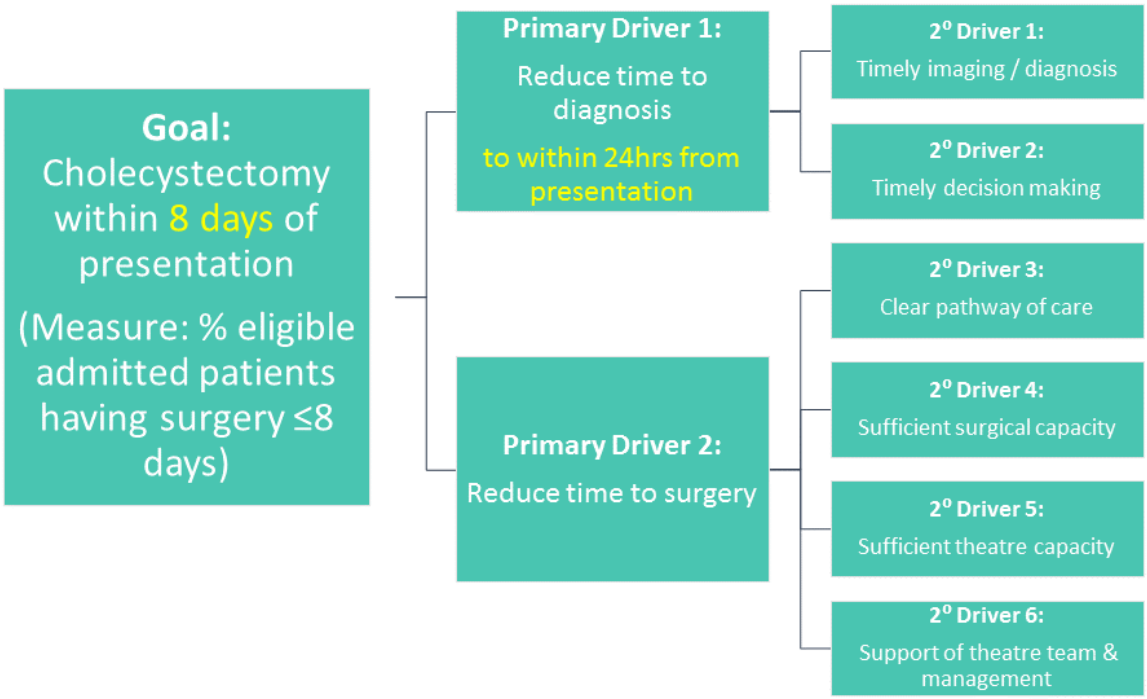


*5.2.2 The CholeQuIC intervention*

The CholeQuIC collaborative was a modified version of the Quality Improvement collaborative approach, incorporating evidence of what works for this type of QI project.[58, 59, 94] This approach was chosen because, although guidance on the optimal time to surgery for these patients exists, there was a lack of evidence on how to achieve this in practice. The goal for the collaborative was set in partnership with an expert reference group. After reviewing available guidance, Surgery within eight days of presentation was chosen to match current National Institute of Clinical Excellence guidelines for acute cholecystitis (surgery within seven days of diagnosis), incorporating an extra day from presentation for diagnosis to occur.[24]

Another researcher (JB, see Publications) and myself visited several hospitals that were known to be managing acute gallstone disease in line with guidance and used learning from these and from other relevant Quality Improvement work (e.g. [60, 103, 104, 135, 138]). In consultation with an expert in this field of surgery (IB), we developed a driver diagram (Figure 5.3) and a theory of change which explain the necessary conditions and the 'how' and 'why' of the intervention (Figure 5.4). The hard core of the CholeQuIC intervention was the focus on development, testing and ultimately implementation of context-specific solutions that would move their service toward achieving the project goal. The soft periphery of the intervention were components including local process measure collection and analysis, stakeholder engagement and learning from others within the collaborative. Teams were supported through this process with a range of activities, detailed in Figure 5.3.

**FIGURE 5.3 – CholeQuIC Driver Diagram**




Consequently, the collaborative was designed to support site teams to develop, test and ultimately implement context-specific solutions that would move their service toward achieving the project goal. Teams were supported throughout the collaborative process (Figures 5.4 & 5.5) in a number of ways: four collaborative meetings, designed to help teams learn about improvement methods and share progress and ideas with each other; webinars between meetings to help maintain momentum and to keep each other updated; and site visits from the CholeQuIC team, to provide in situ support, aid problem solving and facilitate engagement with colleagues. The CholeQuIC team also provided assistance with analysis of locally collected audit data, to help teams track their own progress.

**FIGURE 5.4 – CholeQuIC Theory of Change**

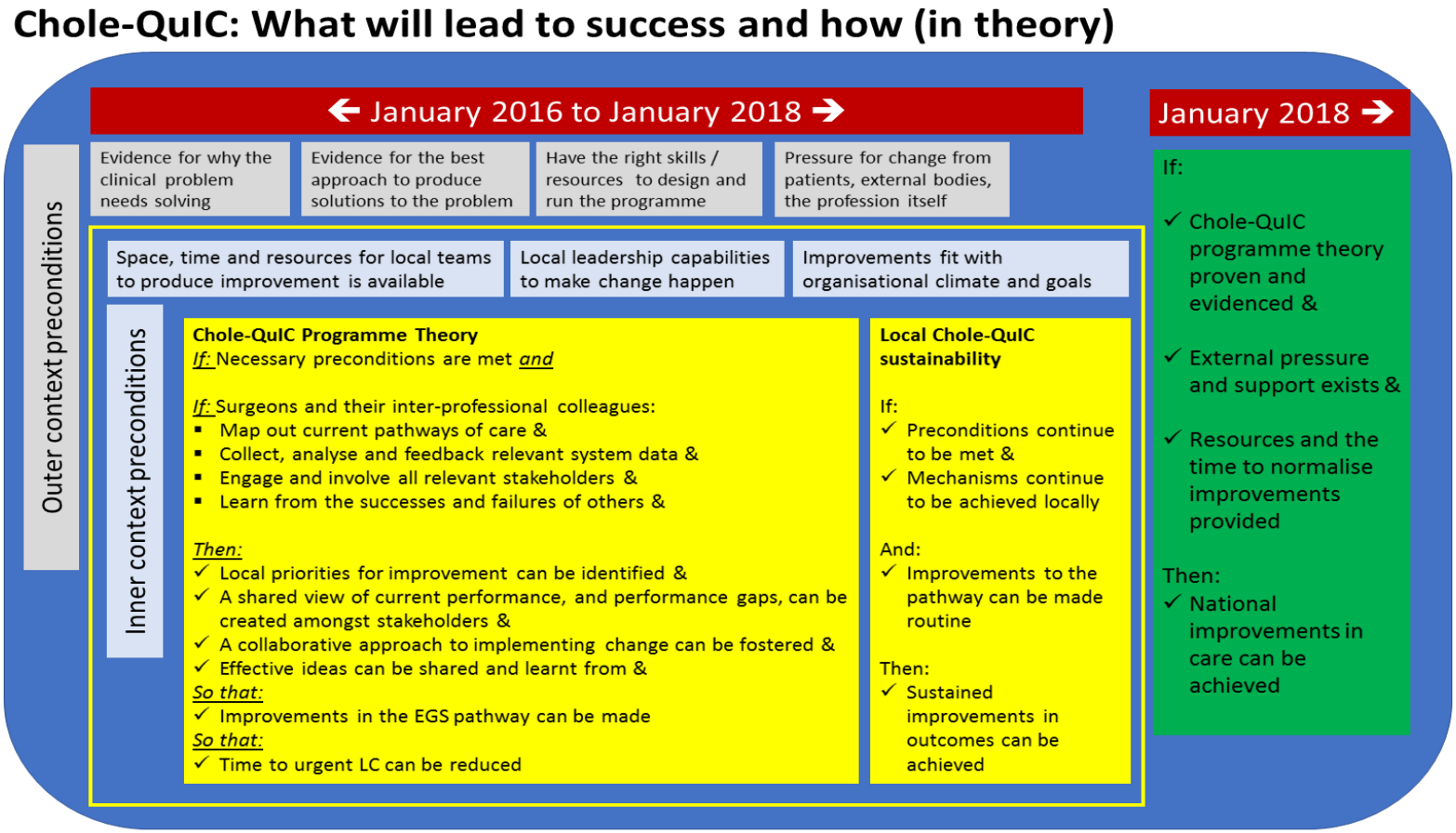
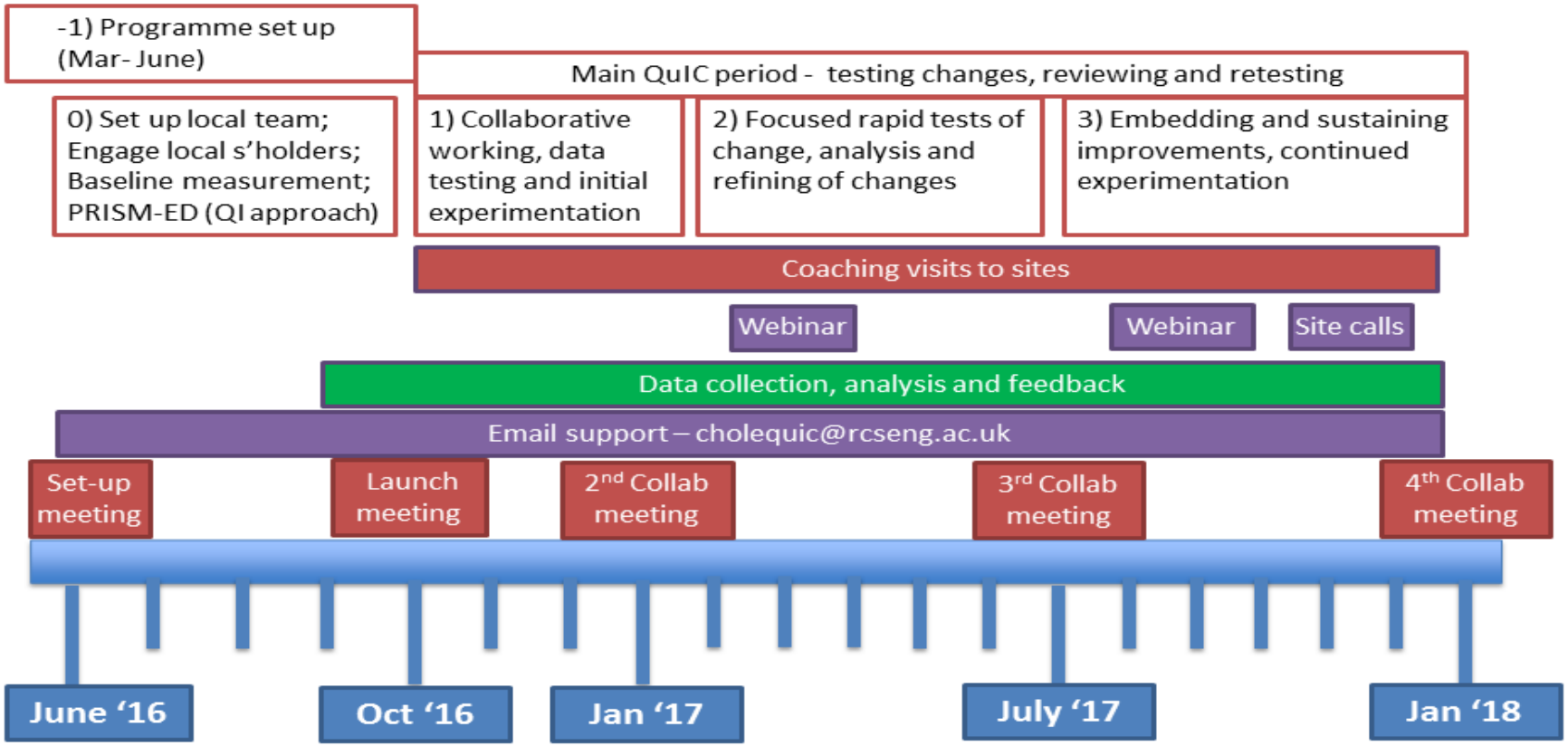


**FIGURE 5.5: CholeQuIC programme structure and key activities**



*5.2.3 Study design*

The mixed-methods evaluation was fully funded by the RCS and was approved by the ethics review board of Queen Mary University of London [QMREC1817a]. NHS Research Ethics Committee approval is not required for the analysis of anonymised routine data for service evaluation. Project findings are reported in accordance with SQUIRE guidelines for the publication of QI work.[66]

*5.2.4 Data collection*

Our evaluation used routine hospital data on all patients admitted as an emergency with acute biliary pain, cholecystitis or gallstone pancreatitis (ICD-10 codes K85.0,.1,.8,.9; ICD-10 codes K80.0,.1.,.2; K81.0,.1,.8,.9; K82.0-.4,.8,.9; ICD-10 code R10 who subsequently have cholecystectomy), from 1 April 2014 to 31 December 2017. For NHS trusts in England, aggregate quarterly figures were derived from the English Hospital Episode Statistics database. For Welsh Health Boards, data were obtained from the Patient Episode Database for Wales. Key variables were the number of patients admitted as an emergency with eligible conditions and number of patients who had emergency or elective surgery (Office of Population Censuses and Surveys code J18) within eight days of emergency admission. Some quarterly values were masked as less than five; for analysis, values between one and four for the masked values were imputed by multiplying the emergency admissions with the typical ratio for that NHS Trust (up to a maximum of four patients). Prospective quality improvement data were collected by hospital teams on all patients on the CholeQuIC clinical pathway (see Figure 5.2), including patient eligibility, time to surgery and whether they received an inpatient cholecystectomy, were discharged for an urgent cholecystectomy, or

were either temporarily or permanently unfit for surgery. Anonymised summary data were shared with the RCS core team, using an encrypted email service. Run charts and statistical process control charts were created from these summary data to assess local improvements, including changes to mean and median time to surgery, and three-day and 14-day surgery rates. These data were fed back to teams to support improvement and analysed as part of the evaluation.

*5.2.5 Data analysis*

We produced time series of quarterly activity for: each of the English and Welsh NHS organisations participating in CholeQuIC, for the English and Welsh site CholeQuIC cohort as a whole and a combined English and Welsh control group, consisting of 127 English Acute NHS trusts and 5 Welsh Health Boards. The time series was divided into three segments: quarters from April 2014 to June 2016 were pragmatically considered a sufficient duration to represent the baseline period (baseline) and identify and variations over time; one quarter from July to September 2016 was considered a transition or 'run-in' period, because we expected that while sites would be attempting changes at this point it would be too soon to identify improvements in emergency cholecystectomy rates; five quarters from Oct 2016 to Dec 2017 represented the intervention period (intervention), where we would expect to see improvement in our target measure. A negative binomial regression model was used to assess whether the proportion of patients having surgery within eight days had changed in the intervention period compared with the baseline period. The model only assessed whether there had been a change in the average level of the time-series because it was too short to test for other trends. A second model, with the relative difference adjusted for the

change observed in the control group, was also used. A negative binomial regression model was preferred to a Poisson model as it allowed for over-dispersion.

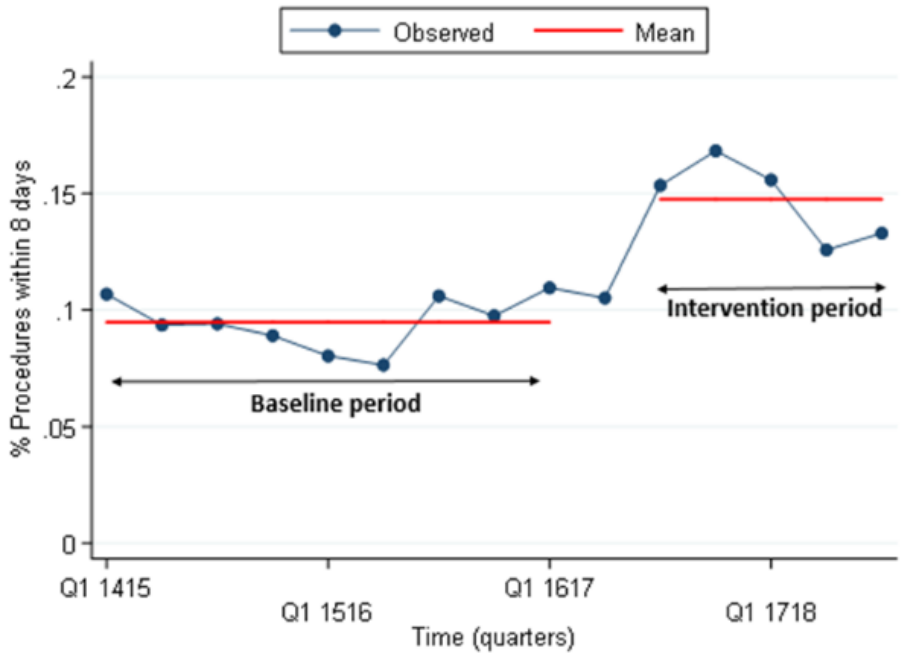
A statistical process control chart was created from the time to surgery for all eligible patient admissions who had a cholecystectomy at any of the 12 participating sites, using locally collected data. Time to surgery was plotted by date of presentation and upper and lower control limits were calculated at three standard deviations from mean time to surgery. Following standard practice for SPC chart interpretation mean and control limits were recalculated when a shift - nine or more successive data points above or below the mean – was identified. A shift is a data signal in an SPC chart that indicates a special cause variation, analogous to a significant, non-random, change in the data ($p = <0.05$).[127, 139, 140]

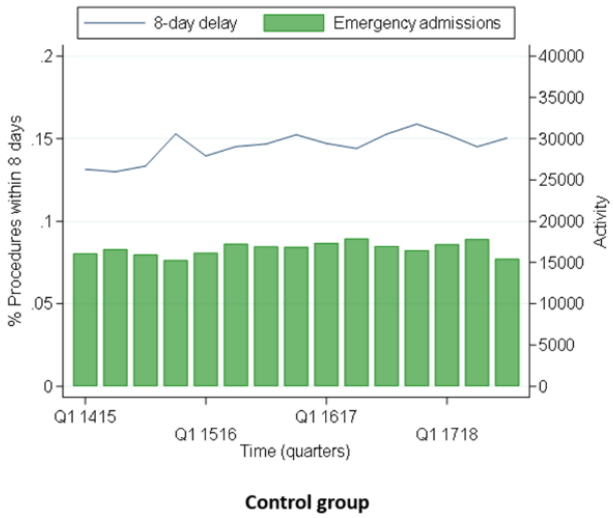
### *5.3 Results*

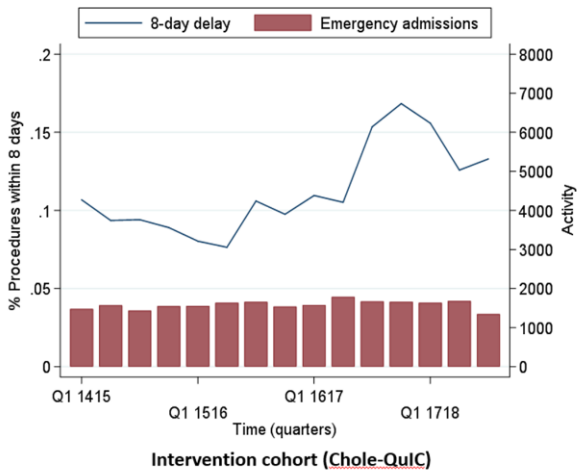
Of the 13 sites invited to join CholeQuIC, 12 fully participated throughout the programme, attending all four collaborative meetings, three webinars, participating in at least one site visit, collecting prospective data and testing improvement ideas. Site 13 withdrew voluntarily after nine months, having engaged to only a limited extent (no attempt at service changes, incomplete local data collection), and hence was included within the control group rather than the intervention group in the main analyses. In total there were just under 7944 acute biliary admissions across the CholeQuIC cohort of 12 during the 15-month intervention period: 5390 with acute biliary pain or cholecystitis, and 2554 with acute pancreatitis. 1160 patients had a cholecystectomy within eight days of their admission.

Our analysis indicates a significant increase in eight-day surgery rate, above any national trend toward improvement, for sites participating in CholeQuIC with an increase in the quarterly average of emergency cholecystectomies for all 12 sites from 145.5 patients in the baseline period to 232.0 patients in the intervention period. While the number of emergency admissions has stayed relatively stable for both the CholeQuIC group and the control group (the other 127 hospitals across England and 5 Health Boards Wales), the percentage of procedures within eight days has risen, with the control group eight-day rate increasing from 14.2 to 15.3% and the CholeQuIC group increasing from an average of 9.4% during the baseline period to 14.6% during the intervention period (Figure 5.6). This improvement represents a relative change of 1.56 more surgeries within eight days compared to 1.08 for the control group; accounting for the national trend toward improvement in the control group gives a relative change of 1.45 ([1.29-1.62], Table 5.1). A post hoc sensitivity analysis, using the same modelling but including site 13 in the intervention rather than the control group was also done. This is shown in Table 5.1; the sensitivity analysis did not affect the overall findings.

**Figure 5.6: Panel a, Time series of quarterly average 8-day rate in CholeQuIC cohort; Panel b: Control and CholeQuIC cohort group, admissions and 8-day rates for baseline and intervention periods**



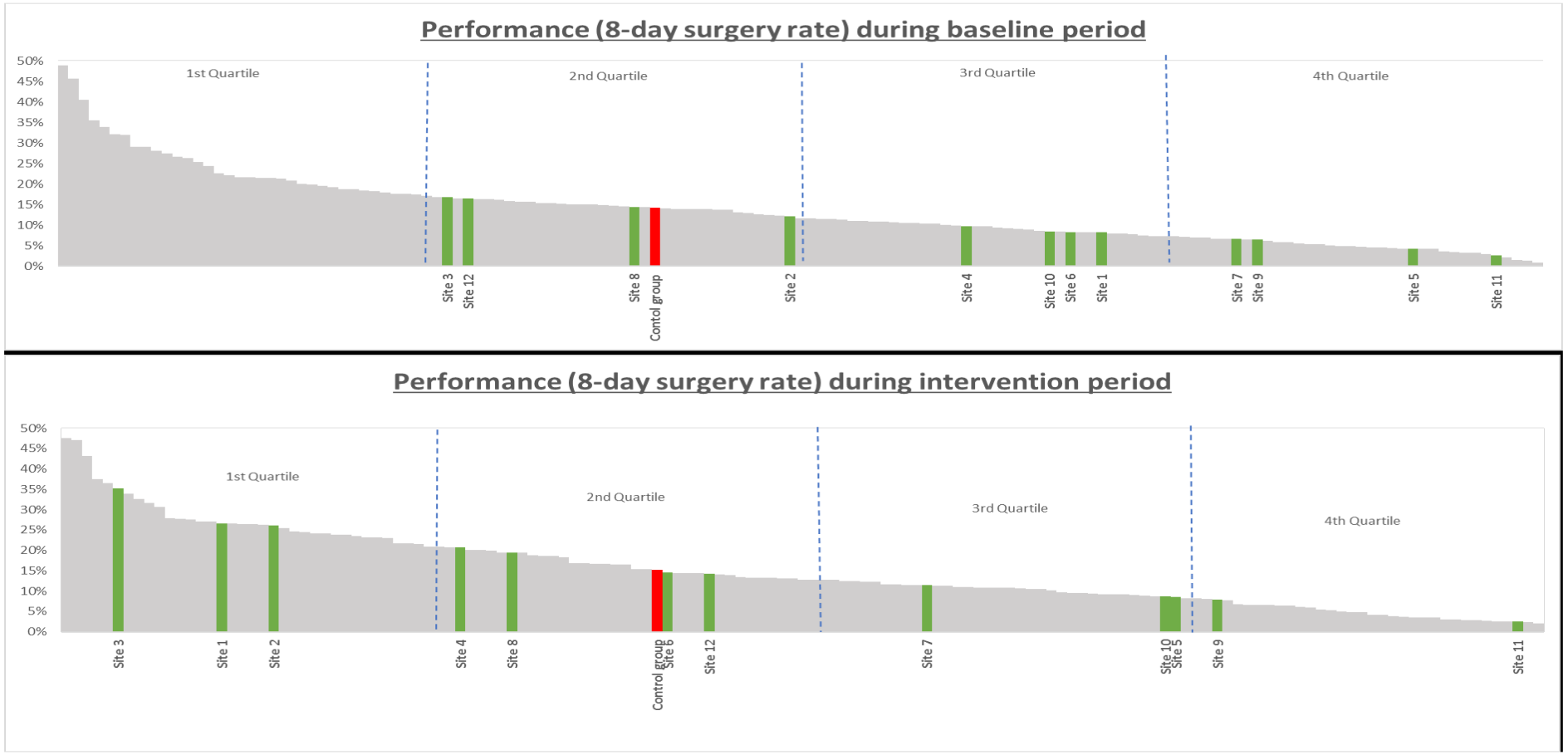



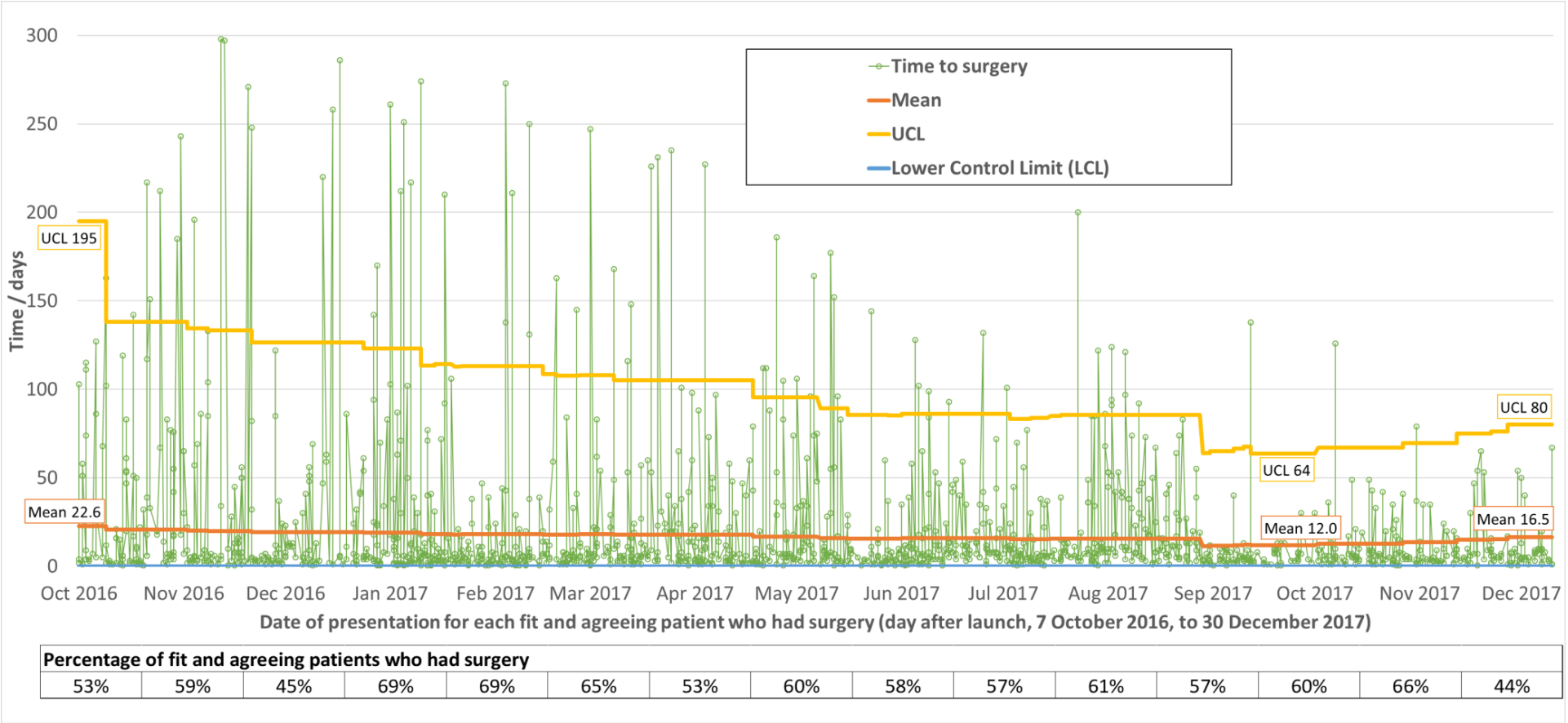
At the individual site level, eight of the 12 CholeQuIC sites had a significant improvement ($p<0.05$) in their eight-day surgery rate above the national trend toward improvement and in four sites the eight-day surgery rate increased to over 20% of all emergency admissions (see Table 5.1). As a comparator, between 2014 and 2018, eight-day surgery rates ranged from between 2-48% across England and Wales, with the top quartile across England and Wales achieving a median of 26% (range 21-48%). Figure 6.7 illustrates the ranking of CholeQuIC sites for the eight-day surgery rate compared to all English and Welsh Trusts over the baseline and intervention periods. All but two (Sites 11 & 12) of the CholeQuIC sites improved their ranking in emergency surgery rates among English and Welsh hospitals (a leftward shift in Figure 6.7), with three moving into the top quartile and two of three moving from the 4th to 3rd quartiles. Sites 11 and 12 saw concurrent reductions in performance.

**Table 5.1: CholeQuIC and control eight-day surgery rate and site by site data** DGH = District General Hospital, HPB = Hepatobiliary surgery service

| | Activity - all admissions for biliary disease | | % Procedures within 8 days (all admissions) | | Relative change from baseline | | Combined model (adjusted for Control Group) | |
|---|---|---|---|---|---|---|---|---|
| | Baseline | Intervention | Baseline | Intervention | Relative change | 95% Confidence Interval | Relative change | 95% Confidence Interval |
| **All CholeQuIC (12 sites)** | **13 929** | **7 944** | **9.4%** | **14.6%** | **1.56*** | **1.38 to 1.75** | **1.45*** | **1.29 to 1.62** |
| Control | 147 495 | 83 391 | 14.2% | 15.3% | 1.08* | 1.02 to 1.14 | | |
| | | | | | | | | |
| Site 1 – DGH | 521 | 301 | 8.8% | 25.9% | 2.94* | 2.02 to 4.27 | 2.73* | 1.88 to 3.96 |
| Site 2 – Tertiary (non-HPB) | 964 | 521 | 12.2% | 26.5% | 2.16* | 1.69 to 2.77 | 2.01* | 1.55 to 2.60 |
| Site 3 - DGH | 513 | 355 | 16.8% | 35.2% | 2.10* | 1.60 to 2.76 | 1.95* | 1.47 to 2.59 |
| Site 4 - DGH | 1 103 | 629 | 9.9% | 20.8% | 2.09* | 1.45 to 3.01 | 1.96* | 1.50 to 2.55 |
| Site 5 – Tertiary (HPB) | 1 333 | 770 | 4.6% | 8.6% | 1.88* | 1.27 to 2.77 | 1.74* | 1.22 to 2.49 |
| Site 6 - DGH | 1 114 | 619 | 8.5% | 14.7% | 1.72* | 1.06 to 2.79 | 1.60* | 1.19 to 2.16 |
| Site 7 - DGH | 1 189 | 627 | 6.7% | 11.2% | 1.68* | 1.06 to 2.65 | 1.54* | 1.11 to 2.15 |
| Site 8 – Tertiary (non-HPB) | 1 413 | 900 | 14.4% | 19.6% | 1.35* | 1.11 to 1.66 | 1.26* | 1.01 to 1.56 |
| Site 9 - DGH | 1 213 | 684 | 6.5% | 8.3% | 1.28 | 0.88 to 1.85 | 1.19 | 0.84 to 1.68 |
| Site 10 – Tertiary (HPB) | 1 476 | 760 | 8.4% | 8.8% | 1.03 | 0.64 to 1.66 | 0.97 | 0.72 to 1.33 |
| Site 11 – Tertiary (non-HPB) | 1 505 | 793 | 2.9% | 3.0% | 1.02 | 0.59 to 1.77 | 0.96 | 0.58 to 1.59 |
| Site 12 – Tertiary (HPB) | 1 585 | 985 | 16.5% | 14.2% | 0.86 | 0.69 to 1.09 | 0.8 | 0.64 to 100 |
| | | | | | | | | |
| **Sensitivity analysis** | | | | | | | | |
| All CholeQuIC (13 sites) | 15 973 | 9 108 | 9.0% | 13.3% | 1.47* | 1.30 to 1.67 | 1.37* | 1.22 to 1.55 |
| Control | 147 495 | 83 391 | 14.2% | 15.3% | 1.08* | 1.01 to 1.14 | | |
| * = significant improvement ($p < 0.05$) | | | | | | | | |

**Figure 5.7: % of procedures within 8 days of admission in baseline and intervention periods. Data representing CholeQuIC sites in the national context (Control group mean performance shown as red bar)**






| Percentage of fit and agreeing patients who had surgery | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 53% | 59% | 45% | 69% | 69% | 65% | 53% | 60% | 58% | 57% | 61% | 57% | 60% | 66% | 44% |

**Figure 5.8: Statistical Process Control chart for all patients who have surgery from the 12 participating sites**

We created a Statistical Process Control chart (Figure 5.8) that presents the locally collected data on time to surgery for the 1 580 emergency admissions who had a cholecystectomy, from a total of 3 001, during the improvement period. Although the percentage of eligible patients receiving a cholecystectomy remained consistent throughout the project (ranging from 44-69%), the average time to surgery for these patients improved over time, and variation reduced. Mean time to surgery reduces from 22.6 days at the project start, to a low of 12.0 days (Sept-Oct 2017, finishing at 16.5 days in December 2017). Variation in time to surgery also reduced substantially over time: from September 2017 breaches to the Upper Control Limit become infrequent, despite a concurrent tightening of the control limits. This indicates that not only was the time to surgery reducing overall across the cohort but that the process was also becoming more reliable.

### *5.4 Discussion*

The main finding of this evaluation was that 8/12 hospitals participating in a Quality Improvement Collaborative were able to significantly increase early cholecystectomy rates for patients with gallstone disease requiring hospital admission, in line with national guidance. As this was a controlled cohort evaluation, we were able to account for any trend toward improvement in the remaining 127 hospitals in England and Wales. Results remained significant when the small national improvement trend was accounted for. When plotted on Statical Process Control charts, locally collected data showed a clear reduction in variability and increase in reliability in providing timely laparoscopic cholecystectomy to this patient group across the cohort.

Looking at eight-day surgery rate, the cohort average of 14.6% remains below the national control group average of 15.3%. This is partly explained by the large range of outcomes between the 12 sites, 3.0-35.2% (reasons for variation are discussed below), and partly explained by the selection criteria for inclusion in the programme: sites were chosen with room for improvement and who were not currently involved in improvement activities in this area. Finally, it is important to remember that this was an improvement programme where changes were introduced throughout the 15 months, not introduced in totality at the start. When interpreting these rates, it is important to recognise that only five of the 127 English Trusts and no Welsh hospitals have achieved more than 35% eight-day surgery rate between 2014 and 2018. These data may appear low as the denominator includes all patients recorded with relevant ICD 10 codes. Clinical assessments by the participating site teams suggest that a large proportion of this denominator are not eligible for surgery, but this is the only viable data set to compare intervention data with control data. For patients who are suitable for early surgery, our study shows that it is possible for hospitals to significantly improve their service to ensure these patients receive the timely surgery that national guidance recommends.

The variation in outcomes between the 12 sites suggests that there are particular influences that aid or hinder successful improvement. Several important themes emerged from the concurrent mixed-methods process evaluation which looked at the characteristics of the four most successful sites in comparison to the four least successful sites. These are described in Chapter 6.

To our knowledge this is the first Quality Improvement intervention that has focused on this patient group. National gallstone registries in both Sweden and Denmark have been credited with facilitating overall improvements in quality of care through benchmarking, [141, 142] but more active efforts to use Quality Improvement methods to improve care across multiple institutions have not been undertaken. Indeed, surgeon-led improvement collaboratives appear rare, with a recent systematic review of published improvement collaborative evaluations only finding a small number focused on improving care for surgical patients and only one located within the speciality of General Surgery.[59] With regards to the success of this collaborative, the robustness of the evaluation design and the results achieved compare favourably with the outcomes of previous Quality Improvement Collaboratives in other areas of health care. This is likely to be due to an effective, evidence-based intervention design combined with the motivation and efforts of the site teams. This reinforces the argument that Quality Improvement Collaboratives should focus on an issue or problem for which is a strong consensus that change is required, especially when both problem and solution identification are supported by a respected professional body (in this case the Royal College of Surgeons of England).

This evaluation has several strengths, including the use of registry data that facilitated both a substantial baseline data period and comparison with a control group to observe both in-cohort and national (secular) trends toward improvements in care for this patient group. We chose to compare our cohort with all other NHS Trusts rather than matched comparisons because a) it would have been difficult to determine which potential matching criteria (e.g. overall bed numbers, surgical bed numbers, number of surgeons who could perform the

procedure, number of operating theatres) would be appropriate and would not introduce inadvertent confounding, b) we wished to evaluate the effect of participation in CholeQuic against an overall national trend and c) we were careful to select a range of hospital sizes and types and geographical locations to be broadly representation of NHS Trusts in England and Wales. The evaluation also has limitations which should be considered. First, sites volunteered to participate and thus a commitment to provide surgical and support time for the duration of the project was present in the sites that may not be present in all hospitals. Second, to demonstrate that improvements could be achieved and sustained in a range of contexts, the design of the programme was to select sites with headroom for improvement; correspondingly most sites had baseline performance with regards to the eight-day rate below the national average. It may be that sites with better baseline performance would be less motivated or able to improve further, even if their current practice was not in line with national guidance. Third, as five of the eight members of the evaluation team were directly involved in delivering this programme, there is a risk of bias in the analysis, which was mitigated by ensuring an independent researcher (DC) carried out the quantitative analysis and that an independent evaluation expert (GM) regularly reviewed the evaluation processes to provide both internal and face validity.

Finally, at this stage some potentially relevant outcome data was not available for analysis from routine national data, including readmission rates, 14-day rates or median and mean time to surgery. Balancing measures such as positive or negative impact on waiting times for biliary patients on elective lists or other patient groups were not looked at but would have been useful to assess for any unintended consequences. However, the logic of CholeQuIC

was that the current status quo was not acceptable and that the move should be toward offering acute biliary patients a better service from the inception of CholeQuIC. By offering patients definitive treatment in a timely fashion, they would receive higher quality care and these patients would also not be added to the pool of patients with biliary disease in the community awaiting surgery. Thus small, short term impacts on elective waiting times were anticipated but were traded off against the immediate and longer-term benefits.

### *5.5 Conclusion*

A surgeon-led Quality Improvement collaborative approach can be effective at improving care for patients requiring emergency cholecystectomy. The learning from this collaborative should be useful for others wishing to improve care for patients with acute gallstone-related disease and potentially other surgical patient groups where current care is below the standards set by national guidance.

# CHAPTER 6: Understanding the influences on successful quality improvement in emergency general surgery: Learning from a process evaluation of the CholeQuIC project

## *6.1 Introduction*

In this chapter, I present the qualitative aspects of the prospective, mixed-methods process and outcome evaluation of the Cholecystectomy Quality Improvement Collaborative (CholeQuIC) project, combined with knowledge of the extent of sites' improvement from our quantitative evaluation, as described in Chapter 6. Noting both the overall positive impact of CholeQuIC and its variability across participating sites, this mixed-methods process evaluation was designed to answer the following questions: 1) How was the collaborative delivered by the faculty and received, understood and enacted by the participants locally? and 2) What influenced teams' ability to improve care for patients requiring emergency cholecystectomy?

## *6.2 Methods*

### *6.2.1 Summary of the CholeQuIC project*

The RCS commissioned the CholeQuIC project in spring 2016. Recruitment to the collaborative was through a competitive application process; 13 of the 29 hospitals that applied were selected. Criteria for selection are described in Chapter 6 and were designed such that the cohort would represent the spectrum of hospital characteristics across the UK National Health Service (NHS).

### *6.2.3 Study design*

Prospectively designed, mixed-method process evaluation, utilising the knowledge and experience of CholeQuIC faculty (TS, JB, NQ, IB) combined with oversight from a senior, external researcher, with expertise in the field of mixed-methods and qualitative evaluation (GM). The evaluation was approved by the ethics review committee of Queen Mary University of London [QMREC1817a]. Project findings are reported in accordance with SQUIRE guidelines.[66]

*6.2.4 Data collection*

We collected a range of data at both the collaborative-wide level (all sites) and in a purposively sampled sub-set of sites. At the collaborative level we collated: 1) field notes, compiled by another researcher (JB) and myself immediately following coaching site visits; 2) ethnographic observations, involving non-participant observation of each of the main collaborative meetings, undertaken by external researchers; 3) notes from webinars and site calls; and 4) project documentation, including slides prepared by teams and summative site reports written by the faculty for each site.

We purposively sampled a sub-set of sites to take part in focus groups with a maximum variation sample (in terms of surgical volume, teaching / specialist status and performance during the project). Focus groups were recruited at two stages (five months into the project and at the end of the project). Focus groups ranged from four to 12 participants, and included lead surgeons (consultant grade), surgical trainees, other junior doctors involved in the project, nursing staff, anaesthetic staff, booking co-coordinators and service managers. Ethnographic observations and focus group recordings were professionally transcribed. All sites and individuals are pseudo-anonymised.

*6.2.5 Data analysis*

To answer question 1 (How was the collaborative delivered by the faculty and received, understood and enacted by the participants locally?) we primarily used data from project documentation and ethnographic observations of the collaborative meetings to understand participants' response to the meetings and programme overall, using a deductive framework approach driven by our question: a) how the programme was delivered by the faculty and b) how it was received, understood and enacted by the participants. To answer question 2, on what influenced improvement in care for patients requiring emergency cholecystectomy, we adopted a comparative case study approach. We identified two sets of cases: 'highly successful hospitals' and 'challenged hospitals', each containing four hospitals, using the main outcome measure for the collaborative (increase in the proportion of patients who had their surgery within eight days of presentation).

Data analysis for question 2 was based on a modified form of the Framework Analysis approach.[143, 144] We generated emergent themes that seemed important to improvement success from the data for all hospitals in the cohort, sensitised by constructs from Normalisation Process Theory,[113, 145] and the CholeQuIC Theory of Change (Figure 5.2**).** Normalisation Process Theory maps out the improvement process as the product of four social mechanisms (see Table 6.1): coherence (what individuals and teams do to make sense of a new practice); cognitive participation (what individuals and teams do to engage with a new practice); collective action (what individuals and teams do to enact a new practice); and reflexive monitoring (what individuals and teams do to appraise the effects of a new

practice). We chose Normalisation Process Theory as a structure, primarily to support our understanding of the social and contextual influences at play within the CholeQuIC sites. Through doing this, we adapted and added to the Normative Process Theory constructs to reflect the themes emerging from our inductive data analysis.

After identifying the sub-sets of 'highly successful' and 'challenged' hospitals as above, we undertook a more structured deductive approach where data were further analysed using this emergent set of themes and influences. Throughout the process, manual coding of data was undertaken separately by three individuals and then discussed with other team members, with codes and aggregate themes agreed during regular meetings. Analysis of the case study data involved developing within-case themes, then identifying cross-case themes and patterns, looking in particular for data that might provide understanding of what enabled or hindered successful improvement, and identifying potential rival explanations. Having familiarised ourselves with the data and emergent themes we developed a visual framework of within- and across-case patterns (Figure 6.1).

**Table 6.1: Description of key influences on success and related Normalisation Process Theory constructs**

| Description of key influences | Overall area of work | Related NPT construct |
|---|---|---|
| Cognitive, relational and behavioural work | | |
| 1. Achieving clarity of purpose amongst site leads and all key stakeholders | Sense-making | Coherence |
| 2. Capacity (time and resources) to lead and effective team working / project support | Relational | Cognitive participation |
| 3. Turing ideas into action | Making change happen | Collective action |
| 4. Learning from own and others' experience | Learning from change | Reflexive monitoring |
| Clinical process | | |
| 5. Creating additional capacity to do | Surgical / theatre capacity | N/A |

| | | |
|---|---|---|
| emergency cholecystectomies | | |
| 6. Co-ordinating / managing the patient pathway | Patient pathway / flow | N/A |

## *6.3 Results*

Of the 13 sites invited to join CholeQuIC, 12 fully participated throughout the programme, attending all four collaborative meetings and three webinars, participating in at least one site visit, collecting prospective data throughout and testing improvement ideas. Site 13 withdrew voluntarily after nine months, having engaged to only a limited extent (no attempt at service changes, incomplete data submission). In total we collected evaluation data comprising six focus group transcripts (out of the eight planned) from five hospitals, field notes from 17 site visits, four transcripts from ethnographic observations of collaborative meetings, and 12 summative site reports. We were unable to convene focus groups with two selected sites due to logistical issues, but had detailed field notes for both from site visits towards the end of the project.

### *6.3.1 Outcome of the CholeQuIC project*

CholeQuIC achieved its aim of demonstrating that gallstone care can be successfully improved in English and Welsh hospitals, although the extent of this improvement varied between the participating sites. Two-thirds (8/12) of participating hospitals improved care significantly from the baseline period (April 2014 to June 2016) to the intervention period (Oct 2016 to Dec 2017) in the main outcome measure for the collaborative, proportion of patents having surgery within eight days of presentation, even after accounting for a secular trend towards improvement among control hospitals (Table 5.1). Note that although the intervention period results still appear low for many sites, our data from the national control group indicate the top performing quartile of hospitals across England and Wales achieve a median of 26% (range 21-48%) of patients having their urgent laparoscopic cholecystectomy within eight days. Four highly successful hospitals achieved a significant

improvement that was well ahead of any secular trend across the rest of England and Wales, at least doubling rates of surgery compared to the baseline period. Moreover, this improvement was sustained for nine months or more during the project. Conversely, four challenged hospitals did not achieve a significant improvement against the eight-day goal in our analyses and their performance on the main outcome measure remained below the national average for the duration of the project.

*6.3.2 Delivery of the collaborative activities*

The collaborative programme was delivered largely as planned, with the exception of one additional set of site calls, added to maintain teams' momentum during Autumn 2017. Site participation at each meeting was complete with no site teams missing a main meeting (and site 13 attended all meetings prior to withdrawing from the collaborative). The median size of each site's team attending was three members of staff (a range between one and six). Regarding calls and webinars, 11 of 13 teams participated in the first webinar and nine of 12 teams participated in the second webinar the site calls. Fewer site visits to coach teams were delivered than we planned. We planned 26 site visits over the course of the programme (two per site) but by the end of the programme had made only 17. All sites had at least one visit and were offered a second. Reasons for declining a second visit included lack of time to host the visit, failure to find a mutually convenient data and time (between site and CholeQuIC teams), or site leads taking the view that a second visit was not necessary.

*6.3.3 How was the collaborative received, understood and enacted by participants?*

Overall, the CholeQuIC project was well received by participants, in terms of both the clinical problem to be fixed and the approach taken to improve outcomes.

"*The consultant said immediately that 'this project was absolutely the right thing to do', he said he knew it was 'the way forward' … [This] seemed to be the consensus of opinion of many of the other consultants attending.*" [Ethnographic notes, January 2018 meeting]

"*As the room fills up, a sense of energy gradually builds. There does seem to be a buzz in the room, a sense of anticipation. I sense that people want to be here*." [Ethnographic notes, July 2017 meeting]

Several site leads reported having wanted, or having tried unsuccessfully, to improve care for this patient group for a long time, and felt that this project, particularly with the associated support from a professional body like the Royal College of Surgeons, was what they needed to drive improvement forward.

"*He said that he had been trying for many years to get urgent cholecystectomies undertaken in his hospital but with no success…He made the comment that he felt that at their hospital they had got all the necessary ingredients to make this work but that unfortunately they 'haven't got the oven'*." [Ethnographic notes, October 2016 meeting]

Attendees at collaborative meetings often reported feeling motivated or re-energised by attending and some found they gained ideas from participants who were doing well, and insights into how to overcome challenges.

*"[Participants told me] it had been very useful listening to others' ideas. They also mentioned that they do not have a pathway as such, except possibly one that is "in their head". So they found the session about pathways particularly useful*." [Ethnographic notes, Jan 17 meeting]

This feeling was not universal, however. Some others site leads enjoyed the social aspects of the meetings but stated that they did not derive benefit from the collaborative approach.

"*…and I like the meetings, it's nice, because I know some of the other guys from the hospitals so it's quite a sociable thing and it's good to go and speak to people […] but I haven't found, I haven't had to collaborate [with other CholeQuIC sites]; I haven't found the need to speak to other units independently from what we're doing"* [Focus group, Highly successful group]

Overall, our data suggest that the most important aspects of the collaborative, from sites leads' point of view, appeared to be: 1) meeting up with like-minded colleagues, 2) the external drive or focus that the collaborative afforded and 3) the legitimacy conferred by CholeQuIC's status as a Royal College of Surgeons initiative. Whilst there was some clear evidence of cross-pollination of ideas and some communication and partnering outside meetings, we found these aspects to be much more limited. In terms of how participants enacted the CholeQuIC programme theory (see Figure 2) locally, our data analysis indicated that all participating sites attempted to follow the recommended steps, but with varying degrees of fidelity. For example, some sites focused much more on using local data collection to drive improvement than others. Through comparison of the case-study hospitals, we next examine this variability, and its positive and negative consequences, in more detail.

**FIGURE 6.1- Presence or absence of main influences on successful improvement in case study sites**
CholeQuIC = Cholecystectomy Quality Improvement Collaborative NPT = Normalisation Process Theory

| Adapted NPT framework | | Coherence | Relational | Operational | Learning | | Demand Management | | Creating Additional Capacity | |
|---|---|---|---|---|---|---|---|---|---|---|
| Key influence | Site # | Clarity of purpose amongst **all** key stakeholders | Capacity (time and resources) to lead and effective team working / project support | Ideas to action | Learning from self | Learning form others | Commonly agreed pathway | Co-ordinating patients | Utilising both elective and emergency lists | Sufficient surgeons and theatre space for urgent cholecystectomies |
| Highly successful sites | 1 | + | + | + (early) | + | + | + | + | + | + |
| | 2 | + | + | + (early) | + | + | + | + | + | + |
| | 3 | + | + | + (early) | + | + | + | + | + | + |
| | 4 | / | + | + | + | / | / | + | / | / |
| Challenged sites | 9 | / | + | / | + | + | / | + | / | - |
| | 10 | - | / | - | - | - | - | - | / | / |
| | 11 | - | - | - | - | - | - | - | / | / |
| | 12 | - | / or + (late) | / | / | + | / | / | / | - |
| Key: | | - Influence minimally or not present | / Influence somewhat or variably present | + Influence clearly present and visible | | | | | | |

*6.3.4 How was CholeQuIC enacted locally by teams and what influenced success?*

The analyses highlighted the extent of cognitive, relational and behavioural work (the 'enactment' of the CholeQuIC programme locally) that site leads and their teams needed to do to improve care for this patient group. This work is described under six descriptors of key influences, alongside their related area of improvement work and related Normalisation Process Theory constructs (Table 6.1). These tasks appeared to relate directly to hospitals' success in achieving the project goal: the four 'highly successful' sites achieved the tasks effectively, while the four challenged hospitals appeared to struggle with them. I describe each separately, although they are interdependent. In particular, no single set of features could be credited for success or lack thereof; rather it was the combination of their presence or absence, and the interaction between them, that appeared important.

*Clarity of purpose amongst site leads and other stakeholders*

Clarity of purpose was much better established in the highly successful than the challenged sites. In some Quality Improvement projects stakeholder engagement may be 'desirable' rather than necessary, but here it seemed vital, as diverse stakeholders were key to creating capacity and unblocking access to theatre lists. In one highly successful site, where the clarity of purpose was palpable amongst key staff we met, the use of a patient story seemed to galvanise everyone into action,

"*I talked about my patient, who waited 18 months to see me in clinic, which is not an unusual wait here in [area] and when I met her she was having an attack of biliary colic in the waiting room, so I admitted her. My clinic is on a Monday, I operated on her the following day… Bringing her along to some of my colleagues on our clinical governance day*

*and just having her speak about her experiences has helped put CholeQuIC right in the forefront of people's minds."* [Focus group, Highly successful group]

In the other highly successful sites, evidence from field notes indicated a shift in culture and behaviour from willingness to change 'in principle' towards modified processes that were endorsed by a range of stakeholders. Leads in two sites deployed a patient story along the lines above; in all four, engagement was a clear strategy of the site lead, extending well beyond emails and convening stakeholder meetings. Leads used multiple strategies including meetings, one-to-one conversations, data feedback, and opportunistic moments (e.g. corridor conversations) to engage colleagues.

Participants from both successful and challenged sites stressed that gaining common understanding and support for the work was difficult. A key point of divergence was that clarity of purpose was absent in all challenged sites in at least one key stakeholder group, i.e. amongst surgical colleagues, senior service managers or those gate-keeping emergency theatre lists. In challenged sites attempts were made to engage with necessary stakeholders, but the clarity of purpose visible in the highly successful sites remained absent. The challenge seemed greater in the larger centres, where the size of these organisations, and the presence of multiple surgical teams within the two specialist centres, meant that achieving coherence across the whole group of surgical stakeholders was harder. In one site, for example, the reported attitude from one key colleague in a different surgical sub-specialty team was:

"*We haven't got a problem, so we don't need to change".* [Focus group, Challenged group, emphasis added]

Field notes also suggested factors that may have affected clarity of purpose. In one site failure to agree on how to treat these patients made progress near impossible; in another, recurrent organisational challenges, both financial and patient-flow-related, meant that creating extra surgical capacity was not a priority for the responsible managers. Again, the challenge seemed greater in the larger centres.

"*Despite what seems like quite a bit of progress, what comes across most in his talk is that 'we have a lot to do to change attitudes'. Key issues seem to be that it's difficult to get into theatre (they are competing with other cases, e.g. cancer, and if they do get slots these tend to be late when it's not safe to operate), there's a reluctance to prioritise emergency patients over 'long waiter patients'.*" [Ethnographic notes, July 2017 meeting]

*Capacity to lead and effective team working / project support*

We identified a divergence between highly successful and challenged sites in their success in ringfencing time for the project. In all highly successful sites, time for the project was included in the lead's job plan; conversely in two challenged sites this was never achieved, and in the other two it was achieved only later in the project. Leading any project alongside existing clinical commitments can be challenging, especially when the lead has a role in motivating and encouraging others. As a highly successful site lead put it,

"*I'm glad I put all this work in…it wouldn't have worked otherwise, but… I'm exhausted*"

[Field note, highly successful group, November 2017 site visit]

In one challenged site, the lead applied to include dedicated time for CholeQuIC in his job plan but this was denied; unsurprisingly he saw this as not only a practical disadvantage but also a signal of the limited commitment of senior management to the project. No site lead attempted to make change happen single-handedly; all leads built teams or support groups of varying sizes and composition. A pattern was apparent, however, in how these teams worked toward project goals. In more successful sites professionals behaved as a coordinated, interdependent team, rather than as a working group or simply colleagues working in parallel. For example,

"*One thing that [site lead] was very clear about was the Friday morning meetings. He said that although they all communicated during the week about patient co-ordination, they tried (generally successfully he says) to meet briefly every Friday morning at the start of the day to catch up on more of the ongoing improvement activities."* [Field note, Highly successful group, November 2016 site-visit]

"*At the final event they depicted their work with a picture of a rugby team in order to illustrate that they had 'successfully managed to get a good team together and that actually we're pleased with ourselves for managing this'*." [Ethnographic notes, January 2018]

One challenged site did have a similarly high functioning team by the end of the programme. However, it only developed later in the project, seemingly limiting progress earlier on.

"*This has been a huge learning process for me along the way, to try and lead it […] so I got some nursing time from one our very senior nurses […] And, between us I thought that the two of us could then go out into the wider [site] audience, and we would be able to manage that. And I think that was probably the wrong approach. We probably should have engaged right from the beginning with a wider team, we should have had a team of four, five, six people.*" [Focus group, Challenged group]

*Turning ideas into action*

The CholeQuIC approach focused on helping teams to develop (or adapt) solutions to overcome local challenges and test them to see if they would work in practice, following the iterative Model for Improvement approach (30). In summary this approach entails: 1) using data to understand the current position; 2) defining an improvement goal; 3) developing ideas to get from current position to the goal; 4) testing ideas in mini 'real world' experiments; 5) refining and further testing ideas that work and discarding ideas that do not; 6) implementing the most effective ideas and continuing to use data to monitor (30).

A notable point of divergence between highly successful and challenged cases was the speed with which the former put ideas into action. In the challenged cases this translation process was much slower, or absent altogether. Four months into the project, three highly successful sites were already submitting data that showed successful moves toward the collaborative goal. The fourth was actively testing out ideas. For example, in one site the team very quickly agreed to trial the ring-fencing of elective slots for emergency cholecystectomies to examine the impact.

"*Both of us said, 'Look, why don't we keep slots free on our lists?'. My list is a Tuesday, [other surgical lead] is a Thursday… So I didn't speak to anybody outside that [immediate theatre team] because my experience of NHS management is if you ask permission then you're waiting six months for an answer. So very much do it, then seek forgiveness."* [Focus group, Highly successful site]

The importance of maintaining this willingness to test and adapt beyond the initial change was also identified; in another highly successful site, not only had staff turned their initial ideas into action early on, but continued to revise and refine their new process over time.

"*Minor process changes were introduced in response to data review and discussion at collaborative meetings, including additional training on the pathway during staff rotations [and they] utilised the 'Whiteboard' idea [from another site*].

"*In July 2017, they made the bold decision to move from elective lists for admitted patients [first change] to using held slots on CEPOD [emergency theatres]…From November 2017 the team are looking to introduce an 'as needed system' of pulling an elective list when demand increases over a two-week period…"* [Field notes, Highly successful site, October 2017 site-visit]

Conversely, in the challenged sites, a lack of early action was evident. This did not seem to reflect a lack of desire to improve care processes; rather, a combination of contextual factors and a reliance on a slower, more methodical planning approach was apparent. For example, in one site data from field notes and meetings showed that several months were spent designing, agreeing and planning the implementation of a new pathway, with

associated paperwork. Unfortunately, however, the pathway proved unsuccessful: colleagues simply did not use it. The project team found it hard to recover from this setback, particularly in the context of a time-limited project, and by the end of the project, little progress toward to the CholeQuIC goal had been made.

*Learning from own and others' experience of change*

Learning from sites' own data was a cornerstone of the CholeQuIC Programme Theory. Whilst all teams collected and collated data for the project, the perceived importance of this data varied. In most highly successful sites, data was collected almost contemporaneously and then reviewed to track progress.

"*Their goal has been to have 80% of patients [in surgery] within eight days and, they tell us, they're heading in the right direction to achieve that. This is at least in part due to data informing how they organise the service and guiding them to focus on "gaps between goal and reality".* [Ethnographic notes, July 17 meeting]

Data used was not always in the form prescribed by the CholeQuIC core team. Field notes indicated that in three of the four highly successful sites a variety of data and other local intelligence was used, including using coding data and more traditional theatre log book checks, to monitor new processes and how these were working. For the challenged sites, data collection and analysis were deprioritised, so that the information available was retrospective in nature, rather than providing timely and actionable insights into the impact of activities. In some sites, collecting data seemed to be an activity undertaken for by RCS project team rather than for teams themselves to analyse and monitor progress.

"[Site lead] *said how useful it was to review their data. I asked how often they had been doing this and he, a little apologetically, replied that they had not really had the time to do this by themselves."* [Field note, Challenged group, October 2017 site visit]

*Creation of additional capacity for emergency cholecystectomies*

Besides the four activities discussed above, relating to the NPT constructs of cognitive and relational and behavioural work, we identified two further influences that distinguished highly successful from challenged sites, relating to clinical processes. First, creating additional surgical capacity was essential. Approaches varied between sites. Successful sites tended to use a dual strategy, ringfencing elective slots for emergency work whilst simultaneously working on engaging colleagues in theatres with the concept that some emergency cholecystectomies belonged on the emergency theatre lists, even though historically these cases had been afforded much lower priority. This two-pronged attempt to create capacity –repurposing elective space plus optimising use of emergency theatres – appeared to produce the most successful results.

Conversely, in the challenged sites, difficulties in creating additional capacity were a major barrier. Here, competing clinical priorities prevented the addition of more emergency cholecystectomies onto already overburdened emergency lists, and made it difficult to find suitable elective lists that could be ringfenced or repurposed for these procedures.

*"*Moderator: *So you said that the ring-fenced capacity has helped but I think you also alluded to the fact there's nowhere near enough capacity. So, what's getting in the way of getting more capacity? [Everyone laughs]*

Participant: *Everybody in the hospital wants more capacity…"*

[Focus group, Challenged group]

*Managing and co-ordinating demand across an agreed pathway*

Second in terms of clinical processes, the four highly successful sites all succeeded in reaching agreement across stakeholders on the appropriate clinical pathway for this patient group and their flow along the pathway. For example, in one site, a shared understanding of the pathway and an effective mechanism for patient coordination were in evidence, through the use of a simple whiteboard in the surgeons' office,

"*The biggest success that we've had has been the CholeQuIC board that [team members] came up with and it was just a board… we put anybody, any patient, on there with a putative diagnosis of acute biliary disease, whether it's right or wrong doesn't matter, it's about getting them up there and then scrubbing those patients out a couple of times a week and saying 'yes this one is, no this one isn't'[...] as the patients get identified and targeted."*

[Focus group, Highly successful group]

Conversely, in three of the four challenged sites, neither a pathway (whether formally documented or informally understood) nor mechanisms of coordination between different

parts of the service were present. In one larger challenged site, recognition of the problem of co-ordination led the lead to build a business case for a biliary co-ordinator, a role that two other (successful) CholeQuIC sites had created. However, by the end of the project, no-one had been appointed and the issue of coordination remained. In another challenged site, a pathway was agreed amongst consultant surgeons, but to the lead's frustration, it was not followed in practice.

"*In fact, many of the surgeons were carrying out practice which ensured that outcome goals could not be met. One of the frustrations outlined by the project lead at the final project meeting was that some surgeons are sending people for MRI scans [outside the agreed pathway], which regularly took over a week to take place."* [Ethnographic notes, January 2018 meeting]

### *6.4 Discussion*

This research described in this chapter shows that a Quality Improvement Collaborative approach can be effective at reducing time to surgery for patients with acute gallstone disease, but that making the approach work is complex and challenging. Our study is distinguished by the use of a robust mixed-methods design which demonstrated the overall impact of the approach compared to a contemporaneous control group and highlighted the differential effectiveness of the approach across 12 participating hospitals. This enabled us to identify aspects of implementation and context associated with greater impact in four 'highly successful' hospitals demonstrating the most statistically significant change, and less impact in four 'challenged' hospitals that were least successful by this measure. Our framework, guided by Normalisation Process Theory,[113, 145] but populated through

comparative analysis of data from the cases of highly successful and challenged sites, suggested six sets of influences that seemed most consequential. Intensive work was required to ensure that all key stakeholders had a shared understanding of, and agreement with, the purpose and benefits of rapid surgical intervention; where this was in doubt, achieving improvement was more challenging. However, clarity of purpose was a necessary, but not sufficient, condition for improvement. Sites also systematically diverged in their handling of practical issues, such as protected time within job plans, functional team-working, and rapidly turning ideas into action. It was a combination of these influences that characterised the highly successful sites in CholeQuIC. Other key factors for success, more specific to emergency surgery and not so readily accounted for within our Normalisation Process Theory-informed framework, included a multi-pronged approached to creating additional theatre capacity, and agreeing a clear pathway with effective coordination mechanisms.

Our findings, interpreted in light of relevant theory and research, allow us to offer some transferable lessons for other practitioners. Firstly, our findings contribute to the growing evidence of what influences the effectiveness of Quality Improvement Collaboratives.[58, 95] Engagement in CholeQuIC was good, with consistent attendance and involvement by between one and six staff from each site at every meeting. In addition, the participants valued the social aspects of the collaborative (e.g. meeting up with like-minded individuals), the external driver for change it provided and the legitimacy conferred on it by the Royal College of Surgeons. However, there was no strong evidence for the level of sharing, partnering and cross-pollination of ideas found to be key mechanisms within some Quality

Improvement Collaboratives.[78, 146] This leads us to tentatively suggest that some 'simpler' Quality Improvement problems, such as that addressed in CholeQuIC, may be just as effectively addressed using lighter-touch Quality Improvement programme approach with fewer meetings and less emphasis on inter-site collaboration. However, the importance of the social aspects probably precludes a move to a remote-contact only 'campaign' approach.[147]

At the site level, seen through the lens of Normalisation Process Theory, there was a certain linearity to the improvement process during CholeQuIC. Achieving clarity of purpose (coherence) has to be the initial step in any work contingent on the actions of multiple stakeholders, followed by efforts to enrol those stakeholders and legitimise the change (cognitive participation). Colleagues have to be willing to see changes made to practice and seniors have to be willing to allocate time and resources to allow a project leader to drive changes through. Thus gaining the support of the organisation for the project required that sufficient numbers of key stakeholders viewed the project goals as aligned with their own. In turn, coherence and cognitive participation shape the ability to collectively act to make change happen. Using data to learn from changes made and monitor progress was also required, but ultimately appeared to be the easiest part of the process.

This finding diverges somewhat from some recent thought on the challenges of improvement in complex environments.[129] It suggests that sometimes, there may be relatively linear routes to change that are likely to achieve success: it is gaining support for and momentum along these routes that is crucial. In itself, the key change in practice that

CholeQuIC required was comparatively simple: the goal was within the gift of surgeons and managers if they all believed it was the right thing to do and if contextual pressures did not present issues which took priority over the CholeQuIC goal. If these conditions were met, other changes would follow relatively easily.

This linear process, however, was more easily activated in some organisations than others. It is noteworthy that three of the challenged hospitals were the busiest in terms of surgical volume; conversely three of the highly successful hospitals had the lowest surgical volume. The size of the challenged organisations, and the presence of multiple surgical teams, meant, first, that achieving coherence across the whole group of surgical stakeholders seemed harder. Second, surgical throughput may be an important constraint on improving practice in an area which, as noted above, relies to a large extent on leads' ability to make capacity available for extra emergency procedures. The results across the entire cohort (Table 5.1) demonstrate high volume centres can achieve significant improvement in care for this patient group, but it should be recognised that challenges in doing so may well be greater than in lower-volume sites.

Another key point of divergence between highly successful and challenged sites was their willingness and ability to turn ideas into action and in particular to do this early on in the project. In Normalisation Process Theory, the Collective Action component suggests the importance of both the ease with which new processes can be adopted (interactional workability) and their fit within the local workflow and context (contextual integration). The challenge in Quality Improvement is to find solutions that are workable and easy to adopt

whilst also improving patient care and outcomes. Our findings point to a potentially effective variation on the widely used iterative approach promoted by the Model for Improvement (30); there needs to be recognition of the time required for generating thoughtful potential solutions, based on an understanding of local systems and context, but this needs to be combined with a subsequent willingness to get on and test these, refining, adapting or discarding as appropriate, in an action-oriented and iterative manner. In our case studies the time spent on deliberation upfront appeared to make for better solutions that needed fewer rounds of testing, but that an openness within the team to testing and iterative adaption was also vital. This aligns with recent thinking in complexity in healthcare organisation, which suggests that the multiplicity of agents involved in any change effort, and the unpredictability of interactions between different parts of a dynamic system, may frustrate even the most thoughtfully developed plans.[129] In such circumstances, acting "scientifically and pragmatically" through a trial-and-error-based approach, of the kind exemplified by the highly successful sites in CholeQuIC, may be a more effective way of finding a solution that fits local circumstance.[133] Learning from rapid cycles of improvement is a specific skill and provides the foundation for the concept that local Quality Improvement data holds valuable lessons and supports a 'turning ideas into action' mindset. Technical skills in iterative testing and analysing time-series data may be useful, as are the communication skills needed to generate clarity of purpose and motivate colleagues to change. A variety of capabilities are thus required to achieve success in improvement.[108] Quality Improvement programme designers should be mindful of developing programme-level interventions, such as training and coaching, to support the development of those tasked with leading Quality Improvement projects at the frontline.

This evaluation has several strengths, including its mixed-methods approach, drawing on a wide range of data to add deeper understanding to the findings of the quantitative evaluation, and the use of a partnered evaluation, capitalising on the rich knowledge of the project team whilst using external oversight to maintain scientific rigour. In particular, the use of disaggregated data demonstrating the differential impact of CholeQuIC across sites allowed us to examine our qualitative data for systematic differences between the most and least successful sites, such that we could highlight those factors consistently associated with better and worse performance. This enabled us to provide recommendations for others in a field that until now has been characterised by very limited understanding of the 'active ingredients' of successful collaboratives, and the work needed to make them work.[58]

The evaluation also has several limitations. Analysis and interpretation of data took place in the light of the identification of the highly successful and challenged groups, and so was guided by (rather than blind to) the results of the quantitative evaluation. There is also a risk of bias in data collection and analysis by those directly involved in running the collaborative. The choice to use an internal evaluation model was a pragmatic decision based upon the limited funding envelope for the study and weighing the in-depth knowledge of the evaluation team regarding the project against the challenges in maintaining a degree of dispassionate objectivity as researchers. This is a known tension with process evaluation work and an ongoing challenge due to limited funding for process evaluations.[72, 115] This was mitigated by partnering with external research expertise (GM), who offered critical peer review from a more detached position with no investment in the project being successful or

not. I also remained cognisant of the challenges inherent in holding the dual role of project leader and evaluator and ensured the team I was leading also recognised these challenges. Reflexivity was key to the 'de-biasing' process combined with meaningful efforts to consider rival explanations, including detailed presentation of the chain of evidence for emergent findings at data meetings with our external expert and a member checking exercise, where the manuscripts for both CholeQuIC publications were sent to the participants prior to submission for publication (only very minor changes to aid clarity were made following this exercise). We also deployed a narrow, prospectively defined measure of success to guide analysis. Accordingly, we sought to identify those factors with the most consistent apparent relationship with the impact of CholeQuIC, whilst also recognising that other plausible explanations may exist.

### *5.5 Conclusion*

Collaborative-based quality improvement is a viable strategy for emergency surgery, but its impact rests on the deployment of both effective clinical and improvement strategies by project leads with their colleagues. Achieving clarity of purpose about the proposed changes amongst key stakeholders is a vital precursor to improvement, while protected time and support to enact improvement solutions, and the ability to learn from the experience of doing so were also associated with greater impact within this cohort. We found the use of objective performance data to identify successful sites, and a theoretical lens to interpret the data, helpful in understanding what works within surgical quality improvement, and would recommend this as an approach for improvement project evaluations.

# CHAPTER 7: Discussion

This final chapter offers overarching reflections on the work presented in the PhD. I begin with a summary of findings, then discuss my interpretation of these findings and how they address the main research question of this thesis. I end with reflections on the research process and recommend areas for future research.

## *7.1 Summary of main findings*

The principal finding from the EPOCH trial (described in Chapter 3) was that there was no survival benefit associated with a national Quality Improvement programme to implement an evidence-based care pathway for patients undergoing emergency abdominal surgery. Furthermore, there was no beneficial effect on 180-day mortality, hospital stay or hospital readmission. At a national level, there were only modest improvements amongst the ten measures selected to reflect key processes of care within the pathway. This suggested that implementation failure was the main cause of the lack of effect upon patient outcomes.

The principal finding from the EPOCH trial process evaluation (described in Chapter 4) was that implementation failure likely stemmed from a goal (pathway implementation) that was too ambitious for the time and resources that local QI leads had available to them. There were only 11 clinical processes which more than half of teams attempted to improve from the clinical pathway (the hard core of the intervention) and only half of the trial cohort reported using five or all six of the Quality Improvement strategies (the soft periphery of the intervention) designed to enable pathway implementation. Ethnographic findings indicated that QI leads predicted, and subsequently experienced, multiple and often significant challenges as they attempted to lead change in their hospitals. These challenges seemed to

shape, to greater or lesser extents, which components of the pathway they chose to focus on first and how they approached implementation. Major implementation barriers included limited time and scarce resources to support QI leads and, connected to this, an onerous burden of data collection which limited capacity to subsequently use these data for improvement. In particular, the effective use of data for improvement involved a substantial social aspect, to help colleagues understand and be motivated by the data, but many QI leads found this aspect challenging as well as time-consuming. Analysis of individual hospital level improvement using time-series (Chapter 5) found that no hospital in the EPOCH trial reliably implemented the care pathway within six months of the end of the intervention period.

Some areas of improvement were identified. In total, 279 (of a possible 800) care processes were improved by hospitals through participation in the EPOCH trial and a small group of hospitals (17.5%, 14/80) were successful in improving more than six care processes. Effect sizes overall were marginal, but with substantial variance for each process across trial hospitals. The hospital teams in our study that achieved greater care-process improvement also reported using more of the implementation strategies recommended by the Quality Improvement programme, suggesting improvement was possible using a multi-faceted approach to improvement. The main findings from our evaluation of the CholeQuIC collaborative (Chapters 6 & 7) were that that 8/12 participating hospitals were able to significantly increase early cholecystectomy rates for patients with gallstone disease requiring hospital admission, in line with national guidance. Results remained significant when the small national improvement trend was accounted for. When plotted on SPC

charts, locally collected data showed a clear reduction in variability and increase in reliability in providing timely laparoscopic cholecystectomy to this patient group across the study cohort. We identified aspects of implementation and context associated with greater impact in four 'highly successful' hospitals demonstrating the most statistically significant change, and less impact in four 'challenged' hospitals that were least successful by this measure. Intensive work was required to ensure that all key stakeholders had a shared understanding of, and agreement with, the purpose and benefits of rapid surgical intervention; where this was in doubt, achieving improvement was more challenging. However, while clarity of purpose was a necessary condition for improvement, it was not sufficient alone. Successful and challenged sites also systematically diverged in their handling of practical issues, such as protected time within job plans, functional team-working, and rapidly turning ideas into action. It was a combination of these influences that characterised the highly successful sites in CholeQuIC. Other key factors for success, more specific to emergency surgery, included a multi-pronged approach to creating additional theatre capacity, and agreeing a clear pathway with effective coordination mechanisms.

## *7.2 What influences the effectiveness of Quality Improvement in emergency general surgery in the NHS?*

In this section I integrate and synthesise the findings on what influenced the effectiveness of the two large Quality Improvement programmes detailed in this thesis in order to provide more transferable findings that may be applicable to future Quality Improvement in emergency general surgery in the NHS. The synthesis is based on a modified Thematic Synthesis approach, which involves review and free-coding of primary data, development of

descriptive themes and then organising these into higher-level (or 'third order') analytical themes.[148] As the number of primary studies for synthesis was small, and were coded primarily by myself, I did not formally undertake the first step (although I was  it was open to this possibility during the review) and the process was undertaken manually, using a large table and post-it notes, rather than using a data organising technology such as Nvivo. Each chapter was read and re-read (plus returning to source data or documentation as necessary, although ultimately no re-coding was done) and the key influences from each work organised into descriptive themes. Following this, the relative influence of these key themes, was considered within and across the two QI programmes, and also how these influences may have interacted, or have been inter-related, within the programmes. These were then aggregated up into high level analytical themes. This process would have been easier had both qualitative analyses been theory based (i.e. the CholeQuIC analysis was informed by Normalisation Process Theory and EPOCH analysis was structured around the trial programme theory only). Notwithstanding this, the findings of the synthesis process are presented below. I first present two major influences that reinforce and add further detail to those already present within the extant improvement and implementation literature. These are undoubtedly important influences but, from the perspective of this as a doctoral thesis, are not entirely novel findings. Following this I present in more detail two emergent influences on the effectives of QI in surgery, currently not discussed in the extant literature.

The evaluations identified the need for a multi-faceted approach to improvement efforts. It is not just effectively engaging colleagues,[105] or good use of audit and feedback,[52] or iterative testing of ideas in the Plan-Do-Study-Act (PDSA) style,[56, 133] but rather a carefully

considered combination of these approaches in concert and often more besides. This finding is not novel and indeed was integrated into the programme theories for both EPOCH and CholeQuIC (Figure 2.3 and Figure 6.1), and the guidance provide to teams, with the latter already being informed by learning from the former. Findings from the evaluations of both suggested that those who were able to follow this guidance were able to achieve greater process improvement. In the both the EPOCH and CholeQuIC evaluations we did not see any superiority of certain Quality Improvement methods over others, but rather it was the *combination of methods* that was associated with improvement. However, the CholeQuIC findings provided a nuanced perspective on this, as stakeholder engagement appeared as a key necessary precursor to any subsequent improvement efforts. Therefore, the first major influence is **using a multi-faceted approach to improvement that focuses the application of multiple improvement methods around achieving stakeholder engagement**. Viewing Quality Improvement methods as a set of social, or socio-technical, interventions is recognised in the literature (e.g. [60, 92]), but the relative importance of the social aspect may not be emphasised sufficiently in tandem with the technical aspect. For example, a recent and well conducted study of data usage in Quality Improvement identified measurement as challenging technical task, requiring a degree of expertise that healthcare staff underestimate, and often do not possess.[149] **This is a pervasive issue that needed highlighting and was a common theme between both EPOCH and CholeQuIC.** No mention was made, however, of the challenges in sharing that data with colleagues in a way that creates and sustains engagement with improvement and the skills required to do that effectively. Yet the 'feedback' element of audit and feedback, the social element, is likely just as important if it is to support improvement because healthcare processes are generally

the result of human behaviour.[52, 134, 150] There are data on how to optimise feedback for individuals but relatively little on how to tailor this for teams despite healthcare, including emergency general surgery, being very much a team-orientated endeavour.[55] Regarding other Quality Improvement methods, a recent evaluation of the application of the PDSA cycle identified both technical and social challenges.[57] Technical challenges appeared surmountable over time with coaching and increased experience. Social aspects, such as engaging multidisciplinary stakeholders in the PDSA process and the outcome of each cycle, appeared much harder to achieve in parallel to acquiring the technical skills in PDSA application. **The second major influence can be described as paying sufficient attention to the technical and socio-technical aspects of Quality Improvement methods, in particular feedback of data to colleagues.**

Two additional influences that have emerged as important from the analyses have been added to summary of evidence-based Quality Improvement presented in Chapter 6 (in **bold** in Figure 7.1). These two new, emergent influences as they are less often accounted for and described in the literature yet are highly relevant to both frontline teams and Quality Improvement programme planners.

**Figure 7.1: Important influences on the effectiveness of QI in emergency general surgery**

*Choosing the right problem to be fixed*

- **Choosing a quality problem that is amenable to Quality Improvement**
- Choosing a problem with common agreement that it needs fixing (defined by stakeholders)
- Clearly defining and articulating the problem

*Appropriate measuring and monitoring*

- Parsimonious data collection yet sufficient to understand the size of the local quality problem
- Ongoing data analysis and feedback to monitor progress and motivate colleagues over time

*Sufficient support and collaboration*

- Sharing of ideas and outcomes with others
- Expert clinical and QI support, training and coaching
- **Time provided to lead QI efforts locally**

*Developing and testing the right solutions*

- Generating context specific solutions (supported by evidence of any previous solutions)
- Testing these solutions, and adapting to what works well or does not

The first emergent influence on Quality Improvement effectiveness regards the type and scale of the quality problem to be fixed. In term of the type of quality problem, the Donabedian assessment of structure, process and outcome remains very useful today.[68, 151] **The findings of this thesis suggest that Quality Improvement programmes are suitable to address *process* problems but less likely to improve structural ones.** This is most likely because clinician- led Quality Improvement has the opportunity, if approached and resourced correctly, to change local processes and colleague behaviour. As such, improvements may be within the gift of those stakeholders involved to change or influence. Conversely, structural problems are generally 'money' problems that require much larger scale changes (e.g. including business cases to apply for more funds). These may be much less likely to be amenable to direct action from clinicians to change.

In the hospital level time series analysis of the EPOCH trial, care processes such as pre-operative risk assessment (57/80 [71%]), intra-operative lactate measurement (42/80 [53%]) and cardiac-output guided fluid therapy (32/80 [40%]) were the most frequently improved process measures (Table 5.2). Conversely, those processes that were potentially constrained by a structural element too, such as consultant-led decision-making (14/80 [18%]), consultant review before surgery (17/80 [21%]) and time from decision to operate to surgery (14/80 [18%]) were the least likely to improve (Table 5.2). In CholeQuIC we saw a similar, if more nuanced picture. Ostensibly the CholeQuIC quality problem appeared very similar to one difficult process within EPOCH, the time from decision to operate to surgery. However, the difference in the case of CholeQuIC was that those patients who required laparoscopic cholecystectomy were generally likely to have this surgery electively *at some point in the future*, just not on an urgent basis. Thus the overall capacity for a hospital to deliver these operations largely existed anyway and the successful sites in CholeQuIC were able to repurpose this capacity in a way that enabled surgery to be delivered on an urgent basis. Thus, a structural problem (theatre and surgeon availability) became a process problem (effectively re-organising and managing existing capacity to deliver urgent surgery).

With a couple of notable, and very recent exceptions.[152, 153] there is surprisingly little explicit discussion in the literature regarding choosing the right type of problem to address with Quality Improvement, despite the fundamental importance of this choice. Additionally, even when a problem is correctly chosen, Quality Improvement programme designers need to consider the scale and scope of the problem to be fixed against the resources provided by the programme to support local efforts. Such resources include Quality Improvement

training, data analysis support and meetings to facilitate the sharing of experiences and potential effective solutions. The EPOCH trial was designed to mirror Quality Improvement programmes popular in the NHS at that time,[91, 92, 154] with minimal in-person contact and support from QI experts. This type of light-touch Quality Improvement programme would later be defined as campaign-style Quality Improvement, to differentiate it from the more intensive Quality Improvement collaborative approach.[147] The learning from EPOCH, and subsequently from systematic reviews of Quality Improvement programmes and collaboratives, is that campaign-style Quality Improvement is only suitable for (very) simple quality problems, limited in their scale and scope, and process orientated in nature (as above).[58, 59] The same review found that for larger-scale, process-orientated, quality problems, more intensively resourced and supported Quality Improvement collaboratives may still be effective and appropriate to use and the robust evidence provided by the controlled evaluation of CholeQuIC would support that conclusion. In summary, there are lessons for those commissioning and designing large-scale Quality Improvement to improve emergency surgery and those wishing to attempt Quality Improvement locally. Both need to be aware of the limitations of the potential of Quality Improvement to improve quality. Future efforts should focus solely on process-orientated problems that are of a scale suitable to be addressed by clinician- led projects with appropriate levels of support and resource available.

**The second emergent influence identified was the amount of time QI leads had available to consider, lead and actually 'do' improvement in these respective projects. A lack of sufficient time appeared to have a negative impact whilst those that managed to secure protected time appeared to be able to lead more effective QI projects.** This may appear a

fairly mundane observation. However, the influence of clinician time appeared fundamental to the ability to improve care in both EPOCH and CholeQuIC, yet it is rarely discussed in the literature or Quality Improvement guidance. For example, guidance discussing the Model for Improvement fails to mention that application of this model to structure improvement efforts may require a substantial time commitment at both the planning and action stages.[41, 57] Similarly theoretical models for improvement and implementation do not foreground the influence of time in a meaningful way. The Normalisation Process Theory alludes to the Collective Action construct as being the 'operational work' of improvement but does not specifically reference the influence of time.[113] The widely used Consolidated Framework for Implementation Research mentions time in just one of over 20 constructs (Readiness for Implementation, Available Resources).[112] Taking as an example one of the most successful Quality Improvement programmes in history, the Michigan Keystone project, the original publication makes no reference to the significant time commitment of physicians and nurses that was required.[74] A subsequent theoretical analysis of the programme's success did detail the time commitment required but did not mention this as a major influence.[60] I suggest these omissions, and many others across the improvement and implementation science literature, have created a situation where improvement projects are seen as relatively easy 'quick-fix' endeavours. It is not hard to see, therefore, why proponents of Quality Improvement have been accused of 'magical thinking' by those taking a hard look at the evidence for if and how Quality Improvement approaches actually work to improve care, or not.[81, 155] This issue is both pervasive and pernicious because the 'quick-fix' mentality of Quality Improvement often prevents those tasked with leading improvement asking for or being granted protected time to lead and do improvement. Sadly, this means

these clinicians are perhaps set up to fail, or certainly not achieve the full potential of their improvement efforts, from the outset. The influence of time as a necessary resource for any serious improvement efforts needs much greater attention in the future although this in turn has potential implications for the funding and conduct of Quality Improvement studies. Acknowledging the time required will necessarily require additional funding for such studies, to ensure clinicians have sufficient job-planned time to 'undertake the labour' required in improvement work. Given that at least some of this labour is leadership for improvement that is necessary for achieving stakeholder engagement, this requires Consultant-grade funding for the lifetime of any study. This has significant cost implications that may make Quality Improvement programmes less attractive to funders and senior NHS management without more substantial evidence of effectiveness.

## *7.3 Applying the learning from EPOCH and CholeQuIC*

As examples of well delivered and robustly evaluated Quality Improvement, the findings of, and learnings from, the EPOCH trial and CholeQuIC have been influential in subsequent efforts to improve the quality of emergency general surgery in the UK.

### *7.3.1 How has learning already been applied?*

Latest figures show around 25,000 patients undergo emergency laparotomy in NHS hospitals each year, with 30-day mortality rate reduced now to 9.6%.[156] Since the EPOCH trial there have been significant efforts focused on addressing the issues of high mortality following emergency laparotomy, both within the UK, with the National Emergency Laparotomy Audit (NELA) and the Emergency Laparotomy Collaborative, [156, 157] and internationally.[158] Starting immediately before the EPOCH trial, NELA is an ongoing major

national audit commissioned by the Healthcare Quality Improvement Partnership (HQIP) and funded by NHS England and the Welsh government.[14] NELA was one of the first national audits to provide real-time data to inform clinicians about the process of patient care and subsequent outcomes at both the local (hospital) and national (NHS) level.

The Emergency Laparotomy Collaborative (ELC) was a large Quality Improvement project funded by the Health Foundation, adopting implementation science to improve patient care.[157] The project was led by the Kent Surrey Sussex Academic Health Science Network (http://www.kssahsn.net) in 28 NHS hospitals, and used the breakthrough collaborative (BTS) model to help teams to implement a laparotomy care-bundle.[58, 159] Using Statistical Process Control (SPC) chart analysis, [139, 140] the ELC demonstrated an improvement in care-bundle compliance with a concurrent association with decreased 30-day mortality (to 8.3%) during the course of the ELC project. However, the ELC was designed as a Quality Improvement project rather than a clinical trial and the observational nature of the study and lack of control group, or a controlled evaluation, means that a causal relationship between the intervention and improved outcomes cannot be confirmed. This remains a common challenge in current improvement science research. [80, 99]

In the context of overall improvement in care for this patient group during the last decade, the Quality Improvement intervention in the EPOCH trial did not impact upon patient survival, and yet the research around the EPOCH trial can be seen to have catalysed learning about how we can improve care for this patient group using Quality Improvement.[160-162] Regarding gallstone disease, and improving delivery of emergency laparoscopic

cholecystectomies, the relative success of, and learning from, CholeQuIC stimulated the Royal College of Surgeons of England to run a second larger collaborative, CholeQuiC Extended Reach (CholeQuIC-ER). CholeQuIC-ER was a 24-site collaborative, with the same improvement goal and using the methods developed and refined during CholeQuIC, which ended in December 2020.

### *7.4 Reflections on methodology in Quality Improvement research*

As complex social interventions, Quality Improvement programmes are challenging to robustly evaluate because of adaptations and variations that necessarily occur as local teams try and 'fit' the QI intervention into their local context. Thus it becomes hard to know what exactly is being trialled and what are the 'active ingredients' that lead to success. This problem has dogged the field of improvement science, leading to scores of weak observational studies of QI with high risk of bias and the inability to robustly evaluate cause and effect.[58, 59, 78, 114] I have three main reflections on research methodology and the research process undertaken for this thesis.

First, my experience with the EPOCH trial conclusively demonstrates, to my mind at least, why trials or other outcome evaluations of complex interventions require concurrent explanatory process evaluations. Without the process evaluation in EPOCH, we would have identified that the intervention didn't work and that there was a large factor of implementation failure, identified by the lack of change in the process measures. Notwithstanding the interpretation of the process measure findings, we would have had no credible data on why implementation failure had occurred and therefore what could be learnt for next time. This would have essentially been a substantial waste of NHS funds and

left difficult questions about the status of future Quality Improvement efforts due to the uncertainty caused by the EPOCH findings. This learning had a direct impact on the CholeQuIC study as it persuaded the Royal College of Surgeons of England to fund the Quality Improvement collaborative and the concurrent evaluation. However, the future of studies (whether trials or controlled evaluations) of complex interventions such as Quality Improvement programmes or collaboratives may lie in iterative study designs where embedded process evaluations allow for learning about intervention implementation to take place 'in vivo' with subsequent pauses to allow for tailoring and adaptation of the intervention as necessary.[124, 163] This is also where the granular time series analysis used in Chapter 5 would be of most use, as it would provide programme leaders and local teams with intelligence on what aspects of the intervention are being improved and which are not. Whilst the iterative design poses methodological and logistical challenges, my experience over the course of writing this thesis suggests it would provide the best chance for studies of Quality Improvement to produce findings of maximal usage. The need to robustly test the effectiveness of an intervention, in situ, needs to be balanced with having an intervention that is likely to succeed and the iterative approach feels like the most likely approach to achieving this balance. Few studies have yet to adopt this approach however.[124]

My next reflection is on the value of a mixed-method approach to process evaluation, even if historically these evaluations have been considered agnostic with regards to research methods used.[72] Although used in different ways, a mixed-method approach added value to both the process evaluations in this thesis. In the EPOCH trial, I was able to augment the very detailed ethnographic data we had on six of the 93 trial sites with substantially more data from a much larger number of trial sites. In itself, the ethnographic data would have

been methodically robust but would have potentially left some questions about how typical the experiences of the six pre-selected sites were to the cohort overall. This is a challenge inherent in ethnographic and other in-depth qualitative study where data collection is necessarily context-situated and time-consuming. In contrast, the data collected from the Quality Improvement programme and from the questionnaire would have informed us about programme delivery both nationally and locally, and provided some (fairly weak) data on the perspectives and experiences of those leading the Quality Improvement locally. In combination, however, the depth of the ethnography findings and the breadth of the process evaluation, and the ability to compare and triangulate these data provided a rich set of findings that we could feel confident represented what happened during the EPOCH trial. The way in which research methods were 'mixed' in the process evaluation of CholeQuIC was different, because the data in the process evaluation was largely qualitative in nature but were able to use quantitative data from the outcome evaluation to shape and focus our case studies. In doing so, we were able to take the novel approach of re-organising the data we had collected and initially analysed from across the study sites and create the cases based upon their performance in improving care (the case of the successful and the challenged hospitals.). This then allowed a useful comparison of influences between the two cases, centred about objective data on how sites had actually performed during the collaborative. Learning from success as an evaluative paradigm or perspective isn't completely new, but has not been widely adopted within the field of evaluation as yet. As discussed in the limitations in Chapter 6, we had to work hard to minimise the risk of researcher bias that could be generated by knowledge of performance but the strategies we put in place to mitigate this enabled us to produce trustworthy findings.

Regarding outcome evaluation methods for complex interventions such as Quality Improvement programmes and collaboratives, I reflect on my experience in designing CholeQuIC where we had a national data-set, the Hospital Episodes Statistics database, which allowed us to use all other NHS hospitals that provided laparoscopic cholecystectomies as a control group. Controlled evaluations are superior to simple before and after studies because of the increased ability to detect the effects of an intervention, and to control for confounders and secular trends. The issue of the secular trend toward improvement that occurs concurrent to a Quality Improvement programme or 'rising tide effect' is one that has been highlighted by previous controlled improvement studies, including the Safer Patients Initiative and Matching Michigan.[76, 92] When we paired our outcome evaluation with the mixed-methods process evaluation we were able to produce robust findings on the effectiveness of CholeQuIC in reducing time to surgery, ahead of any national secular trend plus concurrent explanatory findings without the need for a trial. On reflection, I wonder whether we could have designed EPOCH using a controlled evaluation approach, at a similar national scale but using NELA data to create a control group of non-participant hospitals. This could have had some advantages over the stepped-wedge trial design, including the ability to give all hospitals substantially longer time-frame to attempt improvement (the longest period in the EPOCH trial was 80 weeks, with some hospitals only having a 10 week period). Improvement-focused national audits such as NELA may prove to be the most effective centrally organised approach to the improvement of complex quality issues due to the open ended time period (rather than a time-constrained 'project'), and also due to the potential impact on structural issues in the medium to longer term, such as the number of operating theatres, through national reporting and benchmarking. Wrapping

evaluations around such long-term projects may prove a highly informative and cost-effective approach to studying improvement efforts.

Another, more epistemological, advantage of a controlled evaluation is that improvement interventions may tend to change over time in response to required adaptations and on-going learning yet trial findings generally are founded on the assumption of a stable, well-defined intervention. As discussed above, a study design that allows iterative adaption of the intervention may have advantages in improvement science research. Notwithstanding newer trial designs, such as Sequential Multiple Assignment Randomised (SMART) trials,[164] an iterative approach may be more easily facilitated by a controlled evaluation design plus process evaluation than a randomised trial. The main disadvantages to this approach remain the challenges in attributing causation through non-randomised study designs, due to the increased risks of bias and confounding (most likely due to selecting poorly matched intervention and control groups). This issue is not insurmountable with careful study design and controlled evaluations with built-in scope for intervention iteration and adaptation could facilitate the generation of many more robust studies of Quality Improvement without the associated costs and logistical constraints of the randomised trial.

### *7.5 Future research opportunities*

With Improvement Science still in its infancy, there a clear need for more mixed methods research to learn how to improve the application, and evaluation, of Quality Improvement methods at the frontline. An ideal vehicle for this would be a controlled evaluation of a Quality Improvement collaborative with a concurrent ethnographic evaluation in multiple

participating sites. The ethnographic approach would be of specific benefit given the need to better understand how to focus improvement methods to achieve stakeholder engagement. Funding timelines should allow for an iterative study design, so that both barriers to the application of improvement and also subsequent solutions and adaptations could be evaluated, potentially even through multiple iterations. An evaluation of the impact of differing lengths or amounts of funding could also be incorporated, specifically to evaluate the impact of well-funded vs. minimally funded QI on the ultimate effectiveness of such programmes. We are developing a large body of knowledge regarding barriers to improvement and future research on how to effectively overcome these is necessary .These findings could be used to build practical yet evidence based guidance for clinical staff on the most likely routes to achieving effective Quality Improvement.

# REFERENCES


1. Abbott TEF, Fowler AJ, Dobbs TD, Harrison EM, Gillies MA, Pearse RM. Frequency of surgical treatment and related hospital procedures in the UK: a national ecological study using hospital episode statistics. *Br J Anaesth*. Aug 2017;119(2):249-257. doi:10.1093/bja/aex137
2. Weiser TG, Haynes AB, Molina G, et al. Estimate of the global volume of surgery in 2012: an assessment supporting improved health outcomes. *Lancet*. Apr 2015;385 Suppl 2:S11. doi:10.1016/S0140-6736(15)60806-6
3. Pearse RM, Moreno RP, Bauer P, et al. Mortality after surgery in Europe: a 7 day cohort study. *Lancet*. Sep 2012;380(9847):1059-65. doi:10.1016/S0140-6736(12)61148-9
4. group ISOS. Global patient outcomes after elective surgery: prospective cohort study in 27 low-, middle- and high-income countries. *Br J Anaesth*. Oct 2016;117(5):601-609. doi:10.1093/bja/aew316
5. Meara JG, Leather AJ, Hagander L, et al. Global Surgery 2030: evidence and solutions for achieving health, welfare, and economic development. *Lancet*. Aug 2015;386(9993):569-624. doi:10.1016/S0140-6736(15)60160-X
6. Fowler AJ, Abbott TEF, Prowle J, Pearse RM. Age of patients undergoing surgery. *Br J Surg*. 07 2019;106(8):1012-1018. doi:10.1002/bjs.11148
7. Nepogodiev D, Martin J, Biccard B, Makupe A, Bhangu A, Surgery NIfHRGHRUoG. Global burden of postoperative death. *Lancet*. 02 2019;393(10170):401. doi:10.1016/S0140-6736(18)33139-8
8. Finks JF, Osborne NH, Birkmeyer JD. Trends in hospital volume and operative mortality for high-risk surgery. *N Engl J Med*. Jun 2 2011;364(22):2128-37. doi:10.1056/NEJMsa1010705
9. Shah N, Hamilton M. Clinical review: Can we predict which patients are at risk of complications following surgery? *Crit Care*. May 7 2013;17(3):226. doi:10.1186/cc11904
10. Sankar A, Beattie WS, Wijeysundera DN. How can we identify the high-risk patient? *Curr Opin Crit Care*. Aug 2015;21(4):328-35. doi:10.1097/MCC.0000000000000216
11. Anderson I, Eddlestone J, Lees N. *The Higher Risk General Surgical Patient - Towards Improved Care for a Forgotten Group.* 2011.
12. England TRCoSo. *Emergency Surgery: Standards for unscheduled care*. 2011.
13. Ireland AoSoGBaN. *EMERGENCY GENERAL SURGERY: THE FUTUREA CONSENSUS STATEMENT*. 2007.
14. Team NP. The First Patient Report of the National Emergency Laparotomy Audit. London: RCoA; 2015.
15. CholeS Study Group WsMRC. Population-based cohort study of outcomes following cholecystectomy for benign gallbladder diseases. *Br J Surg*. 11 2016;103(12):1704-1715. doi:10.1002/bjs.10287
16. Saunders DI, Murray D, Pichel AC, Varley S, Peden CJ, Network UEL. Variations in mortality after emergency laparotomy: the first report of the UK Emergency Laparotomy Network. *Br J Anaesth*. Sep 2012;109(3):368-75. doi:10.1093/bja/aes165
17. Faiz O, Warusavitarne J, Bottle A, et al. Nonelective excisional colorectal surgery in English National Health Service Trusts: a study of outcomes from Hospital Episode Statistics Data between 1996 and 2007. *J Am Coll Surg*. Apr 2010;210(4):390-401. doi:10.1016/j.jamcollsurg.2009.11.017
18. Clarke A, Murdoch H, Thomas MJ, Cook TM, Peden CJ. Mortality and postoperative care after emergency laparotomy. *Eur J Anaesthesiol*. Jan 2011;28(1):16-9. doi:10.1097/EJA.0b013e32833f5389
19. . AoUGS. *Pathway for the Management of Acute Gallstone Diseases ;*. 2015. Accessed 16th October 2020. http://www.augis.org/wp-content/uploads/2014/05/Acute-Gallstones-Pathway-Final-Sept-2015.pdf

20. Glasgow RE, Cho M, Hutter MM, Mulvihill SJ. The spectrum and cost of complicated gallstone disease in California. *Arch Surg*. Sep 2000;135(9):1021-5; discussion 1025-7. doi:10.1001/archsurg.135.9.1021
21. Wu XD, Tian X, Liu MM, Wu L, Zhao S, Zhao L. Meta-analysis comparing early versus delayed laparoscopic cholecystectomy for acute cholecystitis. *Br J Surg*. Oct 2015;102(11):1302-13. doi:10.1002/bjs.9886
22. Sanders G, Kingsnorth AN. Gallstones. *BMJ*. Aug 2007;335(7614):295-9. doi:10.1136/bmj.39267.452257.AD
23. CholeS Study Group WsMRC. Population-based cohort study of variation in the use of emergency cholecystectomy for benign gallbladder diseases. *Br J Surg*. Nov 2016;103(12):1716-1726. doi:10.1002/bjs.10288
24. NICE. *Gallstone Disease: Diagnosis and Management Clinical Guideline CG188*. 2014. Accessed 16th October 2020. https://www.nice.org.uk/guidance/cg188
25. NCEPOD. *Treat the Cause: A review of the quality of care provided to patients treated for acute pancreatitis*. 2016.
26. Zhou MW, Gu XD, Xiang JB, Chen ZY. Comparison of clinical safety and outcomes of early versus delayed laparoscopic cholecystectomy for acute cholecystitis: a meta-analysis. *ScientificWorldJournal*. 2014;2014:274516. doi:10.1155/2014/274516
27. Cao AM, Eslick GD, Cox MR. Early Cholecystectomy Is Superior to Delayed Cholecystectomy for Acute Cholecystitis: a Meta-analysis. *J Gastrointest Surg*. May 2015;19(5):848-57. doi:10.1007/s11605-015-2747-x
28. Gurusamy KS, Davidson C, Gluud C, Davidson BR. Early versus delayed laparoscopic cholecystectomy for people with acute cholecystitis. *Cochrane Database Syst Rev*. Jun 2013;(6):CD005440. doi:10.1002/14651858.CD005440.pub3
29. Murray AC, Markar S, Mackenzie H, et al. An observational study of the timing of surgery, use of laparoscopy and outcomes for acute cholecystitis in the USA and UK. *Surg Endosc*. 07 2018;32(7):3055-3063. doi:10.1007/s00464-017-6016-9
30. Polo M, Duclos A, Polazzi S, et al. Acute Cholecystitis-Optimal Timing for Early Cholecystectomy: a French Nationwide Study. *J Gastrointest Surg*. Nov 2015;19(11):2003-10. doi:10.1007/s11605-015-2909-x
31. Blohm M, Österberg J, Sandblom G, Lundell L, Hedberg M, Enochsson L. The Sooner, the Better? The Importance of Optimal Timing of Cholecystectomy in Acute Cholecystitis: Data from the National Swedish Registry for Gallstone Surgery, GallRiks. *J Gastrointest Surg*. 01 2017;21(1):33-40. doi:10.1007/s11605-016-3223-y
32. Berwick DM. Controlling variation in health care: a consultation from Walter Shewhart. *Med Care*. Dec 1991;29(12):1212-25. doi:10.1097/00005650-199112000-00004
33. Donabedian A. *T he Definition of Quality and Approaches to Its Assessment. Vol 1. Explorations in Quality Assessment and Monitoring*. Health Administration Press; 1980.
34. America IoMUCoQoHCi. Crossing the Quality Chasm: A New Health System for the 21st Century. 2001.
35. Board NQ. *Shared commitment to quality*. 2016. Accessed 16th October 2020. https://www.england.nhs.uk/publication/shared-commitment-to-quality-from-nqb/
36. Fitzsimons B, Cornwell J. What can we learn from patients' perspectives on the quality and safety of hospital care? *BMJ Qual Saf*. Sep 2018;27(9):671-672. doi:10.1136/bmjqs-2018-008106
37. Batalden PB, Davidoff F. What is "quality improvement" and how can it transform healthcare? *Qual Saf Health Care*. Feb 2007;16(1):2-3. doi:10.1136/qshc.2006.022046
38. Langley A, Denis JL. Beyond evidence: the micropolitics of improvement. *BMJ Qual Saf*. Apr 2011;20 Suppl 1:i43-6. doi:10.1136/bmjqs.2010.046482
39. Neuhauser D. The quality of medical care and the 14 points of Edward Deming. *Health Matrix*. 1988;6(2):7-10.

40. Best M, Neuhauser D. W Edwards Deming: father of quality management, patient and composer. *Qual Saf Health Care*. Aug 2005;14(4):310-2. doi:10.1136/qshc.2005.015289
41. Langley G, Moen R, Nolan K. *The improvement guide: a practical approach to enhancing organizational performance*. 2nd edition ed. Jossey-Bass; 2009.
42. Fund TK. Embedding a Culture of Quality Improvement. London: The Kings Fund; 2017.
43. Kinsman L, Rotter T, James E, Snow P, Willis J. What is a clinical pathway? Development of a definition to inform the debate. *BMC Med*. May 27 2010;8:31. doi:10.1186/1741-7015-8-31
44. Lawal AK, Rotter T, Kinsman L, et al. What is a clinical pathway? Refinement of an operational definition to identify clinical pathway studies for a Cochrane systematic review. *BMC Med*. Feb 23 2016;14:35. doi:10.1186/s12916-016-0580-z
45. Ljungqvist O, Young-Fadok T, Demartines N. The History of Enhanced Recovery After Surgery and the ERAS Society. *J Laparoendosc Adv Surg Tech A*. Sep 2017;27(9):860-862. doi:10.1089/lap.2017.0350
46. Rotter T, Kinsman L, James E, et al. Clinical pathways: effects on professional practice, patient outcomes, length of stay and hospital costs. *Cochrane Database Syst Rev*. Mar 17 2010;(3):CD006632. doi:10.1002/14651858.CD006632.pub2
47. Allen D, Gillen E, Rixson L. The Effectiveness of Integrated Care Pathways for Adults and Children in Health Care Settings: A Systematic Review. *JBI Libr Syst Rev*. 2009;7(3):80-129. doi:10.11124/01938924-200907030-00001
48. Resar R, Pronovost P, Haraden C, Simmonds T, Rainey T, Nolan T. Using a bundle approach to improve ventilator care processes and reduce ventilator-associated pneumonia. *Jt Comm J Qual Patient Saf*. May 2005;31(5):243-8. doi:10.1016/s1553-7250(05)31031-2
49. Weiss CH. Why do we fail to deliver evidence-based practice in critical care medicine? *Curr Opin Crit Care*. Oct 2017;23(5):400-405. doi:10.1097/MCC.0000000000000436
50. Schweizer ML, Chiang HY, Septimus E, et al. Association of a bundled intervention with surgical site infections among patients undergoing cardiac, hip, or knee surgery. *JAMA*. Jun 2 2015;313(21):2162-71. doi:10.1001/jama.2015.5387
51. Lavallee JF, Gray TA, Dumville J, Russell W, Cullum N. The effects of care bundles on patient outcomes: a systematic review and meta-analysis. *Implement Sci*. Nov 29 2017;12(1):142. doi:10.1186/s13012-017-0670-0
52. Ivers N, Jamtvedt G, Flottorp S, et al. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev*. Jun 2012;(6):CD000259. doi:10.1002/14651858.CD000259.pub3
53. Hysong SJ. Meta-analysis: audit and feedback features impact effectiveness on care quality. *Med Care*. Mar 2009;47(3):356-63. doi:10.1097/MLR.0b013e3181893f6b
54. Brown B, Gude WT, Blakeman T, et al. Clinical Performance Feedback Intervention Theory (CP-FIT): a new theory for designing, implementing, and evaluating feedback in health care based on a systematic review and meta-synthesis of qualitative research. *Implement Sci*. 04 2019;14(1):40. doi:10.1186/s13012-019-0883-5
55. Hysong SJ, Francis J, Petersen LA. Motivating and engaging frontline providers in measuring and improving team clinical performance. *BMJ Qual Saf*. 05 2019;28(5):405-411. doi:10.1136/bmjqs-2018-008856
56. Taylor MJ, McNicholas C, Nicolay C, Darzi A, Bell D, Reed JE. Systematic review of the application of the plan-do-study-act method to improve quality in healthcare. *BMJ Qual Saf*. Apr 2014;23(4):290-8. doi:10.1136/bmjqs-2013-001862
57. McNicholas C, Lennox L, Woodcock T, Bell D, Reed JE. Evolving quality improvement support strategies to improve Plan-Do-Study-Act cycle fidelity: a retrospective mixed-methods study. *BMJ Qual Saf*. 05 2019;28(5):356-365. doi:10.1136/bmjqs-2017-007605

58. Nadeem E, Olin SS, Hill LC, Hoagwood KE, Horwitz SM. Understanding the components of quality improvement collaboratives: a systematic literature review. *Milbank Q*. Jun 2013;91(2):354-94. doi:10.1111/milq.12016
59. Wells S, Tamir O, Gray J, Naidoo D, Bekhit M, Goldmann D. Are quality improvement collaboratives effective? A systematic review. *BMJ Qual Saf*. 03 2018;27(3):226-240. doi:10.1136/bmjqs-2017-006926
60. Dixon-Woods M, Bosk CL, Aveling EL, Goeschel CA, Pronovost PJ. Explaining Michigan: developing an ex post theory of a quality improvement program. *Milbank Q*. Jun 2011;89(2):167-205. doi:10.1111/j.1468-0009.2011.00625.x
61. Campbell DA, Englesbe MJ, Kubus JJ, et al. Accelerating the pace of surgical quality improvement: the power of hospital collaboration. *Arch Surg*. Oct 2010;145(10):985-91. doi:10.1001/archsurg.2010.220
62. Glasgow JM, Scott-Caziewell JR, Kaboli PJ. Guiding inpatient quality improvement: a systematic review of Lean and Six Sigma. *Jt Comm J Qual Patient Saf*. Dec 2010;36(12):533-40. doi:10.1016/s1553-7250(10)36081-8
63. D'Andreamatteo A, Ianni L, Lega F, Sargiacomo M. Lean in healthcare: A comprehensive review. *Health Policy*. Sep 2015;119(9):1197-209. doi:10.1016/j.healthpol.2015.02.002
64. Mason SE, Nicolay CR, Darzi A. The use of Lean and Six Sigma methodologies in surgery: a systematic review. *Surgeon*. Apr 2015;13(2):91-100. doi:10.1016/j.surge.2014.08.002
65. Kaplan HC, Provost LP, Froehle CM, Margolis PA. The Model for Understanding Success in Quality (MUSIQ): building a theory of context in healthcare quality improvement. *BMJ Qual Saf*. Jan 2012;21(1):13-20. doi:10.1136/bmjqs-2011-000010
66. Ogrinc G, Davies L, Goodman D, Batalden P, Davidoff F, Stevens D. SQUIRE 2.0 (Standards for QUality Improvement Reporting Excellence): revised publication guidelines from a detailed consensus process. *BMJ Qual Saf*. 12 2016;25(12):986-992. doi:10.1136/bmjqs-2015-004411
67. Oliver CM, Bassett MG, Poulton TE, et al. Organisational factors and mortality after an emergency laparotomy: multilevel analysis of 39 903 National Emergency Laparotomy Audit patients. *Br J Anaesth*. Dec 2018;121(6):1346-1356. doi:10.1016/j.bja.2018.07.040
68. Donabedian A. The quality of care. How can it be assessed? *JAMA*. 1988 Sep 23-30 1988;260(12):1743-8. doi:10.1001/jama.260.12.1743
69. Coles E, Anderson J, Maxwell M, et al. The influence of contextual factors on healthcare quality improvement initiatives: a realist review. *Syst Rev*. Apr 2020;9(1):94. doi:10.1186/s13643-020-01344-3
70. Kaplan HC, Brady PW, Dritz MC, et al. The influence of context on quality improvement success in health care: a systematic review of the literature. *Milbank Q*. Dec 2010;88(4):500-59. doi:10.1111/j.1468-0009.2010.00611.x
71. Stevens DP, Shojania KG. Tell me about the context, and more. *BMJ Qual Saf*. Jul 2011;20(7):557-9. doi:10.1136/bmjqs-2011-000206
72. Moore GF, Audrey S, Barker M, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*. Mar 2015;350:h1258.
73. Haynes AB, Weiser TG, Berry WR, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med*. Jan 2009;360(5):491-9. doi:10.1056/NEJMsa0810119
74. Pronovost P, Needham D, Berenholtz S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med*. Dec 2006;355(26):2725-32. doi:10.1056/NEJMoa061115
75. Benning A, Ghaleb M, Suokas A, et al. Large scale organisational intervention to improve patient safety in four UK hospitals: mixed method evaluation. *BMJ*. Feb 2011;342:d195.
76. Benning A, Dixon-Woods M, Nwulu U, et al. Multiple component patient safety intervention in English hospitals: controlled evaluation of second phase. *BMJ*. Feb 2011;342:d199.

77. Ahmed J, Khan S, Lim M, Chandrasekaran TV, MacFie J. Enhanced recovery after surgery protocols - compliance and variations in practice during routine colorectal surgery. *Colorectal Dis*. Sep 2012;14(9):1045-51. doi:10.1111/j.1463-1318.2011.02856.x
78. Schouten LM, Hulscher ME, van Everdingen JJ, Huijsman R, Grol RP. Evidence for the impact of quality improvement collaboratives: systematic review. *BMJ*. Jun 28 2008;336(7659):1491-4. doi:10.1136/bmj.39570.749884.BE
79. Auerbach AD, Landefeld CS, Shojania KG. The tension between needing to improve care and knowing how to do it. *N Engl J Med*. Aug 2007;357(6):608-13. doi:10.1056/NEJMsb070738
80. Grady D, Redberg RF, O'Malley PG. Quality Improvement for Quality Improvement Studies. *JAMA Intern Med*. 02 2018;178(2):187. doi:10.1001/jamainternmed.2017.6875
81. Dixon-Woods M, Martin GP. Does quality improvement improve quality? *Future Hosp J*. Oct 2016;3(3):191-194. doi:10.7861/futurehosp.3-3-191
82. Clarke-Jones J. Productive ward: the way forward. *Health Serv J*. Apr 2007;117(6052):suppl 18-9.
83. Shepherd S. Rolling out the productive ward. *Nurs Times*. 2009 Mar 10-16 2009;105(9):suppl 12-3.
84. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q*. 2004;82(4):581-629. doi:10.1111/j.0887-378X.2004.00325.x
85. Robert G, Morrow E, Maben J, Griffiths P, Callard L. The adoption, local implementation and assimilation into routine nursing practice of a national quality improvement programme: the Productive Ward in England. *J Clin Nurs*. Apr 2011;20(7-8):1196-207. doi:10.1111/j.1365-2702.2010.03480.x
86. Morrow E, Robert G, Maben J, Griffiths P. Implementing large-scale quality improvement: lessons from The Productive Ward: Releasing Time to Care. *Int J Health Care Qual Assur*. 2012;25(4):237-53. doi:10.1108/09526861211221464
87. Wright S, McSherry W. A systematic literature review of Releasing Time to Care: The Productive Ward. *J Clin Nurs*. May 2013;22(9-10):1361-71. doi:10.1111/jocn.12074
88. Improvement. TIfH. *Technical Report of the Health Foundation's Safer Patients Initiative Phase 2.* .
89. Benn J, Burnett S, Parand A, Pinto A, Iskander S, Vincent C. Perceptions of the impact of a large-scale collaborative improvement programme: experience in the UK Safer Patients Initiative. *J Eval Clin Pract*. Jun 2009;15(3):524-40. doi:10.1111/j.1365-2753.2009.01145.x
90. Foundation TH. *Learning Report: The Safer Patients Initiative*. 2011.
91. Dixon-Woods M, Leslie M, Tarrant C, Bion J. Explaining Matching Michigan: an ethnographic study of a patient safety program. *Implement Sci*. Jun 2013;8:70. doi:10.1186/1748-5908-8-70
92. Bion J, Richardson A, Hibbert P, et al. 'Matching Michigan': a 2-year stepped interventional programme to minimise central venous catheter-blood stream infections in intensive care units in England. *BMJ Qual Saf*. Feb 2013;22(2):110-23. doi:10.1136/bmjqs-2012-001325
93. Power M, Tyrrell PJ, Rudd AG, et al. Did a quality improvement collaborative make stroke care better? A cluster randomized trial. *Implement Sci*. Apr 2014;9(1):40. doi:10.1186/1748-5908-9-40
94. Improvement IfH. *The breakthrough series: IHI's collaborative model for achieving breakthrough improvement.* 2003. *Innovation Series White Paper,* .
95. Carter P, Ozieranski P, McNicol S, Power M, Dixon-Woods M. How collaborative are quality improvement collaboratives: a qualitative study in stroke care. *Implement Sci*. Mar 2014;9(1):32. doi:10.1186/1748-5908-9-32
96. Craig P, Dieppe P, Macintyre S, et al. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*. Sep 2008;337:a1655.

97. Brown C, Hofer T, Johal A, et al. An epistemology of patient safety research: a framework for study design and interpretation. Part 2. Study design. *Qual Saf Health Care*. Jun 2008;17(3):163-9. doi:10.1136/qshc.2007.023648
98. Øvretveit J, Gustafson D. Evaluation of quality improvement programmes. *Qual Saf Health Care*. Sep 2002;11(3):270-5.
99. Poots AJ, Reed JE, Woodcock T, Bell D, Goldmann D. How to attribute causality in quality improvement: lessons from epidemiology. *BMJ Qual Saf*. 11 2017;26(11):933-937. doi:10.1136/bmjqs-2017-006756
100. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ*. Mar 2014;348:g1687.
101. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. Apr 2008;336(7650):924-6. doi:10.1136/bmj.39489.470347.AD
102. Quiney N, Huddart S, Peden C, Dickinson M. Use of a care bundle to reduce mortality following emergency laparotomy. *Br J Hosp Med (Lond)*. Jun 2015;76(6):358-62. doi:10.12968/hmed.2015.76.6.358
103. Dixon-Woods M, McNicol S, Martin G. Ten challenges in improving quality in healthcare: lessons from the Health Foundation's programme evaluations and relevant literature. *BMJ Qual Saf*. Oct 2012;21(10):876-84. doi:10.1136/bmjqs-2011-000760
104. Gotlib Conn L, McKenzie M, Pearsall EA, McLeod RS. Successful implementation of an enhanced recovery after surgery programme for elective colorectal surgery: a process evaluation of champions' experiences. *Implement Sci*. Jul 2015;10:99. doi:10.1186/s13012-015-0289-y
105. Mills PD, Weeks WB. Characteristics of successful quality improvement teams: lessons from five collaborative projects in the VHA. *Jt Comm J Qual Saf*. Mar 2004;30(3):152-62.
106. Huddart S, Peden CJ, Swart M, et al. Use of a pathway quality improvement care bundle to reduce mortality after emergency laparotomy. *Br J Surg*. Jan 2015;102(1):57-66. doi:10.1002/bjs.9658
107. Pronovost PJ, Watson SR, Goeschel CA, Hyzy RC, Berenholtz SM. Sustaining Reductions in Central Line-Associated Bloodstream Infections in Michigan Intensive Care Units: A 10-Year Analysis. *Am J Med Qual*. May 2016;31(3):197-202. doi:10.1177/1062860614568647
108. Gabbay J, le May A, Connell C, J. K. Skilled for improvement? : The Health Foundation; 2014.
109. Alexander JA, Hearld LR. The science of quality improvement implementation: developing capacity to make a difference. *Med Care*. Dec 2011;49 Suppl:S6-20. doi:10.1097/MLR.0b013e3181e1709c
110. Perla RJ, Provost LP, Murray SK. The run chart: a simple analytical tool for learning from variation in healthcare processes. *BMJ Qual Saf*. Jan 2011;20(1):46-51. doi:10.1136/bmjqs.2009.037895
111. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. Jan 13 1996;312(7023):71-2.
112. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci*. Aug 2009;4:50. doi:10.1186/1748-5908-4-50
113. Murray E, Treweek S, Pope C, et al. Normalisation process theory: a framework for developing, evaluating and implementing complex interventions. *BMC Med*. Oct 2010;8:63. doi:10.1186/1741-7015-8-63
114. Jones EL, Lees N, Martin G, Dixon-Woods M. How Well Is Quality Improvement Described in the Perioperative Care Literature? A Systematic Review. *Jt Comm J Qual Patient Saf*. May 2016;42(5):196-206.

115. Grant A, Treweek S, Dreischulte T, Foy R, Guthrie B. Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. *Trials*. Jan 2013;14:15. doi:10.1186/1745-6215-14-15
116. Boynton PM, Greenhalgh T. Selecting, designing, and developing your questionnaire. *BMJ*. May 2004;328(7451):1312-5. doi:10.1136/bmj.328.7451.1312
117. McColl E JA, Thomas L,. *Design and use*

*of questionnaires: a review of best practice applicable to surveys of health*

*service staff and patients.* . 2001.
118. Peden CJ, Stephens T, Martin G, et al. A national quality improvement programme to improve survival after emergency abdominal surgery: the EPOCH stepped-wedge cluster RCT. 2019.
119. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs*. Apr 2008;62(1):107-15. doi:10.1111/j.1365-2648.2007.04569.x
120. Blencowe NS, Brown JM, Cook JA, et al. Interventions in randomised controlled trials in surgery: issues to consider during trial design. *Trials*. Sep 2015;16:392. doi:10.1186/s13063-015-0918-4
121. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci*. Nov 2007;2:40. doi:10.1186/1748-5908-2-40
122. Spurgeon P, Mazelan PM, Barwell F. Medical engagement: a crucial underpinning to organizational performance. *Health Serv Manage Res*. Aug 2011;24(3):114-20. doi:10.1258/hsmr.2011.011006
123. Pannick S, Athanasiou T, Long SJ, Beveridge I, Sevdalis N. Translating staff experience into organisational improvement: the HEADS-UP stepped wedge, cluster controlled, non-randomised trial. *BMJ Open*. Jul 18 2017;7(7):e014333. doi:10.1136/bmjopen-2016-014333
124. Burke RE, Shojania KG. Rigorous evaluations of evolving interventions: can we have our cake and eat it too? *BMJ Qual Saf*. Apr 2018;27(4):254-257. doi:10.1136/bmjqs-2017-007554
125. Taylor SL, Dy S, Foy R, et al. What context features might be important determinants of the effectiveness of patient safety practice interventions? *BMJ Qual Saf*. Jul 2011;20(7):611-7. doi:10.1136/bmjqs.2010.049379
126. Anhøj J, Olesen AV. Run charts revisited: a simulation study of run chart rules for detection of non-random variation in health care processes. *PLoS One*. 2014;9(11):e113825. doi:10.1371/journal.pone.0113825
127. Provost L, Murray S. *The Health Care Data Guide: Learning from Data for Improvement*. 1st edition ed. Jossey-Bass;; 2011.
128. Eveleigh MO, Howes TE, Peden CJ, Cook TM. Estimated costs before, during and after the introduction of the emergency laparotomy pathway quality improvement care (ELPQuIC) bundle. *Anaesthesia*. 11 2016;71(11):1291-1295. doi:10.1111/anae.13623
129. Braithwaite J, Churruca K, Long JC, Ellis LA, Herkes J. When complexity science meets implementation science: a theoretical and empirical analysis of systems change. *BMC Med*. 04 2018;16(1):63. doi:10.1186/s12916-018-1057-z
130. Anhøj J. Diagnostic value of run chart analysis: using likelihood ratios to compare run chart rules on simulated data series. *PLoS One*. 2015;10(3):e0121349. doi:10.1371/journal.pone.0121349
131. Resar RK. Making noncatastrophic health care processes reliable: Learning to walk before running in creating high-reliability organizations. *Health Serv Res*. Aug 2006;41(4 Pt 2):1677-89. doi:10.1111/j.1475-6773.2006.00571.x
132. Team NP. *The Second*

*Patient Report of the National Emergency Laparotomy Audit ( NELA ).* . 2016. .
133. Reed JE, Howe C, Doyle C, Bell D. Simple rules for evidence translation in complex systems: A qualitative study. *BMC Med*. 06 2018;16(1):92. doi:10.1186/s12916-018-1076-9

134. Johnson MJ, May CR. Promoting professional behaviour change in healthcare: what interventions work, and why? A theory-led overview of systematic reviews. *BMJ Open*. Sep 2015;5(9):e008592. doi:10.1136/bmjopen-2015-008592
135. Stephens TJ, Peden CJ, Pearse RM, et al. Improving care at scale: process evaluation of a multi-component quality improvement intervention to reduce mortality after emergency abdominal surgery (EPOCH trial). *Implement Sci*. 11 2018;13(1):142. doi:10.1186/s13012-018-0823-9
136. Pearse RM, Harrison DA, MacDonald N, et al. Effect of a perioperative, cardiac output-guided hemodynamic therapy algorithm on outcomes following major gastrointestinal surgery: a randomized clinical trial and systematic review. *JAMA*. Jun 2014;311(21):2181-90. doi:10.1001/jama.2014.5305
137. Clarke CM, Cheng T, Reims KG, et al. Implementation of HIV treatment as prevention strategy in 17 Canadian sites: immediate and sustained outcomes from a 35-month Quality Improvement Collaborative. *BMJ Qual Saf*. May 2016;25(5):345-54. doi:10.1136/bmjqs-2015-004269
138. Huddart S, Peden CJ, Swart M, et al. Use of a pathway quality improvement care bundle to reduce mortality after emergency laparotomy. *Br J Surg*. Jan 2015;102(1):57-66. doi:10.1002/bjs.9658
139. Benneyan JC, Lloyd RC, Plsek PE. Statistical process control as a tool for research and healthcare improvement. *Qual Saf Health Care*. Dec 2003;12(6):458-64. doi:10.1136/qhc.12.6.458
140. Mohammed MA. Using statistical process control to improve the quality of health care. *Qual Saf Health Care*. Aug 2004;13(4):243-5. doi:10.1136/qhc.13.4.243
141. Enochsson L, Thulin A, Osterberg J, Sandblom G, Persson G. The Swedish Registry of Gallstone Surgery and Endoscopic Retrograde Cholangiopancreatography (GallRiks): A nationwide registry for quality assurance of gallstone surgery. *JAMA Surg*. May 2013;148(5):471-8. doi:10.1001/jamasurg.2013.1221
142. Harboe KM, Bardram L. Nationwide quality improvement of cholecystectomy: results from a national database. *Int J Qual Health Care*. Oct 2011;23(5):565-73. doi:10.1093/intqhc/mzr041
143. Lewis J. *Qualitative research practice: a*

*guide for social science students and researchers*. Sage; 2003.
144. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol*. Sep 2013;13:117. doi:10.1186/1471-2288-13-117
145. May CR, Cummings A, Girling M, et al. Using Normalization Process Theory in feasibility studies and process evaluations of complex healthcare interventions: a systematic review. *Implement Sci*. 06 2018;13(1):80. doi:10.1186/s13012-018-0758-1
146. Hulscher ME, Schouten LM, Grol RP, Buchan H. Determinants of success of quality improvement collaboratives: what does the literature show? *BMJ Qual Saf*. Jan 2013;22(1):19-31. doi:10.1136/bmjqs-2011-000651
147. Khodyakov D, Ridgely MS, Huang C, DeBartolo KO, Sorbero ME, Schneider EC. Project JOINTS: what factors affect bundle adoption in a voluntary quality improvement campaign? *BMJ Qual Saf*. Jan 2015;24(1):38-47. doi:10.1136/bmjqs-2014-003169
148. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Med Res Methodol*. Jul 10 2008;8:45. doi:10.1186/1471-2288-8-45
149. Woodcock T, Liberati EG, Dixon-Woods M. A mixed-methods study of challenges experienced by clinical teams in measuring improvement. *BMJ Qual Saf*. Feb 2021;30(2):106-115. doi:10.1136/bmjqs-2018-009048
150. Brehaut JC, Colquhoun HL, Eva KW, et al. Practice Feedback Interventions: 15 Suggestions for Optimizing Effectiveness. *Ann Intern Med*. Mar 2016;164(6):435-41. doi:10.7326/M15-2248
151. Donabedian A. Evaluating the quality of medical care. 1966. *Milbank Q*. 2005;83(4):691-729. doi:10.1111/j.1468-0009.2005.00397.x

152. Ubeda SRG. An Approach to Identify "Right" Problems for Initial Quality Improvement Projects. *Global Journal on Quality and Safety in Healthcare* 2020;3(1)doi:https://doi.org/10.4103/JQSH.JQSH_16_19
153. Soong C, Cho HJ, Shojania KG. Choosing quality problems wisely: identifying improvements worth developing and sustaining. *BMJ Qual Saf*. 10 2020;29(10):1-2. doi:10.1136/bmjqs-2020-011054
154. Power M, Fogarty M, Madsen J, et al. Learning from the design and development of the NHS Safety Thermometer. *Int J Qual Health Care*. Jun 2014;26(3):287-97. doi:10.1093/intqhc/mzu043
155. Marshall M, de Silva D, Cruickshank L, Shand J, Wei L, Anderson J. What we know about designing an effective improvement intervention (but too often fail to put into practice). *BMJ Qual Saf*. 07 2017;26(7):578-582. doi:10.1136/bmjqs-2016-006143
156. Team NP. *Fourth*

*Patient Report of the National Emergency Laparotomy Audit*. 2018.
157. Aggarwal G, Peden CJ, Mohammed MA, et al. Evaluation of the Collaborative Use of an Evidence-Based Care Bundle in Emergency Laparotomy. *JAMA Surg*. 05 2019;154(5):e190145. doi:10.1001/jamasurg.2019.0145
158. Vester-Andersen M, Waldau T, Wetterslev J, et al. Randomized multicentre feasibility trial of intermediate care versus standard ward care after emergency abdominal surgery (InCare trial). *Br J Surg*. May 2015;102(6):619-29. doi:10.1002/bjs.9749
159. Improvement IoH. *The Breakthrough Series: IHIs collaborative model for acheiving breakthrough improvement*. 2003. *Innovation Series White Paper*.
160. Philp MM, Pitt HA. Emergency general surgery: can we do better? *Lancet*. 06 2019;393(10187):2178-2180. doi:10.1016/S0140-6736(18)32982-9
161. Cornelius V, Johnston CL. The Enhanced Peri-Operative Care for High-risk patients trial: an independent discussion and commentary. *Br J Anaesth*. Sep 2019;123(3):261-266. doi:10.1016/j.bja.2019.06.003
162. Murray D, Pearse RM, Quiney N. Towards a coordinated national strategy to improve survival after emergency laparotomy. *Br J Anaesth*. Oct 2019;123(4):399-401. doi:10.1016/j.bja.2019.06.023
163. Shojania KG. Conventional evaluations of improvement interventions: more trials or just more tribulations? *BMJ Qual Saf*. Nov 2013;22(11):881-4. doi:10.1136/bmjqs-2013-002377
164. Kilbourne AM, Almirall D, Eisenberg D, et al. Protocol: Adaptive Implementation of Effective Programs Trial (ADEPT): cluster randomized SMART trial comparing a standard versus enhanced implementation strategy to improve outcomes of a mood disorders program. *Implement Sci*. Sep 2014;9:132. doi:10.1186/s13012-014-0132-x
165. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials*. Feb 2007;28(2):182-91. doi:10.1016/j.cct.2006.05.007
166. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ*. Feb 7 2011;342:d40. doi:10.1136/bmj.d40
167. Kahan BC, Jairath V, Dore CJ, Morris TP. The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies. *Trials*. Apr 23 2014;15:139. doi:10.1186/1745-6215-15-139
168. Kahan BC, Rushton H, Morris TP, Daniel RM. A comparison of methods to adjust for continuous covariates in the analysis of randomised trials. *BMC Med Res Methodol*. Apr 11 2016;16:42. doi:10.1186/s12874-016-0141-3
169. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. *Stat Med*. Apr 15 2005;24(7):993-1007. doi:10.1002/sim.1981
170. Morgan KE, Forbes AB, Keogh RH, Jairath V, Kahan BC. Choosing appropriate analysis methods for cluster randomised cross-over trials with a binary outcome. *Stat Med*. Jan 30 2017;36(2):318-333. doi:10.1002/sim.7137

171. Hemming K, Taljaard M, Forbes A. Analysis of cluster randomised stepped wedge trials with repeated cross-sectional samples. *Trials*. Mar 4 2017;18(1):101. doi:10.1186/s13063-017-1833-7
172. Stephens TJ, Peden CJ, Haines R, et al. Hospital-level evaluation of the effect of a national quality improvement programme: time-series analysis of registry data. *BMJ Qual Saf*. Aug 2020;29(8):623-635. doi:10.1136/bmjqs-2019-009537
173. Møller MH, Adamsen S, Thomsen RW, Møller AM, group PUPPt. Multicentre trial of a perioperative protocol to reduce mortality in patients with peptic ulcer perforation. *Br J Surg*. Jun 2011;98(6):802-10. doi:10.1002/bjs.7429
174. Tengberg LT, Bay-Nielsen M, Bisgaard T, et al. Multidisciplinary perioperative protocol in patients undergoing acute high-risk abdominal surgery. *Br J Surg*. Mar 2017;104(4):463-471. doi:10.1002/bjs.10427
175. Presseau J, Mackintosh J, Hawthorne G, et al. Cluster randomised controlled trial of a theory-based multiple behaviour change intervention aimed at healthcare professionals to improve their management of type 2 diabetes in primary care. *Implement Sci*. 05 2018;13(1):65. doi:10.1186/s13012-018-0754-5
176. Williams L, Daggett V, Slaven JE, et al. A cluster-randomised quality improvement study to improve two inpatient stroke quality indicators. *BMJ Qual Saf*. Apr 2016;25(4):257-64. doi:10.1136/bmjqs-2015-004188
177. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci*. Apr 2011;6:42. doi:10.1186/1748-5908-6-42

# APPENDIX

1) The EPOCH trial: main trial paper
2) EPOCH Process Evaluation Questionnaire
3) CholeQuIC Application form and selection criteria
4) CholeQuIC Focus Group topic guide
5) CholeQuIC structured site visit field note

# The EPOCH trial: A stepped-wedge cluster randomised trial of a Quality Improvement programme, designed to reduce mortality from emergency abdominal surgery in the UK NHS

## *3.1 Introduction*

Prior to the EPOCH trial, emergency abdominal surgery was associated with worrying poor post-operative outcomes. Around 30,000 patients underwent this type of surgery each year in the UK National Health Service with 30-day mortality rates in excess of 10%, with wide variation in standards of care between hospitals. Several groups had studied the effect of Quality Improvement initiatives to implement individual interventions or 'care bundles' of several treatments, and so improved care for these patients. Overall, the findings of these small studies suggested a survival benefit, but most utilised uncontrolled cohort designs associated with a high risk of bias. The feasibility and benefit of a national Quality Improvement programme to implement a more extensive acute care pathway for this patient group remained uncertain. We conducted a stepped-wedge cluster randomised trial, with an embedded process evaluation, to evaluate the effect of implementing this pathway on survival following emergency abdominal surgery in NHS hospitals

## *3.2 Methods*

### *3.2.1 Study design and participants*

EPOCH was a multi-centre, stepped-wedge cluster randomised trial of a Quality Improvement intervention to promote the implementation of a perioperative care pathway for patients undergoing emergency abdominal surgery. The trial protocol was published prospectively by the Lancet (Protocol 13PRT/7655) and on the trial website

(www.epochtrial.org/protocol). The trial was prospectively registered at isrctn.com on 27th February 2014 but a registration number was not issued until 7th March 2014 (ISRCTN80682973).

NHS hospitals delivering an emergency general surgical service were eligible for inclusion provided they undertook a significant volume of emergency abdominal surgery cases and contributed data to the National Emergency Laparotomy Audit (NELA). Hospitals were required to nominate specialty leads from surgery, anaesthesia and critical care, and to secure support from their NHS Trust Board or equivalent. Hospitals which were already implementing a care pathway to improve treatment for this patient group were excluded. Patients were eligible for inclusion in the data analysis if they were 40 years or older, and undergoing emergency open abdominal surgery in a participating hospital during the 85-week trial period from 3rd March 2014 to 19th October 2015. Patients were excluded from the analysis if they were undergoing a simple appendectomy, surgery related to organ transplant, gynaecological surgery, laparotomy for traumatic injury, treatment of complications of recent elective surgery or if they had previously been included in the EPOCH trial.

*3.2.2 Data collection*

Trial data were collected through the NELA database (www.nela.org.uk), and then linked using unique patient identifiers to Hospital Episode Statistics and the Office for National Statistics in England and Wales, and the Information Services Division of NHS Scotland, to provide data describing mortality and hospital re-admissions. The trial was approved by the

East Midlands (Nottingham 1) Research Ethics Committee (Ref: 13/EM/0415). Data were analysed without individual patient consent in accordance with section 251 of the National Health Services Act 2006.

*3.3.3 Randomisation and masking*

We planned to include 15 geographical clusters of five to seven hospitals. The Quality Improvement intervention lasted 85 weeks with one geographical cluster commencing the intervention each five-week step from the 2nd to the 16th time-period. Clusters were randomly assigned to one of 15 start dates for the intervention by an independent statistician using a computer-generated random allocation sequence. Because each geographical area started in the usual care group, and ended in the Quality Improvement group, there were 17 time periods in total. Local investigators in each geographical area were notified 12 weeks in advance of activation of the Quality Improvement programme at their hospital. Because they were engaged in delivery of the intervention, it was not possible to mask hospital staff. Patients were masked to study group allocation. The organisation of hospitals into geographical clusters minimised any contamination between sites due to natural workforce movements between hospitals.

*3.3.4 Trial intervention*

The EPOCH care pathway and Quality Improvement methodology are described in Chapter 2.

*3.3.5 Outcome measures*

The primary outcome measure was all-cause mortality within 90 days following surgery. Secondary outcomes were all-cause mortality within 180 days following surgery, duration of hospital stay after surgery and hospital re-admission within 180 days of surgery. We selected ten predefined process measures (key components of the care pathway) for inclusion in the main report: 1) consultant-led decision to operate, 2) consultant review of patient before surgery, 3) pre-operative documentation of risk, 4) time from decision to operate to entry into operating theatre, 5) patient entered operating theatre within time-frame specified by their urgency (less than 2 hours, 2-6 hours, 6-18 hours, or more than 18 hours), 6) consultant surgeon present in operating theatre, 7) consultant anaesthetist present in operating theatre, 8) cardiac output guided fluid therapy used during surgery, 9) serum lactate measured at end of surgery and 10) critical care admission immediately after surgery.

*3.3.6 Statistical analysis*

A stepped-wedge design was chosen to improve statistical power by facilitating within-cluster comparison. Sample size calculations were based on the Hussey & Hughes approach,[165] for an analysis with fixed time effects and random cluster effects, modified to exclude data collected during the five-week period in which the intervention commenced in individual clusters. Using Hospital Episodes Statistics data, we estimated that 27,540 eligible

patients would be registered across 90 NHS hospitals over 85 weeks, with a 90-day mortality rate of 25% in the usual care group, and a between-hospital coefficient of variation of 0.15. Assuming a constant case load (18 patients per five weeks per hospital), independent hospital effects and a 5% significance level, the trial would have 92% power to detect a reduction in 90-day mortality from 25% to 22%. If the assumption of independent hospital effects was not met, and the 15 geographical clusters functioned effectively as 15 large hospitals, power would be reduced to 83%.

All analyses were conducted according to intention-to-treat principles. All eligible patients with available outcome data were included in the analysis, and analysed according to the randomisation schedule.[166] Patients who presented during the five-week time period immediately after Quality Improvement activation were excluded from the analysis. Hospitals that initially agreed to participate but subsequently withdrew prior to the trial start date were excluded, however hospitals which withdrew after the trial start date, or did not implement the intervention, were included in the analysis. Hospitals which merged with other hospitals during the trial period were included in the analysis up to the point of the merger.

We were unable to procure data describing survival status after hospital discharge for patients in Wales. We therefore changed our primary analysis from binary to a time-to-event approach allowing inclusion of mortality events censored at hospital discharge. This affected 909 patients in Wales, 179 (20%) of whom died in hospital, and 730 (80%) who were censored at hospital discharge. All analyses included time period as a fixed effect using

indicator variables, and adjusted for age, gender, and indication for surgery using fixed factors.[167] Age was included as a continuous covariate, assuming a linear association with outcome.[168] Missing baseline data for indication for surgery were handled using a missing indicator approach.[169] All-cause mortality within 90 days of surgery was analysed using a mixed-effects parametric survival model with a Weibull survival distribution. The model included random-intercepts for geographical area, hospital and hospital-period in the Quality Improvement programme. This allowed additional correlation between patients in the same hospital and the same period, compared to patients in other periods, as is recommended.[170, 171] All-cause mortality within 180 days was analysed using the same approach. Duration of hospital stay was analysed using competing risk time-to-event models, with mortality before the outcome event acting as the competing risk, and robust standard errors to account for clustering by geographical area. The hazard ratio from this analysis measures the relative probability of hospital discharge between treatment arms, with HR<1 indicating a lower probability of discharge in the QI group (and therefore longer hospital stay). Hospital readmission within 180 days was analysed using the same approach (with a HR <1 indicating a lower probability of re-admission).

We performed two secondary analyses for the primary outcome. The first evaluated the effect of the intervention over time. This analysis included patients who presented to hospital during the five-week period immediately after implementation of the intervention. We analysed patients according to the following four groups; (a) no Quality Improvement implemented (usual care group); (b) Quality Improvement implemented less than five weeks; (c) Quality Improvement implemented between five and ten weeks; and (d) Quality

Improvement implemented more than ten weeks. Our second analysis evaluated the intervention in other patient populations which may have been affected by the intervention. This included patients who either underwent laparoscopic surgery or were aged 18-40 years, and who met all other eligibility criteria. Due to the small number of patients in this group, results are summarised descriptively rather than undertaking a formal statistical analysis.

## *3.4 Results*

Fifteen geographic areas underwent randomisation including 97 NHS hospitals. Four hospitals withdrew before the start of the trial, leaving 93 participating. Between 3rd March 2014 and 19th October 2015, 15,873 eligible patients underwent surgery in participating hospitals with data recorded in the NELA database (usual care 8490 patients, QI 7383 patients; Figure 3.1, Table 3.1). Baseline characteristics were similar between groups (Table 3.2).

### *3.4.1 Process measures*

91/93 (98%) hospitals were represented at the initial Quality Improvement meeting for the relevant geographical cluster and 53/93 (57%) were represented at the follow-up Quality Improvement meeting. This representation included a named hospital QI lead for 89/93 (96%) hospitals at the first meeting and 47/93 (51%) hospitals at the second. Most meetings (n=13/15) occurred within two weeks of the activation date. Patient-level process measures are described in Table 3.3. In accordance with our analysis plan; we did not test these for statistical significance.

### *3.4.2 Clinical outcomes*

Complete primary outcome data were available for more than 99% of patients (Figure 3.1). The primary outcome of 90-day mortality occurred in 1393 usual care group patients (16%) compared with 1210 QI group patients (16%) (Hazard ratio, QI vs usual care: 1.11 [0.96-1.28]) (Figure 3.2 and Table 3.4). Results were similar for mortality 180-day mortality (HR 1.12 [0.98-1.28]) (Figure 3.3). Patients in the Quality Improvement group had a lower probability of hospital discharge (Hazard ratio for hospital discharge 0.90 [0.83-0.97]), leading to a marginally longer hospital stay (days in hospital, usual care: 8 [13-23] days vs.

QI: 8 [13-24] days), although this difference was not clinically meaningful (Figure 3.4 and table 3.6). There was no difference between groups in hospital re-admission within 180 days (usual care 1618 (20%) vs. QI 1242 (18%); Hazard ratio for re-admission 0.87 [0.73-1.04]) (Figure 3.5 and table 3.6). In a secondary analysis, we found no evidence that the QI strategy became more effective the longer it had been adopted (Table 3.7). To assess the impact of missing mortality data following hospital discharge from patients in Wales, we assessed the number of mortality events which occurred after hospital discharge but before 90 days in English and Scottish hospitals. Only 5% (631/13,034) of patients died between hospital discharge and 90 days, suggesting few outcome events in Wales were missed. Analysis of the effect of the intervention over time is presented in table 3.6 and of the inclusion of younger patients and those undergoing laparoscopic surgery in Table 3.7.

**Figure 3.1. Inclusion of hospitals and patients in the trial.**

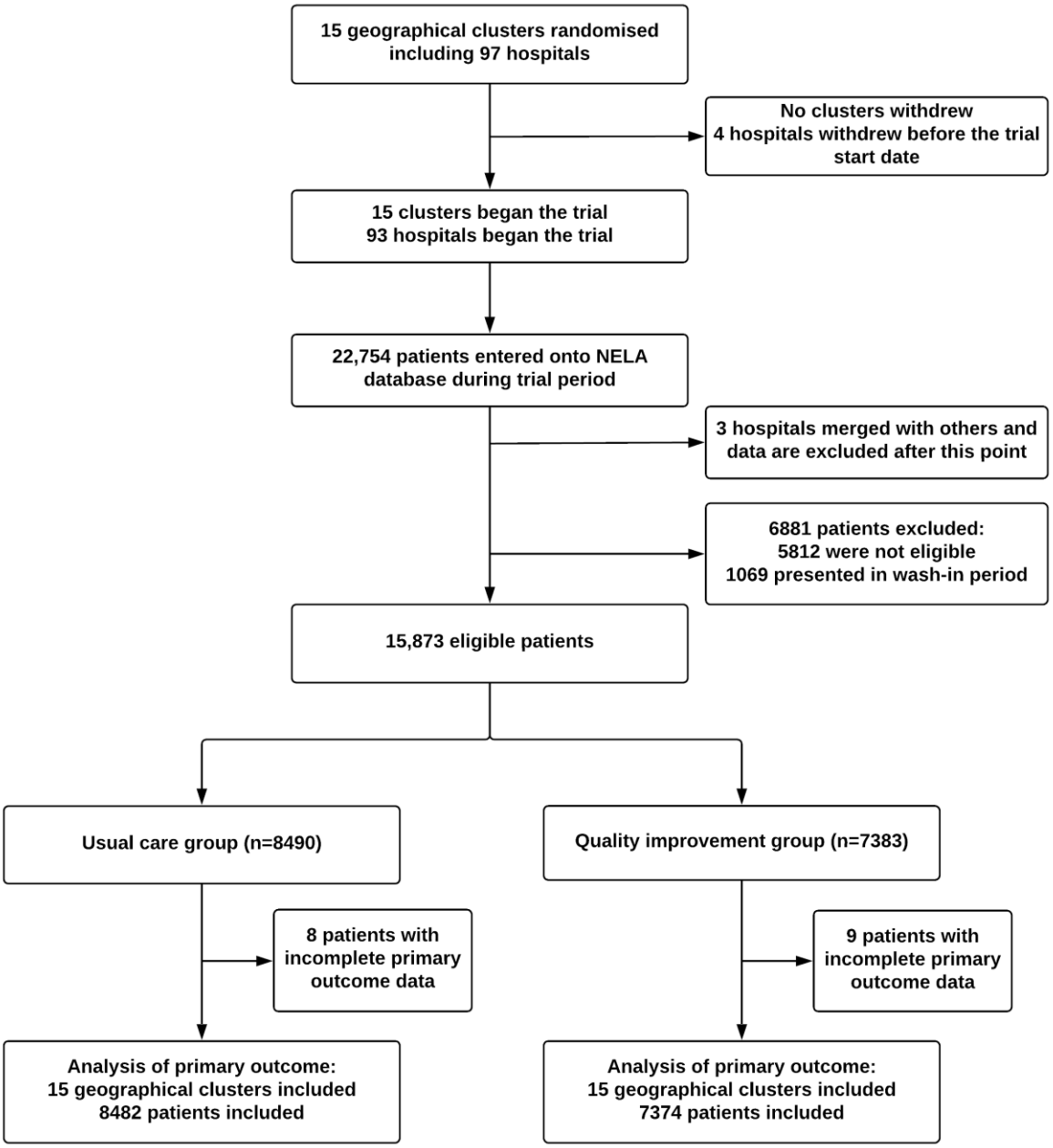


**Table 3.1. Number of patients recruited in each cluster in each period*[a]**

| | **Period** | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Geographical area (cluster)** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **Total** |
| **1** | 53 | - | 39 | 54 | 41 | 46 | 39 | 36 | 44 | 35 | 41 | 28 | 37 | 43 | 39 | 31 | 37 | *643* |
| **2** | 62 | 52 | - | 57 | 32 | 45 | 52 | 54 | 34 | 58 | 43 | 37 | 59 | 58 | 50 | 56 | 39 | *788* |
| **3** | 89 | 92 | 100 | - | 93 | 98 | 95 | 88 | 91 | 105 | 84 | 83 | 107 | 79 | 69 | 62 | 60 | *1395* |
| **4** | 60 | 49 | 52 | 67 | - | 55 | 68 | 55 | 75 | 67 | 78 | 64 | 82 | 72 | 74 | 71 | 55 | *1044* |
| **5** | 23 | 26 | 31 | 30 | 34 | - | 24 | 27 | 20 | 46 | 30 | 33 | 34 | 48 | 25 | 36 | 34 | *501* |
| **6** | 59 | 59 | 64 | 52 | 61 | 69 | - | 53 | 57 | 52 | 65 | 52 | 33 | 59 | 58 | 54 | 62 | *909* |
| **7** | 74 | 62 | 76 | 79 | 55 | 64 | 79 | - | 68 | 79 | 63 | 84 | 73 | 71 | 78 | 68 | 83 | *1156* |
| **8** | 64 | 56 | 66 | 70 | 50 | 63 | 59 | 60 | - | 73 | 44 | 58 | 55 | 65 | 55 | 55 | 54 | *947* |
| **9** | 104 | 91 | 95 | 88 | 69 | 72 | 98 | 77 | 86 | - | 86 | 76 | 83 | 72 | 70 | 58 | 69 | *1294* |
| **10** | 65 | 64 | 71 | 90 | 68 | 69 | 94 | 83 | 79 | 80 | - | 68 | 85 | 76 | 75 | 61 | 80 | *1208* |
| **11** | 82 | 79 | 91 | 111 | 90 | 117 | 77 | 94 | 91 | 102 | 117 | - | 96 | 117 | 103 | 91 | 85 | *1543* |
| **12** | 85 | 79 | 82 | 80 | 64 | 69 | 58 | 75 | 66 | 60 | 70 | 73 | - | 82 | 90 | 89 | 89 | *1211* |
| **13** | 55 | 60 | 59 | 61 | 57 | 61 | 76 | 54 | 70 | 56 | 56 | 69 | 52 | - | 60 | 67 | 55 | *968* |
| **14** | 55 | 60 | 54 | 57 | 56 | 65 | 50 | 43 | 45 | 69 | 62 | 68 | 66 | 55 | - | 43 | 46 | *894* |
| **15** | 95 | 95 | 74 | 98 | 79 | 69 | 72 | 68 | 65 | 87 | 91 | 85 | 86 | 118 | 101 | - | 89 | *1372* |
| **Total** | *1025* | *924* | *954* | *994* | *849* | *962* | *941* | *867* | *891* | *969* | *930* | *878* | *948* | *1015* | *947* | *842* | *937* | ***15873*** |

*Shaded cells denote periods after intervention implementation. Cells with '-' denote the five-week period immediately after intervention implementation when participants were excluded. [a] Geographical areas are: (1) North East London; (2) South London; (3) North East England; (4) Thames Valley; (5) North West London; (6) South Wales; (7) North Lancaster/Cumbria; (8)Manchester/Merseyside/Yorkshire; (9) East Anglia; (10) Peninsula; (11) East Midlands; (12) Kent/Surry/Sussux; (13) Wessex; (14) Scotland; (15) West Midlands.

**Table 3.2. Baseline patient characteristics. Data presented as n (%) unless otherwise indicated. ASA: American Society of Anesthesiologists physical status score; P-POSSUM: Portsmouth Physiological and Operative Severity Score for the enumeration of Mortality and morbidity score.**

| | **Summary measure** | |
|---|---|---|
| | **Usual care** | **Quality improvement** |
| ***Baseline characteristics*** | | |
| Female | 4550 (54) | 3938 (53) |
| Age – mean (SD) | 68 (13) | 68 (13) |
| Indication for surgery | | |
| Peritonitis | 352 (4) | 251 (3) |
| Perforation | 765 (9) | 693 (9) |
| Intestinal obstruction | 3840 (45) | 3379 (46) |
| Haemorrhage | 213 (3) | 149 (2) |
| Ischaemia | 366 (4) | 332 (5) |
| Abdominal infection | 296 (3) | 239 (3) |
| Other | 523 (6) | 472 (6) |
| Multiple indications | 2122 (25) | 1863 (25) |
| ***Pre-operative characteristics*** | | |
| Estimated risk of death | | |
| Not documented | 3762 (45) | 2468 (34) |
| Low (<5%) | 1354 (16) | 1646 (22) |
| Medium (5-10%) | 1019 (12) | 1102 (15) |
| High (>10%) | 2197 (26) | 2145 (29) |
| ASA grade | | |
| I | 615 (7) | 533 (7) |
| II | 2815 (34) | 2461 (33) |
| III | 3112 (37) | 2745 (37) |
| IV | 1605 (19) | 1465 (20) |
| V | 187 (2) | 156 (2) |
| P-POSSUM score (median [IQR]) | 7.6 (2.9-22.7) | 7.4 (2.8-22.9) |
| Systolic blood pressure (mean [SD]) | 128 (24) | 128 (25) |
| Glasgow coma score (mean [SD]) | 14.8 (1.4) | 14.7 (1.5) |
| Blood lactate (median [IQR)] | 1.6 (1.1-2.8) | 1.5 (1.0-2.6) |

**Table 3.3. Patient level process measures. Data presented as n (%).**

[a] 29 patients in the usual care group and 27 patients in the QI group died during surgery.

| | **Summary measure** | |
|---|---|---|
| **Process measure** | **Usual care** | **Quality improvement** |
| Consultant decision to operate | 7472 (90) | 6589 (90) |
| Consultant reviewed patient at time of decision | 5961 (85) | 5271 (84) |
| Pre-operative documentation of risk | 4570 (55) | 4893 (66) |
| Patient entered operating theatre within specified urgency time frame | 5636 (75) | 5515 (79) |
| Consultant surgeon present in operating theatre | 7117 (85) | 6472 (88) |
| Consultant anaesthetist present in operating theatre | 6313 (76) | 5832 (79) |
| Goal directed fluid therapy used during surgery | 3942 (47) | 4329 (59) |
| Serum lactate measured at end of surgery | 4474 (54) | 4431 (60) |
| Time from decision to operate to entry into operating theatre (hours) | 5.0 (2.1-16.8) | 4.3 (2.0-15.3) |
| Critical care admission immediately after surgery[a] | 5395 (65) | 5050 (69) |

**Table 3.4. Patient outcomes. Data presented as median (IQR), n (%) or hazard ratio with 95% confidence intervals.**

| | Summary outcome measure | | |
|---|---|---|---|
| | **Usual care** | **Quality improvement** | **Hazard ratio (QI vs. usual care)** |
| All-cause mortality within 90 days of surgery | 1393 (16) | 1210 (16) | 1.11 (0.96-1.28) |
| All-cause mortality within 180 days of surgery | 1698 (20) | 1440 (20) | 1.12 (0.98- 1.28) |
| Duration of hospital stay (days) | 8 (13-23) | 8 (13-24) | 0.90 (0.83-0.97) |
| Hospital re-admission within 180 days of surgery | 1618 (20) | 1242 (18) | 0.87 (0.73-1.04) |

**Figure 3.3. Mortality within 90 days of emergency abdominal surgery.**

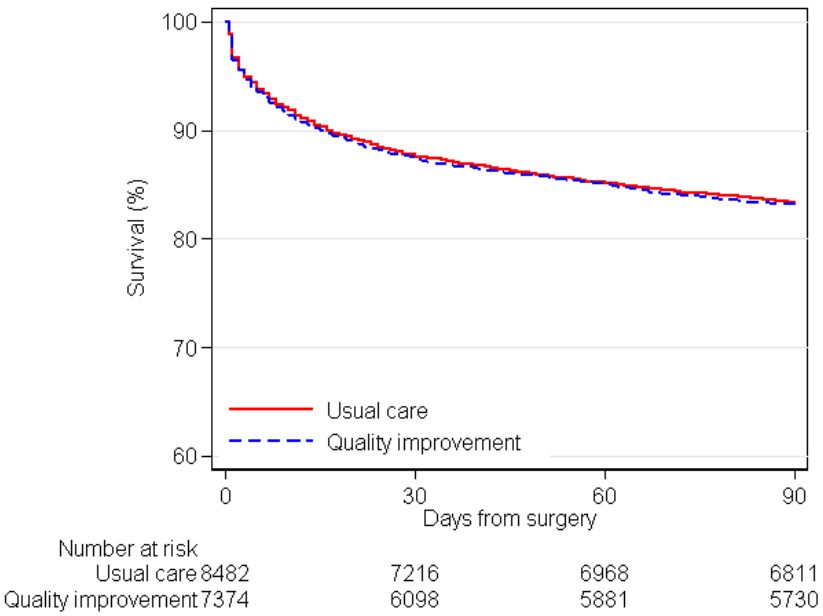
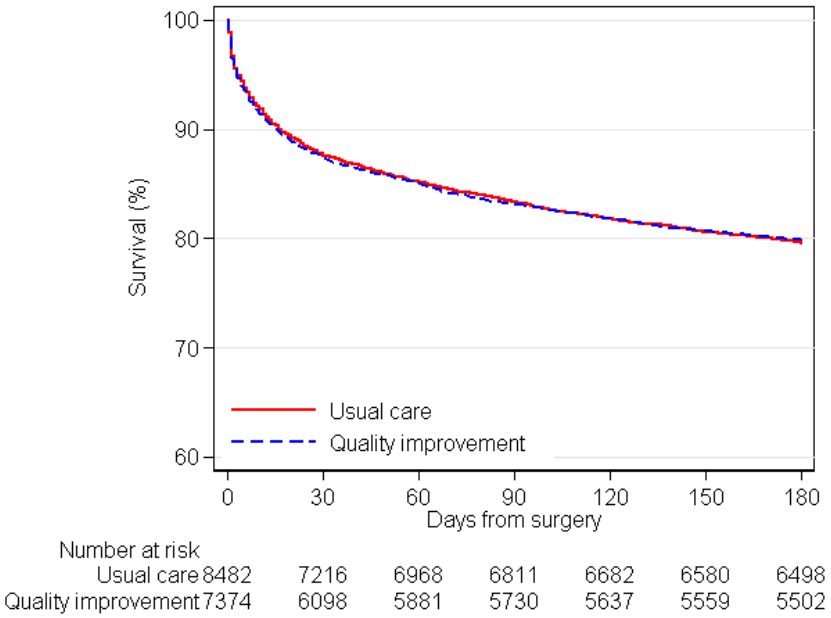
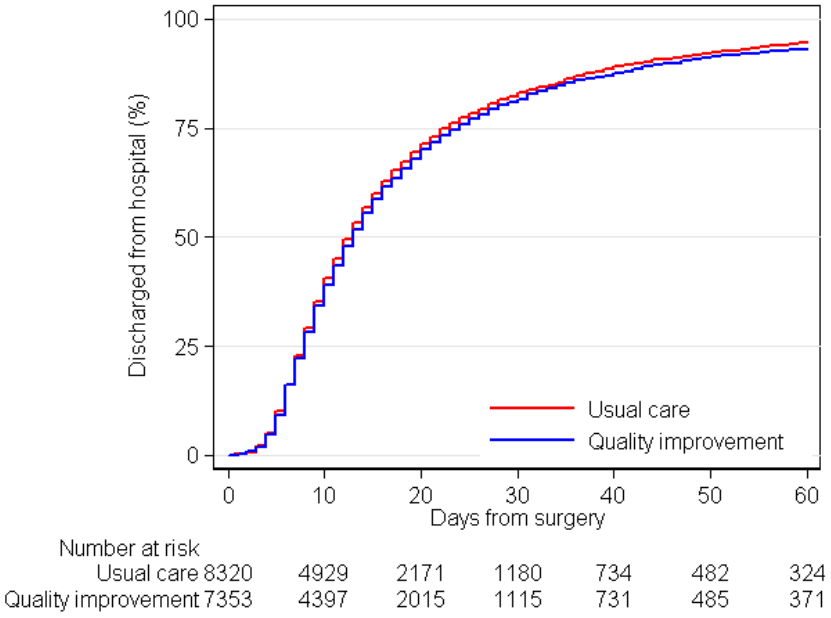


**Figure 3.4. Mortality within 180 days of emergency abdominal surgery.**




| Number at risk | | | | | | | |
|---|---|---|---|---|---|---|---|
| Usual care | 8482 | 7216 | 6968 | 6811 | 6682 | 6580 | 6498 |
| Quality improvement | 7374 | 6098 | 5881 | 5730 | 5637 | 5559 | 5502 |

**Figure 3.5. Duration of hospital stay after emergency abdominal surgery.**




| Number at risk | | | | | | | |
|---|---|---|---|---|---|---|---|
| Usual care | 8320 | 4929 | 2171 | 1180 | 734 | 482 | 324 |
| Quality improvement | 7353 | 4397 | 2015 | 1115 | 731 | 485 | 371 |

**Figure 3.6. Time to hospital re-admission.**

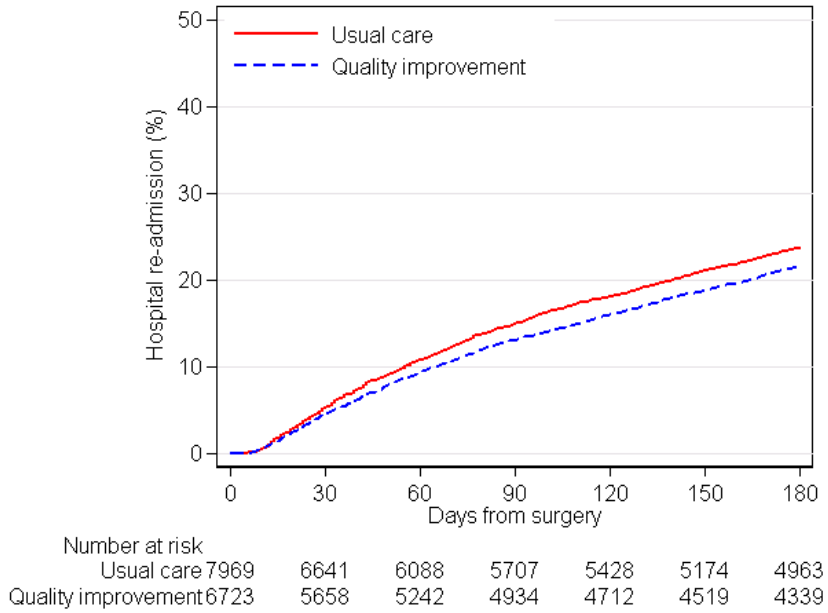



**Table 3.5. Number of hospitals included in each analysis.**

| | **Hospitals (n=93)** |
|---|---|
| All-cause mortality within 90 days of surgery (primary) | 93 (100) |
| All-cause mortality within 180 days of surgery | 93 (100) |
| Duration of hospital stay (days) | 91 (98) |
| Hospital re-admission within 180 days of surgery | 87 (94) |

**Table 3.6. Summary statistics for duration of hospital stay and hospital re-admission.**

| | **Usual care** | **Quality improvement** |
|---|---|---|
| **Duration of hospital stay** | | |
| Censored while in-hospital | 29 (<1) | 102 (1) |
| Discharged | 7195 (86) | 6250 (85) |
| Died in hospital | 1096 (13) | 1001 (14) |
| **Hospital re-admission within 180 days of surgery** | | |
| No | 4954 (62) | 4331 (64) |
| Yes | 1618 (20) | 1242 (18) |
| Died without admission before 180 days | 1397 (18) | 1150 (17) |

**Table 3.7. Summary effect of QI intervention over time.**

| | **90 day mortality** | **Hazard ratio (95% CI)** | **p-value (overall)** |
|---|---|---|---|
| No QI | 1393/8482 (16) | Reference | 0.15 |
| QI <5 weeks | 198/1069 (19) | 1.11 (0.94, 1.32) | - |
| QI 5-10 weeks | 185/983 (19) | 1.21 (1.01, 1.44) | - |
| QI >10 weeks | 1025/6391 (16) | 1.05 (0.90, 1.23) | - |

### *3.5 Discussion*

The principal finding of this trial was that there was no survival benefit associated with a national Quality Improvement programme to implement an evidence-based care pathway for patients undergoing emergency abdominal surgery. Furthermore, there was no beneficial effect on 180-day mortality, hospital stay or hospital readmission. At a national level, there were only modest improvements amongst the ten measures selected to reflect key processes of care within the pathway. In some cases, the baseline rate of adherence to process measures was higher than anticipated. Experience from individual hospitals suggested wide variations in which of the 37 pathway elements local QI teams chose to tackle, the rate of change they achieved, and their eventual success. The baseline contexts of participating hospitals also differed. Implementation of change was slower where existing relationships within and beyond the perioperative team were weaker, and so QI leads had to spend time developing relationships with stakeholders.

At the time of trial design, the EPOCH care pathway was widely agreed to represent an achievable standard of care that informed clinicians would wish to deliver for their patients, but commonly failed to provide because of poor awareness amongst the perioperative team. Our findings reveal that implementation of such an extensive care pathway was a more complex challenge than expected by our clinical community. It is important to interpret the results of this trial alongside those of the process evaluations (outlined in the next two chapters),[135, 172] which together suggest that Quality Improvement programmes designed to implement complex care pathways require more resources, with dedicated time for clinical teams to focus on making change happen.

There are several published reports of the impact of small-scale Quality Improvement projects to improve outcomes for patients undergoing emergency abdominal surgery. In the UK, the ELPQuiC group examined the implementation of a care bundle of five interventions in four NHS hospitals in an uncontrolled before and after study.[138] These authors reported a reduction in mortality (risk ratio 0.61) amongst 726 patients. This study design is more prone to bias than a stepped-wedge cluster randomised trial. The difference in findings may additionally relate to the simpler intervention, and stronger pre-existing relationships between staff leading implementation in these early adopter hospitals. The simpler objective was more readily achieved than that of the national EPOCH trial which set more ambitious targets in hospitals where there may have been a less favourable context for change. Researchers from Denmark reported differing results from three separate studies of perioperative quality improvement interventions for patients undergoing emergency abdominal surgery. The PULP trial group used an uncontrolled before-and-after design with historical controls to study the effect of a 'multidisciplinary perioperative care protocol' in seven hospitals and reported a considerable reduction in 30-day mortality in comparison.[173] However, 56 of the 173 patients allocated to the trial intervention were excluded from the analysis because they did not receive the full intervention, making it harder to interpret these findings. The InCare group did not identify any beneficial effect on 30-day survival from admission to an intermediate unit (critical care) amongst 286 patients undergoing emergency abdominal surgery in seven hospitals.[158] This intervention appeared to change the process of patient care in the 48 hours following surgery, but the trial was stopped for futility partly because of a lower than expected mortality rate in both treatment arms. Finally, the AHA group again studied the effect of a multidisciplinary protocol in a single-

centre uncontrolled before and after study with historical controls, finding a more modest reduction in 30-day mortality from 22% amongst 600 control patients to 16% amongst 600 intervention patients.[174] It is possible that a background trend to improved mortality may explain the findings of these previous studies, especially given the growing international focus on poor patient outcomes following emergency abdominal surgery. Whilst our analysis accounts for temporal trends during the EPOCH trial, it is possible that a decreasing mortality beforehand may explain why the mortality rate was lower than that predicted from NHS registry data.

Meanwhile, recent studies of Quality Improvement in other clinical areas have delivered mixed results.[92, 123, 175, 176] These findings suggest that more focused, discrete clinical interventions may be more successfully implemented than interventions that include larger numbers of care processes. The evidence is less clear in defining the optimal improvement methods. There are several theoretical models of implementation including the Consolidated Framework for Implementation Research and the COM-B model.[112, 177] These provide frameworks for designing and evaluating effective implementation, clinical process and behaviour change. However, none of these models gives emphasis to institutional support or protected leadership time. Our findings suggest these more practical considerations are essential for clinicians to successfully lead quality improvement projects. In the EPOCH trial, teams were encouraged to begin with easier interventions, before building towards full pathway implementation. However, our process evaluation reveals that many teams did not have the time or capacity to progress beyond simpler interventions (e.g. documentation of patient risk) to implementation of more important but challenging

interventions such as admission to critical care. It is also important to note that the National Emergency Laparotomy Audit was launched only three months before the EPOCH trial commenced. Our process evaluation findings (see in particular Chapter 4) suggest that the task of collecting and entering data into the National Emergency Laparotomy Archive (NELA) database was more time-consuming than expected, leaving some QI leads with little time to focus on change. We allowed a five-week period for the transition between usual care and the launch of the quality improvement programme in each cluster. Longer transition and intervention periods with dedicated time for QI leads to plan, negotiate and implement change may have led to more successful implementation. However, we also note that there was no evidence of survival benefit amongst hospitals exposed to the quality improvement programme for longer than 10 weeks, which included hospitals exposed for up to 80 weeks.

The strengths of this trial include wide generalisability (large number of consecutive patients enrolled by many hospitals), robust trial design and the devolved leadership to local clinical QI teams. The EPOCH care pathway was developed through a Delphi consensus process to update national professional guidelines.[11] As with many evidence-based treatment guidelines, some recommendations were graded as strong although the available evidence was weak. The choice of component interventions such as intensive care admission and consultant led care was primarily based on expert opinion; it is unclear how this evidence base could be improved. Partnership with NELA allowed an efficient trial design with no additional data collection for participating staff. However, our final data set required linkage to four national registries in the devolved nations of the UK, and despite completing the trial on time, some organisations involved imposed substantial delays in access to these data

sets. On several occasions, organisations changed their position on information governance regulations, requiring revision of previous agreements between each of the parties involved. In hindsight, we would have encountered fewer problems had we confined the trial to the jurisdictions of fewer organisations with information governance oversight.

Despite the large sample, fewer patients than expected underwent emergency abdominal surgery, and the 90-day mortality rate was lower than anticipated. The sample size calculation was based on Hospital Episodes Statistics data which do not provide a specific diagnostic code for emergency abdominal surgery. Instead we identified a series of codes for relevant procedures. We chose to power the trial to detect a very modest treatment effect partly to accommodate the possibility that these data were poorly representative of the EPOCH trial population. However, the 95% confidence interval for our primary effect estimate was narrow, with a lower limit which indicates a maximum potential mortality reduction of 4%. Our findings are unlikely to change with a larger sample size. Due to difficulty in obtaining post-discharge survival data in Wales, we changed our primary analysis from a binary to a time-to-event approach allowing inclusion of mortality events censored at hospital discharge. However, post-discharge data from England and Scotland suggest few events were missed through this approach. The additional application required to obtain post-discharge mortality data for Wales would have further delayed the trial results by many months.

### *3.6 Conclusions*

In this stepped-wedge cluster randomised trial, we did not identify any survival benefit from a national quality improvement programme to implement an enhanced pathway of care for patients undergoing emergency abdominal surgery. This is likely due to variation between hospitals in fidelity of implementation, prioritisation of pathway components, and the time required to achieve effective change. These findings suggest future quality improvement programmes should implement fewer, more discrete changes and ensure leadership teams have adequate time to achieve sustained improvements in patient care. Undue emphasis on success stories from small early studies may lead us to under-estimate the requirements for successful quality improvement interventions.

# EPOCH Process Evaluation Questionnaire

EPOCH Trial Exit Questionnaire
1. Instructions and purpose

Thank you for taking the time to complete this exit questionnaire. The data you submit will allow us to place into context the results from the main study. As such, your answers are vital to the interpretation of our data.
Some specific points to note:

1. Completion of this questionnaire should be led by the PI or the QI Lead most involved in the running of EPOCH at your site. However, all those involved in EPOCH QI activities should ideally be given the opportunity to contribute to responses.
You may find completing the questionnaire as a team offers the chance to reflect on progress to date.

2. The first section is about the clinical interventions within the EPOCH Pathway. (This section will take the longest to complete.) The following sections focus on the QI activities undertaken as part of the study. There will be then be an opportunity to give the EPOCH Team some feedback.
The final section of the form is about the person completing this form and any others who have contributed.

3. This questionnaire should between 15-30 minutes to complete, depending on how many questions you answer.
We appreciate this is a signifcant time committment. Questions with a red * are required responses (you cannot progress until answered) with all others being optional. Progress can be saved at any time, if you need to stop and continue at a later date.

Although the questionnaire is not anonymous, no individual hospital level data will be presented or published.

Thank you again for your valuable input.

Carol, Rupert, Tim AND The EPOCH Trial team

1. Hospital (study site) name *

2. The EPOCH Clinical intervention

During the trial, hospitals were given some flexiblity as to what clinical interventions & care proceesses to focus on.

In this first section please indicate which clinical interventions and processes from the EPOCH pathway were included in your hospitals improvement activities, once you started the EPOCH study period.

You will then be asked to describe which interventions you found easy to implement and which were more challenging.

2. Using the list below, please indicate which of the pre-operative EPOCH Care Pathway interventions and processes you attempted to improve or implement during EPOCH *

| | Attempted to improve / implement during EPOCH | Did not attempt to improve / implement during EPOCH | Did not attempt to improve / implement: already reliably delivered prior to EPOCH |
|---|---|---|---|
| PRE- OP Consultant-led decision making | ☐ | ☐ | ☐ |
| PRE-OP CT Scan within 2 hours of decision to image | ☐ | ☐ | ☐ |
| PRE-OP Documented mortality risk estimate using formal risk estimate tool (e.g. P-Possum) | ☐ | ☐ | ☐ |
| PRE-OP Patient and relatives given written and oral information about treatment | ☐ | ☐ | ☐ |
| PRE-OP Timely surgery | ☐ | ☐ | ☐ |
| PRE-OP Screening for Sepsis and use of Sepsis 6 as appropriate | ☐ | ☐ | ☐ |
| PRE-OP Analgesia within 1 hour of first medical assessment | ☐ | ☐ | ☐ |
| PRE-OP Screening for coagulopathy and correction as appropriate | ☐ | ☐ | ☐ |
| PRE-OP Maintenance of normothermia | ☐ | ☐ | ☐ |
| PRE-OP Active glucose control | ☐ | ☐ | ☐ |

3. The EPOCH Clinical intervention

3. Using the list below, please indicate which of the intra-operative EPOCH Care Pathway interventions and processes you attempted to improve or implement during EPOCH *

| | Attempted to improve / implement during EPOCH | Did not attempt to improve / implement during EPOCH | Did not attempt to improve / implement: already reliably delivered prior to EPOCH |
|---|---|---|---|
| INTRA-OP Consultant delivered surgery | ☐ | ☐ | ☐ |
| INTRA-OP Consultant delivered anaesthesia | ☐ | ☐ | ☐ |
| INTRA-OP Fluid guided by CO monitoring | ☐ | ☐ | ☐ |
| INTRA-OP Low tidal volume protective ventilation | ☐ | ☐ | ☐ |
| INTRA-OP Serum lactate analysis | ☐ | ☐ | ☐ |
| INTRA-OP Plan for Critical Care admission | ☐ | ☐ | ☐ |
| INTRA-OP WHO checklist performed | ☐ | ☐ | ☐ |
| INTRA-OP Screening for Sepsis and use of appropriate antibiotic therapy | ☐ | ☐ | ☐ |
| INTRA-OP Maintenance of normothermia | ☐ | ☐ | ☐ |
| INTRA-OP Active glucose management | ☐ | ☐ | ☐ |
| INTRA-OP Post-operative analgesia and VTE / N&V prophylaxis prescribed | ☐ | ☐ | ☐ |

4. The EPOCH Clinical intervention

4. Using the list below, please indicate which of the 'End of Surgery ' EPOCH Care Pathway interventions and processes you attempted to improve or implement during EPOCH *

| | Attempted to improve / implement during EPOCH | Did not attempt to improve / implement during EPOCH | Did not attempt to improve / implement: already reliably delivered prior to EPOCH |
|---|---|---|---|
| END OF SURGERY Document core temperature | ☐ | ☐ | ☐ |
| END OF SURGERY Confirm neuromuscular blockade reversal using stimulation device | ☐ | ☐ | ☐ |
| END OF SURGERY Re-check serum lactate | ☐ | ☐ | ☐ |
| END OF SURGERY Re-calculate mortality risk estimate using formal tool (e.g. P-Possum) | ☐ | ☐ | ☐ |
| END OF SURGERY Document fluids given and fluid plan | ☐ | ☐ | ☐ |

5. The EPOCH Clinical intervention

5. Using the list below, please indicate which of the post-operative EPOCH Care Pathway interventions and processes you attempted to improve or implement during EPOCH *

| | Attempted to improve / implement during EPOCH | Did not attempt to improve / implement during EPOCH | Did not attempt to improve / implement: already reliably delivered prior to EPOCH |
|---|---|---|---|
| POST-OP Early pain team review | ☐ | ☐ | ☐ |
| POST-OP Continued antibiotic therapy with microbiology input | ☐ | ☐ | ☐ |
| POST-OP Early nutrition review (surgical / dietician led) | ☐ | ☐ | ☐ |
| POST-OP Physiotherapy on Day 1 after surgery | ☐ | ☐ | ☐ |
| POST-OP Maintenance of normothermia | ☐ | ☐ | ☐ |
| POST-OP Active glucose management | ☐ | ☐ | ☐ |
| POST-OP Daily bloods taken until considered low risk | ☐ | ☐ | ☐ |
| POST-OP Nausea and vomiting prophylaxis given | ☐ | ☐ | ☐ |
| POST-OP VTE prophylaxis given | ☐ | ☐ | ☐ |
| POST-OP Critical Care Outreach Team review on ward | ☐ | ☐ | ☐ |

Please use this space to add any further information you feel is relevant regarding the clinical interventions

6. The EPOCH Clinical intervention

6. Please indicate statement most closely fits your hospitals improvement or implementation activity during EPOCH *

- [ ] We introduced a single pathway of care (across Pre, Intra and Post operative phases)
- [ ] We introduced separate pathways or care bundles for 2 or more phases of the patient admission (e.g. a pre-op pathway plus an intra op care bundle)
- [ ] We introduced separate pathways or bundles for one phase of the patient admission (e.g. pre-op or post op only)
- [ ] We focused on introducing individual / separate interventions
- [ ] Other (please specify):

7. The EPOCH Clinical intervention

7. Please tell us which interventions were easiest to implement / improve and why this was

8. Please tell us which interventions were most challenging to implement / improve and why this was

Well done! That is the hardest and longest part completed.
Now onto some questions about your QI activities.
The majority of these are simple Yes / No or multi-choice questions.

8. Quality Improvement (QI) activities
This second section will cover what QI activities and strategies were used by you and your colleagues when improving care for Emergency Laparotomy patients

9. At your site, was a formal team created to work on QI activities related to EPOCH?

Definition of QI Team:
A group of individuals that work together on the QI project. The team is defined by their shared goals and mutual accountability for the QI project outcome. QI team members are typically responsible for planning and conducting tests of change and/or data collection and management. Members of the QI team may be anyone within the healthcare team, such as doctors, nurses, AHPs, pharmacists, managers, administrative staff. *

- [ ] Yes
- [ ] No
- [ ] Other (please specify):

9. Quality Improvement (QI) activities (continued)

10. Please indicate the approximate size of your QI Team, including yourself *

- [ ] 1
- [ ] 2
- [ ] 3
- [ ] 4
- [ ] 5
- [ ] 6
- [ ] 7
- [ ] 8
- [ ] 9
- [ ] 10
- [ ] >10

11. Which professions and disciplines were involved in your QI Team, including yourself? *

- [ ] Surgeons
- [ ] Anaesthetists
- [ ] Intensivists
- [ ] Radiologists
- [ ] Acute Medicine
- [ ] Emergency Medicine
- [ ] Healthcare of the Elderly physicans
- [ ] Surgeons in training
- [ ] Anaesthetists in training
- [ ] Other doctor in training
- [ ] Nursing - theatres
- [ ] Nursing - ward / critical care
- [ ] Nursing - research (any speciality)
- [ ] Operating Department Practitioners

- [ ] Other Allied Health Professionals
- [ ] Service / departmental managers
- [ ] Senior / executive management
- [ ] Audit / data staff
- [ ] Other (please specify):

10. Quality Improvement (QI) activities (continued)

12. Please indicate which of the methods below best describes your NELA data collection process N.B. If your process has changed significantly over time, pleas indicate which method is in use now and use the comment box to briefly describe this change. *

- [ ] Mostly prospective - the majority of data are collected concurrently and in real time during the peri-operative period by members of the team delivering patient care
- [ ] Mostly retrospective - the majority of data are collected after the peri-operative period by members of the team responsible for that patient care episode
- [ ] Most retrospective - the majority of data are collected after the peri-operative period by other staff not directly involved in that patient care episode
- [ ] Other (please specify):

Comments:

13. Who enters most of the data into the NELA online portal? (you may choose up to 3 options) *

- [ ] EPOCH QI Leads
- [ ] NELA Leads
- [ ] Other clinicians (Consultant grade)
- [ ] Other clinicians (in training)
- [ ] Nursing staff - clinical
- [ ] Nursing staff - research
- [ ] Allied Health Professionals
- [ ] Audit / data staff
- [ ] Other (please specify):

Comments:

14. Were data collected on care processes NOT captured by NELA? e.g. Sepsis screening *

- [ ] Yes
- [ ] No

If YES, please describe briefly:

15. Prior to starting EPOCH did you or your colleagues download and analyse your local NELA data *

- [ ] Yes
- [ ] No
- [ ] Other (please specify):

If YES, please describe briefly:

16. After starting EPOCH did you or your colleagues download and analyse your local NELA data? *

- [ ] Yes
- [ ] No

11. Quality Improvement (QI) activities (continued)

17. What methods did you or your colleagues use to analyse and display your local NELA data? *

- [ ] Run charts
- [ ] Statical Process Control (SPC) charts
- [ ] Bar charts
- [ ] Pie charts
- [ ] Summary statistics
- [ ] Red Amber Green ('RAG') status charts
- [ ] Other (please specify):

12. Quality Improvement (QI) activities (continued)

18. Please indicate approximately how frequently you or your colleagues analysed your local NELA data? *

- [ ] Weekly
- [ ] Fortnightly
- [ ] Monthly
- [ ] Bi-monthly
- [ ] Every 3-4 months
- [ ] Every 6 months
- [ ] Only once - did not update
- [ ] Other (please specify):

13. Quality Improvement (QI) activities (continued)

19. Please use the scale below to rate your agreement with the following statement: "I / we found run-charts helpful when analysing and interpreting our NELA data" *

- [ ] Strongly agree
- [ ] Agree
- [ ] Neutral
- [ ] Disagree
- [ ] Strongly disagree

It would be helpful if you could provide a brief reason for your rating

14. Quality Improvement (QI) activities (continued)

20. Did you feedback your NELA data and analysis to colleagues during EPOCH? *

- [ ] Yes
- [ ] No

Comments:

15. Quality Improvement (QI) activities (continued)

21. Please take a moment to describe how you fedback data to colleagues. Please include:
1. Who you fedback data to;
2. How frequently you did this;
3. What methods you used (e.g. email, EPOCH meetings, departmental events/meetings, posters)
4. Any other activities relevant to feedback of NELA data to colleagues *

22. Please use the scale below to rate your agreement with the following statement: "I / we found run-charts helpful when feeding back NELA data to other colleagues" *

- [ ] Strongly agree
- [ ] Agree
- [ ] Neutral
- [ ] Disagree
- [ ] Strongly disagree
- [ ] Did not use runcharts

It would be helpful if you could provide a brief reason for your rating

23. Please use the scale below to rate your agreement with the following statement: "From my / our experience during EPOCH, feeding back data to colleagues can be an effective strategy to motivate those colleagues to improve care " *

- [ ] Strongly Agree
- [ ] Agree
- [ ] Neutral
- [ ] Disagree
- [ ] Strongly Disagree

Please provide a brief reason for your rating

16. Quality Improvement (QI) activities (continued)

24. Please summarise any board level support you received during the study period

17. Quality Improvement (QI) activities (continued)

25. Did you hold a stakeholder meeting as one of your QI activities? e.g. a meeting for all professionals involved in the care of EmLap patients *

- [ ] Yes
- [ ] No

If YES, please describe briefly:

26. Did you or your colleagues use the "Plan Do Study Act" (PDSA) cycle approach during your QI activities? *

- [ ] Yes, often
- [ ] Yes, occasionally
- [ ] No

18. Quality Improvement (QI) activities (continued)

27. Please use the scale below to rate your agreement with the following statement: " I / we found the PDSA cycle to be a helpful approach to implementation / improvement" *

- [ ] Strongly Agree
- [ ] Agree
- [ ] Neutral
- [ ] Disagree
- [ ] Strongly Disagree

Comments:

19. Your experience of improving care

Nearly done!
In this section, please take a moment to tell us your thoughts on what has worked and not worked for you during EPOCH

28. Reflecting on your experience with EPOCH, please tell us what are the 2 things that you would definitely continue doing if you were to do EPOCH again. *

29. Reflecting on your experience with EPOCH, please tell us 2 things you would do differently if you were to do EPOCH again. *

30. The EPOCH Theory of Change was based upon several key interventions. From your experience with EPOCH please rank these in order of importance. N.B. The 4 choices will move with your ranking decisions. *

| | |
|---|---|
| Using data to drive improvement | |
| Creating the motivation and will to change amongst stakeholders | |
| Fostering inter-professional collaboration and team working | |
| Using QI methods (such as the PDSA cycle) to improve care | |

Comments:

31. You can use this space to tell us more about the barriers and enablers of improvement you have experienced during EPOCH

20. Feedback to the EPOCH Trial team

32. Please rate the support available to you during the EPOCH Trial from the trial team *

- [ ] Very good
- [ ] Good
- [ ] Acceptable
- [ ] Poor
- [ ] Very poor

33. Please tell us what we did that you found helpful

34. Please tell us what we could have done better

35. Please enter your email address here *

36. If others contributed to these responses please list their name and profession or job title here.

37. Would you be willing to be contacted by the EPOCH Trial team to discuss your answers in more detail?

- [ ] Yes
- [ ] No

# CholeQuIC Application form and selection criteria

**Name**
**Address**

X April 2016

Dear X

**Request for expression of interest: Acute gallstone disease quality improvement collaborative**

We are writing to seek your interest in participating in a new acute gallstone disease quality improvement project. This letter sets out what we see as the problem in the delivery of care for acute gallstone disease; what the project will involve, including what participating will mean, and the process for recruiting sites.

If you are interested in participating, please contact Project Manager Erana Sitterlé (esitterle@rcseng.ac.uk or 020 7869 6010) by Friday 29 April and she will be able to explain the next steps. These are also set out in the letter below.

**What is the problem and why do we need to solve it?**
Gallstone related diseases account for approximately one third of emergency general surgery admissions and referrals. There is wide variation in the management of these patients, with cholecystectomy rates within 10 days of first admission (i.e. urgent cholecystectomy) for acute cholecystitis ranging from 0.2% to 35% across England.[1] This is in spite of evidence that shows that urgent cholecystectomy can reduce the rate of recurrent gallstone-related complications[2], and that urgent cholecystectomy is a feasible and safe treatment option in district general hospitals.[3] Currently the average length of stay for these patients is approximately seven days; in analogous health systems (Australia, France) it is 36 hours.[4] We believe that the care of patients admitted with acute gallstone disease could be substantially improved, reducing length of stay and complications and delivering a better standard of care. The overall aims of the project are to reduce time to diagnosis, time to surgery and improve quality of care for patients with acute gallstone disease.

[1] SWORD database: http://www.augis.org/sword/

[2] D. W. da Costa, et al., Same-admission versus interval cholecystectomy for mild gallstone pancreatitis (PONCHO): a multicentre randomised controlled trial. Lancet 2015; 386: 1261-8.

[3] MN Khan et al., Urgent cholecystectomy for acute cholecystitis in a district general hospital – is it feasible? Ann R Coll Surg Engl 2009; 91: 30-34.

[4] AUGIS., Issues in Professional Practice: Pathway for the Management of Acute Gallstone Disease. AUGIS 2015.

**How we propose to solve this problem**

The RCS will fund a multi-site quality improvement project to address the problem outlined above. We will use a quality improvement collaborative approach, with six to ten sites participating in this initial project. An improvement collaborative is an organised, multifaceted approach to quality improvement that involves five essential features[5] to support frontline professionals in improving care:

1. The area of healthcare delivery to be improved has evidence based or expert consensus guidance available but large variations in care exist between these and current practice;
2. Clinical and quality improvement experts provide ideas and support for improvement;
3. A set of inter-professional teams from multiple sites participate to share and learn together and from each other;
4. A model for improvement is used that focuses on setting clear, measurable goals, collecting data and feeding back to relevant stakeholders. Local teams are encouraged to test changes and improvement ideas quickly on a small scale to advance and promote innovation and learning by doing;
5. The collaborative process involves a series of structured activities (meetings, webinars, coaching visits and calls) in a given timeframe to advance improvement.

It is envisaged that the learning from this project would then support the spread of care delivery improvement for this patient group at a wider / national level.

**Your participation in the collaborative**

Involvement in this improvement collaborative will enable you and your colleagues to improve patient outcomes related to acute gallstone disease and develop transferable quality improvement skills.

In order to achieve these aims the teams selected for this project will be expected to participate in the following:

- Four quality collaborative meetings over the project period (June 2016 – January 2018) where site teams will learn quality improvement skills and share ideas and progress. These will be in September 2016, January, July and November 2017 and will most likely be held at the RCS in London;
- Completion of an online quality improvement training programme (www.prism-ed.com);
- Coaching visits and calls with the project improvement team;
- Local measurement of care processes related to acute gallstone disease care (e.g. 'time to ultrasound scanning'). We will define a process measure dataset to augment patient level data obtained via the SWORD database;
- Sharing of both local process and patient level data with the project team and all sites in the collaborative;
- Local application of the recommended improvement methods.

**How to get involved**

[5] Schouten LM, Hulscher ME, van Everdingen JJ, et al. Evidence for the impact of quality improvement collaboratives: systematic review. BMJ 2008;336:1491–4.

The table below sets out what you need to do to get involved in the project.

| **Action required** | **Deadline** |
|---|---|
| Reply to Project Manager, Erana Sitterlé, to confirm interest and availability to attend initial workshop on 28 June at the RCS in London | Friday 29 April |
| Provide the following information:<br>• Signed commitment from Director of Surgery to dedicate staff resource over the course of the project:<br>  o A surgical lead for 1 PA per week<br>  o A project coordinator for at least 1 day per week (this could be a nurse, ODP or trainee, for example)<br>• Confirmation that the project has been discussed with the relevant surgical teams and there is a willingness within the department to participate and change practice<br>• Basic data on patient volumes and workload; we will provide you with a proforma to complete. | Friday 27 May |
| Attend initial project workshop on 28 June at RCS in London – please hold the date for this event, details will be confirmed | Tuesday 28 June |
| If selected to participate in the project, undertake the following project preparation:<br>• Create an inter-professional project team including anaesthetist, nursing staff (ward and theatre, as appropriate), imaging/diagnostics staff, service manager, Board-level sponsor, administrative support<br>• The core project team (surgical lead and project coordinator) to complete the online quality improvement training programme (www.prism-ed.com) | September collaborative meeting (date to be confirmed) |

**Selection of participating hospitals**

We are looking to recruit six to ten hospitals to participate in the project, to ensure the project improvement team can provide sufficient support to each site. We will select the participating sites based on the following:

- ✓ Confirmation of commitment from Director of Surgery received by Friday 27 May, as outlined above
- ✓ Hospitals are not currently undertaking other service improvement activity for acute gallstone disease
- ✓ Priority will be given to sites with a medium to high procedure volume
- ✓ Priority will be given to sites with urgent cholecystectomy rates around or below the national average.

**Further information**

Please contact the RCS Project Manager, Erana Sitterlé, if you have any questions or would like clarification on any points in this letter.

We look forward to hearing from you and possibly working with you on this exciting project.

Yours sincerely

**Mr Ian Beckingham**
Project Clinical Lead

**Mr John Abercrombie**
RCS Council member Lead for Quality Improvement

# CHOLEQuIC Focus groups TOPIC GUIDE

**Welcome and opening points (up to 10 mins)**

1. **Hello – recording – brief**
2. **Purpose of these focus groups**
    a. **CholeQuIc isn't research per se – its an improvement collaborative but we are running a mixed methods evaluation alongside**
        i. **Did it work?**
        ii. **How and why did we achieve the results that we did?**
    b. **Qualitative aspect of evaluation:**
        i. **To learn about what is helping or not helping to achieve the project goals across the cohort**
        ii. **To learn more generally about the experience of surgeons, and their colleagues, in doing QI at the frontline**
        iii. **These focus groups make up part of the data for that**
3. **Confidentiality / use of data**
    a. **Recording**
4. **Recording, incl. test of recording**
    a. **Any questions so far?**
    b. **Get everyone to speak – WHAT WAS THE LAST BOOK THAT YOU READ?**

**CHECK RECORDERS NOW WHILST CONSENT FORMS ARE BEING SIGNED**

5. **Consent forms**
6. **Housekeeping**
7. **Expenses**
8. **How the session will run – timings, may direct the group a little, no right or wrong answers but we may ask for clarification on certain points, really interested in the good, the bad and ugly as that's the way we learn etc**
    a. **Can I check when folks need to leave?**

**DOUBLE CHECK BOTH RECORDERS ARE ON!**

9. **Introductions**
    a. **Name**
    b. **Role in hospital**
    c. **Role in chole-quic**
10. **OK, let's start**

*Let's starts with just describing what you trying to achieve by doing chole-quic in your hospital and detail briefly some of the things you are doing to get you there…*
*Start with you XXX…*

*So now I'd like us to discuss, "**What are some of the things that have helped you so far, as you have been attempting to improve care for patients with gallstones"** and **"What have been the barriers so far, as you have been attempting to improve the care for these patients"***

1. Start with what has helped? (try to limit to X mins)

2. Use probing questions to encourage participants to elaborate

3. Ask focussed questions if necessary to ascertain if other team members, especially those from other disciplines, have similar or different perceptions and thoughts

4. Ask slightly more focussed questions after this to ascertain if there is commonality and/or consensus about what might be the most important / difficult barriers and most important / useful enablers

5. Same for barriers (try to limit to X mins)

6. Summarise and check agreement / correct sense-making

    a. "Fair to say that, based on what you have said, that X, Y and Z have been the common barriers and A has been specific to you (Team 1) and B has been more of an issue for you (Team 2)?"

7. Move onto question 2

<u>Notes</u>

Expected *barriers* to be mentioned (from experience / meeting themes):

- Theatre capacity
- Time pressures
- Lack of priority

Expected *enablers* to be mentioned

- Team work – MDT approach
- Stakeholder engagement
- Understanding current performance and performance gaps - Data
- Learning from each other…
- RCS led

Can mention these if not mentioned by saying

"I was expecting you to mention….Have you not found this a problem"

***Question 2:*** *Next, we want to learn how to make this project, and future projects better. So we would like to give you the opportunity to provide some constructive feedback about Chole-QuIC – about the programme, about the support and resources – to make it better"*
*Cast your mind back to the start of the project – if we were doing a project like this again, what could we do better or differently do you think?*
*"What could we have done better at the start?"*
*"Did you feel sufficiently prepared?"*
*"What could be better about the ongoing support from the project?"*
*"Are the communications you get from us sufficient?"*
*"Have the resources, especially the excel workbooks, been sufficiently useful for you?"*
*"Would you like other things / resources that we are not providing?"*
*"What are we doing well / should keep on doing?"*

**Question 3 -**

*"So finally, we're going to spend a little time thinking about the future QI in surgery and peri-operative care. I'd like you to take a step back from Chole-QuIC and discuss what the key things might be that are required for future surgical improvement projects to be effective"*

Prompt questions:

- The RCS's role in future QI projects
- (The role of the org/ hospital)
- (Commissioners?)
- (The role of the surgical associations??)

- What skills / capabilities are required?
- What support is required?
- What resources are required?
- What training is required?
    - For consultants
    - For trainees
    - For AHPs

Notes

# Chole-QuIC Site Visit – Structured Field Note

| Chole-QuIC Site Visit – Structured Field Note | | | |
|---|---|---|---|
| **Team** | | **Date updated** | |
| **Size (beds)** | | | |
| **DGH / Specialties / HPB** | | | |
| **No. of surgeons who could do the urgent work** | | | |
| **No. of surgeons participating (if known)** | | | |
| **General Reflections (overview paragraph on progress so far)** | | | |
| | | | |
| **Recommended Action in next 3 months** | | | |
| | | | |

| Programme Theory Assessment | |
|---|---|
| **Sense-Making / Coherence (team and hospital wide)** | |
| • Have they got the belief that it's a problem that needs fixing?<br>• What are the short-term and medium goals (& are they clear on them)?<br>• Can they clear on how they can/are achieving these goals?<br>• Are they clear on QI approach? | |
| **Relational / Participation** | |
| Wider System<br>• Are the right people aware and supportive to the work?<br>• Does improvement fit within organisational climate and goals?<br>• Engagement with the system for legitimacy of Chole-QuIC?<br>• Evidence of stakeholders un-blocking or accelerating change? | |
| Team<br>• Are there appropriate leadership capabilities to make change happen?<br>• How well is the team set up to achieve success?<br>• Is the team adapting to needs over time to remain effective? | |
| **Understanding Collective Action** | |
| • Have teams set up the right measurement systems (practical, appropriate and reliable)?<br>• Do they understand the importance of knowing flow?<br>• Are teams developing appropriate ideas for their | |

| | |
|---|---|
| context?<br>• Are they testing iteratively (PDSA)?<br>• Are they making practice agreeable? (i.e. easy for colleagues to adopt) | |
| **Learning / Reflexive Monitoring** | |
| • Are the teams utilising support from others? (Coaches, other teams, experts)<br>• How are the teams review and appraising their progress?<br>• How are they refining their strategies and plans iteratively?<br>• Can they demonstrate effectiveness to others (site stakeholders, broader community)? | |

| Collaborative Goals and drivers | |
|---|---|
| **Reduced Time to Surgery**<br>• What progress are they making in reducing time to surgery?<br>• What progress against 8 day goal?<br>• How sustainable has any reduction been?<br>• Any evidence of other improvements (e.g. patient experience, LoS, Cost)? | |