

This is a repository copy of *A national quality improvement programme to improve survival after emergency abdominal surgery: the EPOCH stepped-wedge cluster RCT*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/151706/>

Version: Published Version

Article:

Peden, Carol J, Stephens, Tim, Martin, Graham et al. (23 more authors) (2019) A national quality improvement programme to improve survival after emergency abdominal surgery: the EPOCH stepped-wedge cluster RCT. Health Services and Delivery Research. ISSN: 2050-4357

<https://doi.org/10.3310/hsdr07320>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Health Services and Delivery Research

Volume 7 • Issue 32 • September 2019

ISSN 2050-4349

A national quality improvement programme to improve survival after emergency abdominal surgery: the EPOCH stepped-wedge cluster RCT

Carol J Peden, Tim Stephens, Graham Martin, Brennan C Kahan, Ann Thomson, Kirsty Everingham, David Kocman, Jose Lourtie, Sharon Drake, Alan Girling, Richard Lilford, Kate Rivett, Duncan Wells, Ravi Mahajan, Peter Holt, Fan Yang, Simon Walker, Gerry Richardson, Sally Kerry, Iain Anderson, Dave Murray, David Cromwell, Mandeep Phull, Mike PW Grocott, Julian Bion and Rupert M Pearse on behalf of the EPOCH trial group



A national quality improvement programme to improve survival after emergency abdominal surgery: the EPOCH stepped-wedge cluster RCT

Carol J Peden,¹ Tim Stephens,² Graham Martin,³
Brennan C Kahan,⁴ Ann Thomson,⁴ Kirsty Everingham,²
David Kocman,³ Jose Lourtie,⁵ Sharon Drake,⁵
Alan Girling,⁶ Richard Lilford,⁷ Kate Rivett,⁸
Duncan Wells,⁹ Ravi Mahajan,¹⁰ Peter Holt,¹¹
Fan Yang,¹² Simon Walker,¹² Gerry Richardson,¹²
Sally Kerry,⁴ Iain Anderson,¹³ Dave Murray,¹⁴
David Cromwell,¹⁵ Mandeep Phull,^{2,16}
Mike PW Grocott,^{17,18} Julian Bion¹⁹
and Rupert M Pearse^{2*} on behalf of the EPOCH
trial group[†]

¹Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

²William Harvey Research Institute, Queen Mary University of London, London, UK

³Health Sciences, University of Leicester, Leicester, UK

⁴Pragmatic Clinical Trials Unit, Queen Mary University of London, London, UK

⁵Royal College of Anaesthetists, London, UK

⁶Institute of Applied Health Research, University of Birmingham, Birmingham, UK

⁷Warwick Medical School, University of Warwick, Coventry, UK

⁸Patient representative, London, UK

⁹Patient representative, Buckinghamshire, UK

¹⁰Faculty of Medicine & Health Sciences, University of Nottingham, Nottingham, UK

¹¹Molecular and Clinical Sciences Research Institute, St George's University of London, London, UK

¹²Centre for Health Economics, University of York, York, UK

¹³Salford Royal Hospital NHS Foundation Trust, Manchester, UK

¹⁴South Tees Hospitals NHS Foundation Trust, Middlesbrough, UK

¹⁵London School of Hygiene and Tropical Medicine, London, UK

¹⁶Department of Anaesthesia and Intensive Care, Queen's Hospital, Romford, UK

¹⁷National Institute for Health Research Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust, Southampton, UK

¹⁸Anaesthesia and Critical Care Research Unit, University of Southampton, Southampton, UK

¹⁹Institute of Clinical Sciences, University of Birmingham, Birmingham, UK

*Corresponding author

†Trial group membership listed in *Appendix 1*

Declared competing interests of authors: Carol J Peden has performed consultancy work for Merck & Co./Merck Sharp & Dohme (Kenilworth, NJ, USA) and for the Institute for Healthcare Improvement (Boston, MA, USA). Graham Martin reports grants from the National Institute for Health Research (NIHR) and personal fees from the BMJ Publishing Group Ltd (London, UK) and is a member of the Health Technology Assessment (HTA) National Stakeholder Advisory Group. Dave Murray reports personal fees from the National Emergency Laparotomy Audit (NELA). David Cromwell reports grants from NELA and is a member of the NELA project team, which provided data and assistance to the project. Mike PW Grocott is a NIHR Clinical Research Network Specialty Lead for Anaesthesia Perioperative Medicine and Pain; the chairperson of the National Institute of Academic Anaesthesia Board; the president of the Critical Care Medicine Section, Royal Society of Medicine; a board member of Evidence Based Perioperative Medicine (EBPOM) UK (London, UK), EBPOM USA (Bannockburn, IL, USA) and EBPOM International; and a board member of Medinspire Ltd (Bridgnorth, UK). Rupert M Pearse holds research grants, has given lectures and/or has performed consultancy work for B.Braun Medical Ltd (Sheffield, UK), GlaxoSmithKline plc (London, UK), Intersurgical Ltd (Wokingham, UK) and Edwards Lifesciences (Irvine, CA, USA). Rupert M Pearse also holds a grant for the submitted work and a NIHR Research Professorship and is a NIHR HTA General Board member.

Published September 2019

DOI: 10.3310/hsdr07320

This report should be referenced as follows:

Peden CJ, Stephens T, Martin G, Kahan BC, Thomson A, Everingham K, *et al.* A national quality improvement programme to improve survival after emergency abdominal surgery: the EPOCH stepped-wedge cluster RCT. *Health Serv Deliv Res* 2019;**7**(32).

Health Services and Delivery Research

ISSN 2050-4349 (Print)

ISSN 2050-4357 (Online)

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HS&DR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hsdr. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Services and Delivery Research* journal

Reports are published in *Health Services and Delivery Research* (HS&DR) if (1) they have resulted from work for the HS&DR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HS&DR programme

The HS&DR programme funds research to produce evidence to impact on the quality, accessibility and organisation of health and social care services. This includes evaluations of how the NHS and social care might improve delivery of services.

For more information about the HS&DR programme please visit the website at <https://www.nihr.ac.uk/explore-nihr/funding-programmes/health-services-and-delivery-research.htm>

This report

The research reported in this issue of the journal was funded by the HS&DR programme or one of its preceding programmes as project number 12/5005/10. The contractual start date was in December 2013. The final report began editorial review in August 2018 and was accepted for publication in May 2019. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HS&DR editors and production house have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the final report document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health and Social Care.

© Queen's Printer and Controller of HMSO 2019. This work was produced by Peden *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

NIHR Journals Library Editor-in-Chief

Professor Ken Stein Professor of Public Health, University of Exeter Medical School, UK

NIHR Journals Library Editors

Professor John Powell Chair of HTA and EME Editorial Board and Editor-in-Chief of HTA and EME journals. Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK, and Honorary Professor, University of Manchester, and Senior Clinical Researcher and Associate Professor, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK

Professor Andrée Le May Chair of NIHR Journals Library Editorial Group (HS&DR, PGfAR, PHR journals) and Editor-in-Chief of HS&DR, PGfAR, PHR journals

Professor Matthias Beck Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Dr Peter Davidson Consultant Advisor, Wessex Institute, University of Southampton, UK

Ms Tara Lamont Director, NIHR Dissemination Centre, UK

Dr Catriona McDaid Senior Research Fellow, York Trials Unit, Department of Health Sciences, University of York, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Wellbeing Research, University of Winchester, UK

Professor John Norrie Chair in Medical Statistics, University of Edinburgh, UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Great Ormond Street Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Ken Stein Professor of Public Health, University of Exeter Medical School, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of editors: www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: journals.library@nihr.ac.uk

Abstract

A national quality improvement programme to improve survival after emergency abdominal surgery: the EPOCH stepped-wedge cluster RCT

Carol J Peden,¹ Tim Stephens,² Graham Martin,³ Brennan C Kahan,⁴ Ann Thomson,⁴ Kirsty Everingham,² David Kocman,³ Jose Lourtie,⁵ Sharon Drake,⁵ Alan Girling,⁶ Richard Lilford,⁷ Kate Rivett,⁸ Duncan Wells,⁹ Ravi Mahajan,¹⁰ Peter Holt,¹¹ Fan Yang,¹² Simon Walker,¹² Gerry Richardson,¹² Sally Kerry,⁴ Iain Anderson,¹³ Dave Murray,¹⁴ David Cromwell,¹⁵ Mandeep Phull,^{2,16} Mike PW Grocott,^{17,18} Julian Bion¹⁹ and Rupert M Pearse^{2*} on behalf of the EPOCH trial group[†]

¹Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

²William Harvey Research Institute, Queen Mary University of London, London, UK

³Health Sciences, University of Leicester, Leicester, UK

⁴Pragmatic Clinical Trials Unit, Queen Mary University of London, London, UK

⁵Royal College of Anaesthetists, London, UK

⁶Institute of Applied Health Research, University of Birmingham, Birmingham, UK

⁷Warwick Medical School, University of Warwick, Coventry, UK

⁸Patient representative, London, UK

⁹Patient representative, Buckinghamshire, UK

¹⁰Faculty of Medicine & Health Sciences, University of Nottingham, Nottingham, UK

¹¹Molecular and Clinical Sciences Research Institute, St George's University of London, London, UK

¹²Centre for Health Economics, University of York, York, UK

¹³Salford Royal Hospital NHS Foundation Trust, Manchester, UK

¹⁴South Tees Hospitals NHS Foundation Trust, Middlesbrough, UK

¹⁵London School of Hygiene and Tropical Medicine, London, UK

¹⁶Department of Anaesthesia and Intensive Care, Queen's Hospital, Romford, UK

¹⁷National Institute for Health Research Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust, Southampton, UK

¹⁸Anaesthesia and Critical Care Research Unit, University of Southampton, Southampton, UK

¹⁹Institute of Clinical Sciences, University of Birmingham, Birmingham, UK

*Corresponding author pearse@qmul.ac.uk

†Trial group membership listed in *Appendix 1*

Background: Emergency abdominal surgery is associated with poor patient outcomes. We studied the effectiveness of a national quality improvement (QI) programme to implement a care pathway to improve survival for these patients.

Objectives: The objectives were to assess whether or not the QI programme improves 90-day survival after emergency abdominal surgery; to assess effects on 180-day survival, hospital stay and hospital readmission; and to better understand these findings through an integrated process evaluation, ethnographic study and cost-effectiveness analysis.

Design: This was a stepped-wedge cluster randomised trial. Hospitals were organised into 15 geographical clusters, and commenced the QI programme in random order over 85 weeks. Analyses were performed on an intention-to-treat basis. The primary outcome was analysed using a mixed-effects parametric survival model, adjusting for time-related effects. Ethnographic and economics data were collected in six hospitals. The process evaluation included all hospitals.

Setting: The trial was set in acute surgical services of 93 NHS hospitals.

Participants: Patients aged ≥ 40 years who were undergoing emergency abdominal surgery were eligible.

Intervention: The intervention was a QI programme to implement an evidence-based care pathway.

Main outcome measures: The primary outcome measure was mortality within 90 days of surgery. Secondary outcomes were mortality within 180 days, length of hospital stay and hospital readmission within 180 days. The main economic measure was the quality-adjusted life-years.

Data sources: Data were obtained from the National Emergency Laparotomy Audit database; qualitative interviews and ethnographic observations; quality-of-life and NHS resource use data were collected via questionnaires.

Results: Of 15,873 eligible patients from 93 NHS hospitals, primary outcome data were analysed for 8482 participants in the usual care group and 7374 in the QI group. The primary outcome occurred in 1393 participants in the usual care group (16%), compared with 1210 patients in the QI group (16%) [QI vs. usual care hazard ratio (HR) 1.11, 95% confidence interval (CI) 0.96 to 1.28]. No differences were found in mortality at 180 days or hospital readmission; there was a small increase in hospital stay in the QI group (HR for discharge 0.90, 95% CI 0.83 to 0.97). There were only modest improvements in care processes following QI implementation. The ethnographic study revealed good QI engagement, but limited time and resources to implement change, affecting which processes teams addressed, the rate of change and eventual success. In some sites, there were challenges around prioritising the intervention in busy environments and in obtaining senior engagement. The intervention is unlikely to be cost-effective at standard cost-effectiveness thresholds, but may be cost-effective over the lifetime horizon.

Limitations: Substantial delays were encountered in securing data access to national registries. Fewer patients than expected underwent surgery and the mortality rate was lower than anticipated.

Conclusions: There was no survival benefit from a QI programme to implement a care pathway for patients undergoing emergency abdominal surgery. The modest impact of the intervention on process measures, despite good clinician engagement, may have been limited by the time and resources needed to improve patient care.

Future work: Future QI programmes must balance intervention complexity with the practical realities of NHS services to ensure that such programmes can be delivered with the resources available.

Trial registration: Current Controlled Trials ISRCTN80682973 and *The Lancet* protocol 13PRT/7655.

Funding: This project was funded by the National Institute for Health Research (NIHR) Health Services and Delivery Research programme and will be published in full in *Health Services and Delivery Research*; Vol. 7, No. 32. See the NIHR Journals Library website for further project information.

Contents

List of tables	xiii
List of figures	xv
List of boxes	xvii
List of abbreviations	xix
Plain English summary	xxi
Scientific summary	xxiii
Chapter 1 Introduction	1
Chapter 2 The EPOCH trial intervention	3
Introduction	3
Trial intervention	3
The cluster-level intervention: the EPOCH trial quality improvement programme	4
The hospital-level interventions	4
<i>Evidence-based care pathway (37 recommended processes of care)</i>	4
<i>The quality improvement intervention, comprising six core improvement strategies</i>	7
Discussion	9
Chapter 3 Stepped-wedge cluster randomised trial	11
Introduction	11
Methods	11
<i>Trial design and participants</i>	11
<i>Data collection</i>	11
<i>Randomisation and masking</i>	11
<i>Trial intervention</i>	12
<i>Outcome measures</i>	12
<i>Statistical analysis</i>	12
<i>Patient and public involvement</i>	13
Results	13
<i>Process measures</i>	13
<i>Clinical outcomes</i>	13
Discussion	20
Chapter 4 Process evaluation	21
Introduction	21
Methods	21
<i>Data sources and data collection</i>	21
<i>Data analysis</i>	21
Results	22
<i>Recruitment of clusters</i>	22
<i>Delivery of the intervention at the cluster level</i>	23
<i>Response to the intervention at cluster level</i>	24
<i>Delivery to individuals at local site (quality improvement intervention fidelity)</i>	25

<i>Use of plan–do–study–act</i>	26
<i>Team approach</i>	26
<i>Use of data feedback and run charts</i>	26
<i>Engagement</i>	26
<i>Segmenting the pathway</i>	26
<i>Response of quality improvement leads: reflections on the change process</i>	26
Discussion	26
Conclusions	30
Chapter 5 Ethnographic evaluation	31
Introduction	31
Methods	32
<i>The EPOCH trial approach as received and understood</i>	34
<i>The EPOCH trial approach as enacted</i>	36
<i>Use of plan–do–study–act</i>	36
<i>Team approach</i>	37
<i>Use of data</i>	38
<i>Engagement</i>	38
The six sites in detail	39
<i>Site 1</i>	39
<i>Site 2</i>	42
<i>Site 3</i>	45
<i>Site 4</i>	48
<i>Site 5</i>	50
<i>Site 6</i>	52
Discussion	55
Chapter 6 Health economic analysis	59
Introduction	59
Methods	59
<i>Overview</i>	59
<i>Data sources</i>	59
<i>Resource use and costs</i>	60
<i>Health-related quality of life</i>	60
<i>Missing data</i>	60
<i>Analysis</i>	61
Results	62
<i>Patient characteristics</i>	62
<i>Missing data</i>	62
<i>Health-care resource use</i>	63
<i>Costs</i>	63
<i>Health-related quality of life</i>	64
<i>Base-case analysis</i>	64
<i>Regression</i>	66
<i>Scenario analysis</i>	67
<i>Sensitivity analysis</i>	68
<i>Subgroup analysis</i>	70
Discussion	72
Conclusion	73
Chapter 7 Discussion	75
Implications for clinical practice	76
Recommendations for future research	76

Acknowledgements	79
References	83
Appendix 1 Members of the EPOCH trial group	89

List of tables

TABLE 1 The EPOCH trial programme theory: main interventions, desired outcomes and evidence for inclusion	8
TABLE 2 Number of patients recruited in each cluster in each period	15
TABLE 3 Baseline patient characteristics	16
TABLE 4 Patient-level process measures	17
TABLE 5 Patient outcomes	18
TABLE 6 Number of hospitals included in each analysis	19
TABLE 7 Summary statistics for duration of hospital stay and hospital readmission	19
TABLE 8 Summary effect of QI intervention over time	20
TABLE 9 90-day mortality in patients who underwent laparoscopic surgery or were aged 18–40 years and met all other inclusion criteria	20
TABLE 10 Data collected for process evaluation	22
TABLE 11 Common themes identified from feedback regarding the QI programme	24
TABLE 12 Reported usage of each QI strategy	25
TABLE 13 Themes emerging from inductive content analysis of QI leads' responses	28
TABLE 14 Characteristics of the six sites sampled	33
TABLE 15 Data collection in all sites	33
TABLE 16 Unit costs of resource use	60
TABLE 17 Baseline characteristics of patients in subsample and EPOCH trial population	62
TABLE 18 Missing data in resource use and HRQoL collected in the subsample	62
TABLE 19 Resources used by the subsample over the trial period	63
TABLE 20 Costs of resources used by the subsample over the trial period, costs of intervention and total costs	63
TABLE 21 EuroQol-5 Dimensions, three-level version scores of the subsample and QALYs over the trial period	64
TABLE 22 Cost-effectiveness analysis in the subsample over trial period	64

TABLE 23 Marginal effect on costs of readmission and costs of outpatient and primary care	66
TABLE 24 Marginal effect on HRQoL decrement at baseline	66
TABLE 25 Changes in HRQoL from randomisation based on the subsample	67
TABLE 26 Estimated costs, EQ-5D-3L scores and QALYs for the EPOCH trial population over the trial period and over the lifetime horizon	67
TABLE 27 Cost-effectiveness analysis in the subsample over the trial period using different assumptions on HRQoL	68
TABLE 28 Baseline characteristics of patients in subsample and complete cases	69
TABLE 29 Cost-effectiveness analysis in the complete-case analysis over trial period	69
TABLE 30 Cost-effectiveness analysis in the subgroups defined by P-POSSUM score: high-risk group	71
TABLE 31 Cost-effectiveness analysis in the subgroups defined by P-POSSUM score: low-risk group	72

List of figures

FIGURE 1 The EPOCH trial care pathway	6
FIGURE 2 Inclusion of hospitals and patients in the trial	14
FIGURE 3 Mortality within 90 days of emergency abdominal surgery	17
FIGURE 4 Mortality within 180 days	18
FIGURE 5 Duration of hospital stay after emergency abdominal surgery	18
FIGURE 6 Time to hospital readmission	19
FIGURE 7 The QI programme as planned and as delivered	23
FIGURE 8 Quality improvement lead attendance at QI meetings	24
FIGURE 9 Clinical process change attempted during intervention period	27
FIGURE 10 Cost-effectiveness acceptability curve for the QI intervention over (a) trial period; and (b) lifetime horizon	65
FIGURE 11 Cost-effectiveness acceptability curve for the QI intervention in the high-risk group (a) within the trial period; and (b) over the lifetime horizon	70

List of boxes

BOX 1 The EPOCH trial care pathway components	5
BOX 2 Summary of programme theory for the EPOCH trial QI intervention	7

List of abbreviations

A&E	accident & emergency	NELA	National Emergency Laparotomy Audit
CI	confidence interval		
ELPQuiC	emergency laparotomy pathway quality improvement care	ONS	Office for National Statistics
		PDSA	plan–do–study–act
EPOCH	Enhanced Peri-Operative Care for High-risk patients	P-POSSUM	Portsmouth Physiological and Operative Severity Score for the enumeration of Mortality and morbidity
EQ-5D-3L	EuroQol-5 Dimensions, three-level version		
GLM	generalised linear model	QALY	quality-adjusted life-year
HES	Hospital Episode Statistics	QI	quality improvement
HR	hazard ratio	R&D	research and development
HRQoL	health-related quality of life	SQUIRE	Standards for Quality Improvement Reporting Excellence
ICER	incremental cost-effectiveness ratio	TIDieR	Template for Intervention Description and Replication
ICU	intensive care unit		
IQR	interquartile range	VLE	virtual learning environment
IT	information technology		

Plain English summary

Background

Emergency abdominal surgery can result in poor outcomes for patients. Around 30,000 patients have this surgery each year in the NHS and 1 in 10 will die within 30 days. There are wide variations in care between hospitals, and researchers have already studied the effect of quality improvement (QI) programmes to encourage the use of routine treatments known to help these patients. These studies suggest improved patient survival, but most used poor-quality research designs, which could produce misleading results.

What we did

We studied whether or not a large national QI programme for patients undergoing emergency abdominal surgery would improve survival. Our research was designed to resolve problems that may have previously caused misleading results. This involved placing hospitals into small groups, or clusters, and launching the QI programme every 5 weeks in one cluster selected at random. We also interviewed and observed staff participating in the study to understand how the programme worked and used an economic analysis to establish whether or not the programme offered value for money.

What we found

We included 15,873 patients aged ≥ 40 years in 93 NHS hospitals. Medical teams engaged positively with the QI programme, but had limited time and resources to implement all the suggested changes. As a result, there were only small changes in patient care and the QI programme did not improve survival. There were wide variations between hospitals in which aspects of patient care teams tried to change. Our analysis of the associated costs suggested that the programme may not be cost-effective, although this was subject to considerable uncertainty.

What this research tells us

Despite the early success of small projects, we found no survival benefit from a national QI programme for patients undergoing emergency abdominal surgery. Our research suggests that insufficient staff time may have been a critical factor, but there are other possible reasons. Similar programmes will succeed only if staff have the time and resources needed to deliver them.

Scientific summary

Background

Emergency abdominal surgery is associated with poor postoperative outcomes. Around 30,000 patients undergo this type of surgery each year in the UK NHS, with 30-day mortality rates in excess of 10% and wide variation in standards of care between hospitals. Several groups have studied the effect of quality improvement (QI) initiatives to implement individual interventions or 'care bundles' of several treatments, thereby improving care for these patients. Overall, the findings of these small studies suggest survival benefit, but most utilised uncontrolled cohort designs, which are associated with a high risk of bias. The feasibility and benefit of a national QI programme to implement a more extensive acute care pathway for this patient group remain uncertain. We studied the effectiveness of a national QI programme to implement a care pathway to improve survival for these patients.

Trial intervention

The Enhanced Peri-Operative Care for High-risk patients (EPOCH) trial care pathway was developed through an evidence-based Delphi consensus process, building on previously published guidelines, to produce a list of 37 component interventions considered best usual practice for this patient group. We used an evidence-based QI programme to implement this care pathway by changing the practice and culture of patient care. QI leads from surgery, anaesthesia and critical care were tasked with leading a hospital-wide improvement programme to implement the care pathway, with the support and guidance of the national EPOCH trial QI team. The key features of the QI methodology were (1) reframing the high mortality rates for these patients as a 'social problem', requiring reorganisation of existing care processes rather than technical innovation; (2) supporting QI leads to engage their front-line staff and executive leaders in the change process; (3) training local QI leads in basic improvement skills; and (4) supporting teams to analyse and feed back key process measure data to their colleagues to drive change. The EPOCH trial QI team provided a 1-day activation and education meeting for each geographical cluster shortly before or during the first week of activation, with further advice and support by telephone and e-mail. All QI resources, including data analysis tools, training materials and promotional documents, were available online through a virtual learning environment (VLE). Follow-up meetings were held 16 weeks after activation, so that QI leads and their teams could meet and share experiences. There were also two national meetings to facilitate shared learning during the trial period. QI leads were eligible to attend these only if their hospital had been activated to the trial intervention.

Methods

We studied the effect of the QI programme to implement the EPOCH trial care pathway in a stepped-wedge cluster randomised trial. This design allowed all participating hospitals to adopt the intervention at some point during the trial, while adjusting for the effect of temporal changes during the trial period. NHS hospitals delivering an emergency general surgical service were eligible for inclusion, provided they undertook a significant volume of emergency abdominal surgery cases. 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 that 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 aged ≥ 40 years and undergoing emergency open abdominal surgery in a participating hospital during the 85-week trial period. Patients were excluded from the analysis if they were undergoing a simple appendicectomy, 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. Sample size calculations

were based on an analysis with fixed time effects and random cluster effects, modified to exclude data collected during the 5-week period in which the intervention commenced in individual clusters. Using Hospital Episode Statistics (HES) 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 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%. Hospitals were organised into 15 geographical clusters and commenced the QI programme in random order, based on a computer-generated random sequence, over an 85-week period. Trial data were collected through the National Emergency Laparotomy Audit database and linked using unique patient identifiers to HES and the Office for National Statistics in England and Wales and the Information Services Division of NHS Scotland. The trial was approved by the East Midlands (Nottingham 1) Research Ethics Committee.

The primary outcome measure was mortality within 90 days of surgery. Secondary outcomes were 180-day mortality, length of hospital stay and hospital readmission within 180 days. Analyses were performed on an intention-to-treat basis. The primary outcome was analysed using a mixed-effects parametric survival model, adjusting for time-related effects. In accordance with our analysis plan, we did not test patient-level process measures for statistical significance.

The trial included an ethnographic study of the adoption of the care pathway in six theoretically sampled sites in the UK and a mixed-method process evaluation of the QI programme and all trial sites. The ethnography combined observation of routine practice with interviews with team members in sampled sites to provide a rich qualitative understanding of the delivery of the QI training and the work of clinical leads in seeking to improve care in their hospitals. Data analysis for the ethnographic data were based on the constant comparative method and informed by sensitising concepts from the literature and discussions among the EPOCH trial team.

For the process evaluation, we collected a range of QI programme activity data (including data on participation in programme activities, such as meetings, and use of the trial VLE) 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. Only one response was required per hospital, but QI leads were asked to complete the questionnaire with colleagues. The programme activity and questionnaire data were analysed and reported using descriptive statistics. Free-text data in the exit questionnaire were coded by two investigators, using both inductive and deductive content analysis techniques.

For the health economic analysis, data describing survival and inpatient stay after surgery, and data on resource use and health-related quality of life (measured using the EuroQol-5 Dimensions, three-level version), collected in a sample of eight trial hospitals, were employed to estimate costs and quality-adjusted life-years (QALYs) of patients receiving the QI intervention and those receiving usual care in the subsample. Within-trial analyses were conducted first and then the cost-effectiveness over the lifetime was assessed by employing assumptions of long-term effects. Using regression models developed based on the subsample, costs and effects were estimated for the full trial population and the cost-effectiveness was assessed within the trial period and over the lifetime horizon.

Findings

In total, 15,873 eligible patients underwent surgery in 93 NHS hospitals between 3 March 2014 and 19 October 2015. Primary outcome data were analysed for 8482 patients in the usual care group and 7374 patients in the QI group. There were only modest improvements among the 10 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. The primary outcome occurred in 1393 patients in the usual care group (16%) compared with 1210 patients in the QI group (16%) [QI vs. usual care hazard ratio (HR) 1.11,

95% confidence interval (CI) 0.96 to 1.28]. We found no differences in 180-day mortality or hospital readmission, and a small increase in hospital stay in the QI group (HR for discharge 0.90, 95% CI 0.83 to 0.97). There were only modest overall improvements in processes of patient care following QI implementation.

The process evaluation data confirmed that the QI programme was delivered as intended, with further activities and resources added in response to the needs of teams in individual sites. Attendance at the initial cluster educational meeting was very good, as was use of the online resources, but attendance at the cluster follow-up meetings was lower. The evaluation identified considerable variation in the use of the of the hospital-level interventions, including differences in which components of the pathway teams attempted to improve and the QI approaches they chose to do this. Reflections from QI leads, captured by the exit questionnaire, suggest that the social aspects of improving care are very important but time-consuming, as are activities related to measurement and data feedback to colleagues.

The ethnographic study revealed near-universal receptivity to the concept of a pathway as a means of improving perioperative processes and outcomes, but concern about the impact on appropriate professional judgement. Teams in the hospitals studied took different approaches to QI, particularly in the extent to which they made use of data and how they used the plan–do–study–act methodology. Some sites were troubled by wider organisational turbulence, which made achieving change through the EPOCH trial challenging; leads in other found it difficult to create a ‘burning platform’ that would prompt behavioural change among their colleagues. In all sites, the EPOCH trial was but one change among many, but there were signs that it could play a significant role in longer-term improvement trajectories.

In the base-case analysis on the subsample of patients in the health economics analysis, the QI intervention was associated with higher costs (mean difference £458) but fewer QALYs gained over the trial period (mean difference –0.002) than usual care; thus, the QI intervention was dominated. However, over the lifetime horizon, the QI intervention was associated with higher costs (mean difference £1508) and more QALYs gained (mean difference 0.131), resulting in an incremental cost-effectiveness ratio (ICER) of £11,511 per QALY, lower than the recommended threshold. The probabilities of being cost-effective at cost-effectiveness thresholds of £13,000, £20,000 and £30,000 per QALY were 51.8%, 56.3% and 58.7%, respectively, indicating considerable uncertainty. For the whole EPOCH trial population, the QI intervention was associated with higher costs and fewer QALYs over the trial period (i.e. it was dominated). However, over the longer time horizon, QI was associated with higher costs and more lifetime QALYs, generating an ICER of £31,632, higher than the highest cost-effectiveness threshold value of £30,000 per QALY. For patients at a higher risk of mortality and morbidity before surgery, this intervention was associated with more QALYs and increased costs within the trial period and over the lifetime horizon. The corresponding ICERs were £119,400 per QALY and £3839 per QALY, respectively.

Conclusions

Despite the success of some smaller projects, there was no survival benefit from a national QI programme to implement a care pathway for patients undergoing emergency abdominal surgery. Through a mixed-methods approach, we were able to identify the reasons for the difficulties in changing patient care in individual hospitals. To succeed, large, national QI programmes need to allow for differences between hospitals and ensure that teams have both the time and the resources needed to improve patient care. The intervention was not cost-effective within the trial period. However, the intervention may be cost-effective for the lifetime horizon, particularly for patients at greatest risk of death.

Trial registration

This trial is registered as ISRCTN80682973 and *The Lancet* protocol 13PRT/7655.

Funding

Funding for this study was provided by the Health Services and Delivery Research programme of the National Institute for Health Research.

Chapter 1 Introduction

More than 1.53 million adults undergo inpatient surgery in the UK NHS each year, with a 30-day mortality of 1.5%.¹ However, patients undergoing emergency abdominal surgery have a much greater risk of death.^{2,3} Around 30,000 patients undergo these procedures in NHS hospitals each year, with 30-day mortality rates in excess of 10%.² There are widespread variations in standards of care between hospitals,^{2,3} including the involvement of senior surgeons and anaesthetists and postoperative admission to critical care. These variations have been associated with differences in mortality rates.^{2,3}

In small studies, quality improvement (QI) initiatives to implement either individual interventions or 'bundles' including several treatments have been associated with improved survival after emergency abdominal surgery.⁴⁻⁷ In a report commissioned by the UK Department of Health and Social Care, the Royal College of Surgeons proposed more extensive improvements to quality of care for this patient group.⁸ Recommendations included consultant-led decision-making, cardiac output-guided fluid therapy and early admission to critical care. However, the feasibility of implementing such an extensive acute care pathway on a national scale, and the benefits of doing so, remain uncertain. There are good examples in which discrete QI interventions have been associated with improved patient outcomes,^{9,10} but other studies yielded disappointing results.¹¹⁻¹³ This is especially true for complex interventions requiring co-ordinated change across a health-care system.¹⁴⁻¹⁷ The benefits of QI initiatives are self-evident to some,¹⁸ but others question the value of these projects, citing high costs, failure to engage clinicians and a lack of scientific rigour.^{19,20} Despite this, the direction in health-care policy is towards ever more widespread use of QI to drive large-scale change.²¹

The launch of the National Emergency Laparotomy Audit (NELA) in December 2013² provided a unique opportunity to study a QI programme to implement a complex care pathway at a national level. We conducted a stepped-wedge cluster randomised trial, with an embedded ethnographic evaluation process and health economic analysis, to evaluate the effect of implementing this pathway on survival following emergency abdominal surgery in NHS hospitals.

Chapter 2 The EPOCH trial intervention

Parts of this report have been reproduced from Stephens *et al.*²² This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated. The text below includes minor additions and formatting changes to the original text.

Introduction

The use of QI approaches to reduce variations in health-care processes and improve the standards of health-care delivery has been increasingly encouraged over the previous 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 different professionals and/or patient groups.²³ There is published guidance on accurately describing complex interventions such as those used in QI [e.g. the Template for Intervention Description and Replication (TIDieR) checklist²⁴ and the Standards for Quality Improvement Reporting Excellence (SQUIRE) guidelines²⁵], but such detailed reporting has not always been the norm in the QI literature. In this chapter we describe the design, rationale and delivery plan for the Enhanced Peri-Operative Care for High-risk patients (EPOCH) trial intervention in line with this guidance. *Chapters 4 and 5* describe how the intervention was delivered in reality, how it worked in practice and how this differed from the intended design.

Trial intervention

The trial intervention was a QI 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 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 trial QI 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 developed a QI programme to train and support QI leads and their colleagues in the delivery of the hospital-level intervention. The hospital-level intervention was what was to be delivered by QI leads and their colleagues in each of the 93 trial sites. QI 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.²⁶ 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 QI 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. In this chapter we detail the development and content of these interventions.

The cluster-level intervention: the EPOCH trial quality improvement programme

We developed a QI 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 trial QI team. The key aims of the programme were to (1) reframe the high mortality rates for these patients as a 'social problem', requiring reorganisation of existing care processes rather than technical innovation; (2) support QI leads to engage their front-line colleagues and executive leaders in the change process; (3) train local QI leads and their colleagues in basic improvement skills based around the model for improvement;²⁷ and (4) support teams to analyse and feed back key process measure data to their colleagues to drive change. The EPOCH trial QI team provided a 1-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 14 weeks before the date of activation to the QI 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 trial QI strategies. This was created to support QI leads who had attended the training and desired further QI 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 trial pathway (Box 1). All hospital QI leads were automatically registered for the VLE 5 weeks prior to activation, and could request that additional colleagues and team members be registered.

Once a cluster was activated, telephone and e-mail support for the intervention was available. Separate e-mail 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 trial 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 eligible to attend these only if their hospital had been activated to the trial intervention.

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.⁸ 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 Grading of Recommendations, Assessment, Development and Evaluations (GRADE) criteria.²⁸ The 37 component interventions are detailed in Box 1 and a graphical display, designed for the QI programme to show how the patient may move along the care pathway, is in Figure 1.

BOX 1 The EPOCH trial care pathway components**Before surgery**

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

During surgery

11. Surgery within 6 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 postoperative analgesia.
20. Prescribe postoperative nausea and vomiting prophylaxis.
21. Prescribe postoperative 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 6 hours of surgery.
28. Analgesia: early review by acute pain team.
29. Continued antibiotic therapy when indicated with microbiology review.
30. Prophylaxis for post-operative nausea and vomiting.
31. Venous thromboembolism prophylaxis.
32. Maintain normothermia.
33. Active glucose management.
34. Daily haematology and biochemistry until mortality risk is low (senior opinion).
35. Nutrition: early dietitian review with consideration of benefits of enteral feeding.
36. Chest physiotherapy review on day 1 after surgery.
37. Critical care outreach review on standard ward with use of early warning scores.

CT, computerised tomography; WHO, World Health Organization.

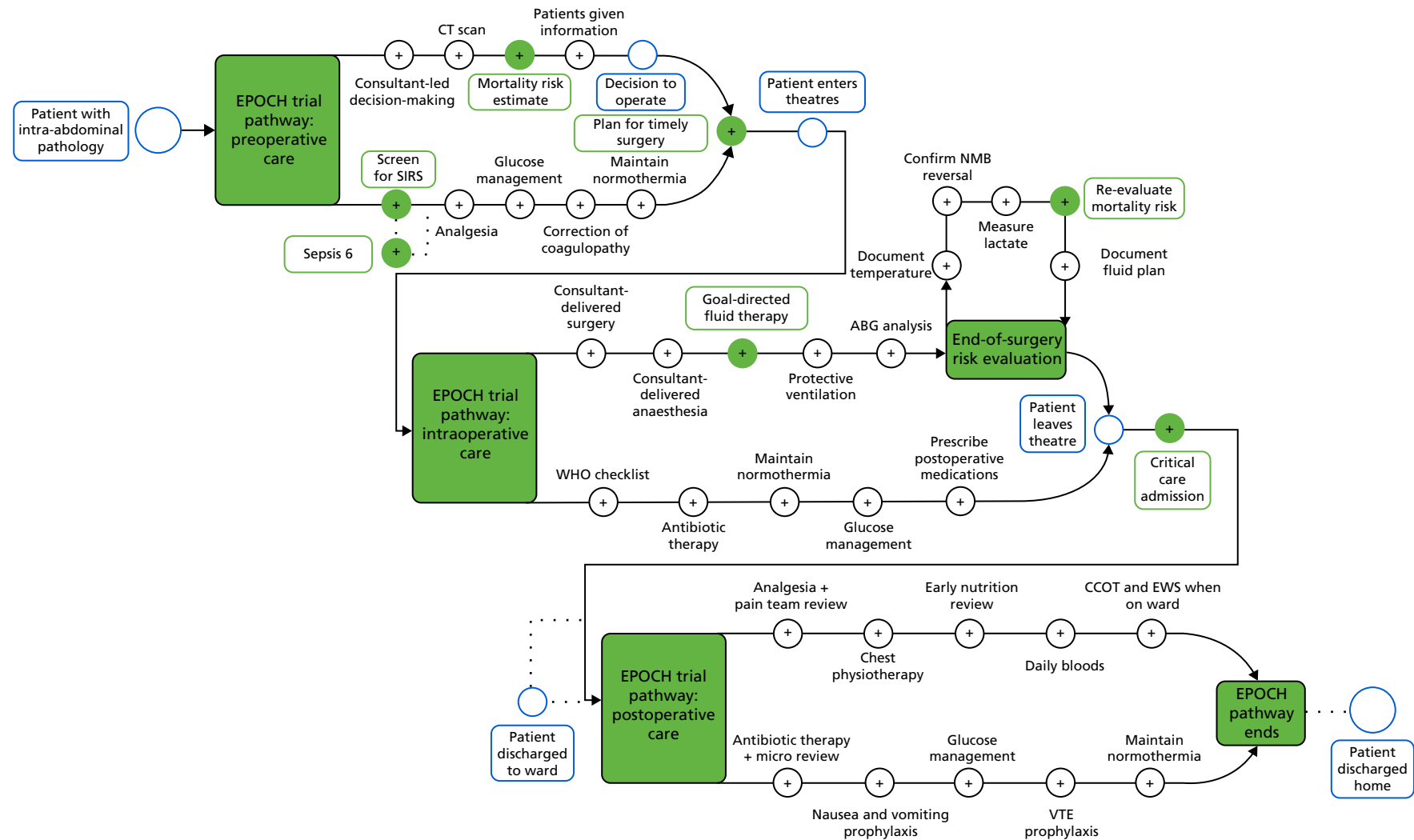


FIGURE 1 The EPOCH trial care pathway. ABG, arterial blood gas; CCOT, critical care outreach team; CT, computerised tomography; EWS, early warning score; NMB, neuromuscular blockade; SIRS, systemic inflammatory response syndrome; VTE, venous thromboembolism; WHO, World Health Organization. Reproduced from Stephens *et al.*²² This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated. The figure includes minor additions and formatting changes to the original figure.

The quality improvement intervention, comprising six core improvement strategies

Two clinicians with QI and training expertise (TS and CJP) developed the programme theory for the QI intervention, defining 'the how' and 'the why' of the QI intervention (Box 2 and Table 1). The EPOCH trial programme theory was based on current evidence and learning from a range of other QI programmes.^{29–32} Six QI 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 interprofessional 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 plan–do–study–act (PDSA) cycles to support process change.

Table 1 details the relationship between the EPOCH trial programme theory, the QI resources available and the QI strategies proposed.

BOX 2 Summary of programme theory for the EPOCH trial QI intervention

If ...

relevant data are reviewed and fed back to teams regularly, and
key professionals come together to form an improvement team, and
QI leads and colleagues learn basic QI approaches, and
relevant stakeholders are made aware of the project and improvement goals ...

then ...

a shared view of performance and improvement gaps can be created, and
professionals can work as a team to define and achieve local improvement goals, and
basic QI approaches can be employed to achieve the improvement goals, and
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.

Reproduced from Stephens *et al.*²² This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated. The box includes minor additions and formatting changes to the original table.

TABLE 1 The EPOCH trial programme theory: main interventions, desired outcomes and evidence for inclusion

Desired outcome	QI strategy	QI programme activity and resources	Evidence for inclusion in programme theory
Motivation for change created among stakeholders and improvement goals clearly understood	QI leads hold a stakeholder meeting after activation (QI strategy 1)	<ul style="list-style-type: none"> Pre-activation checklist (providing guidance for planning of stakeholder meeting) Evidence for QI and need for change provided Presentation on achieving engagement 	<ul style="list-style-type: none"> 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 the Michigan Keystone and Enhanced Recovery programmes)^{30,33} 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 change³⁴
Interprofessional collaboration fostered	Each hospital to form an interprofessional improvement team (QI strategy 2)	<ul style="list-style-type: none"> Team approach promoted QI leads encouraged to invite colleague to EPOCH trial meetings EPOCH trial VLE open to all local QI team members 	<ul style="list-style-type: none"> There is sound theoretical and empirical evidence for the specific role of clinically led QI teams in successful QI^{35,36}
Shared view of current performance created ('situational awareness')	QI leads analyse their own data (NELA data \pm case note reviews and local audit data) and feed this back to colleagues regularly (QI strategy 3)	<ul style="list-style-type: none"> Case note review tool Training on data for improvement Training on how to access and analyse NELA data Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA) workbook programmed to create run charts from NELA data Secure data-sharing site created on VLE 	<ul style="list-style-type: none"> Creating situational awareness regarding clinical performance is seen as fundamental to the model for improvement²⁷ and is the foundation of feedback intervention theory³⁴ Recent empirical data points to data feedback as central to success of several key QI projects^{4,30,31,37} Cochrane reviews on data feedback indicate a positive impact on QI if feedback is appropriate and timely, and when a path to improvement is proposed^{34,38}
Front-line teams develop and use basic QI skills to effect change	QI leads and other team members: <ul style="list-style-type: none"> use time series charts ('run charts') (QI strategy 4) segment the patient pathway (QI strategy 5) Use the PDSA cycles (QI strategy 6) 	<ul style="list-style-type: none"> Introduction to QI skills training provided Links to further reading and training resources for QI Telephone and e-mail support 	<ul style="list-style-type: none"> Application of improvement science approaches, such as the model for improvement, require at least some basic skill acquisition; evidence points to a deficit in this area, putting significant strain on the ability of an improvement project to achieve its potential^{39,40} Time series charts ('run charts') are a simple and robust method of analysing and presenting (for data feedback) changes to care processes⁴¹ Segmentation of the proposed patient pathway involves introducing interventions to 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 trialable and lower risk⁴² The IHI model for improvement, including the PDSA cycle, is an internationally accepted approach to QI^{27,43}

IHI, Institute for Healthcare Improvement; PDSA, plan-do-study-act.

Reproduced from Stephens *et al.*²² This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated. The table includes minor additions and formatting changes to the original table.

Discussion

We designed a complex, evidence-based intervention that we hypothesised would improve the care pathway, and ultimately reduce mortality, for this patient group. Like many complex interventions, it was designed to operate on multiple levels, was to be adapted to fit local contexts, required behaviour change by a range of different actors and needed varying levels of skill to achieve this, in both the recipient and the delivery teams.²³ The intervention drew on current evidence about 'what works' in QI and was also designed 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 e-mail/telephone support. The hard core of the hospital-level intervention, the care pathway of recommendations, was more extensive and was based largely on current Department of Health and Social Care and Royal College of Surgeons guidance.⁸ The soft periphery of the hospital-level intervention, the QI intervention, was conceived with the non-QI expert clinician in mind and designed to be as easy as possible to use. As such, we viewed the EPOCH trial programme and interventions, overall, as enabling clinicians to implement guidance that was in line with what already existed, but that had not been put into practice nationally. Several observers have argued that this pathway is not 'evidence based'. We believe this relates to a misunderstanding of what evidence-based medicine is. The EPOCH trial pathway 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.⁴⁴ Strengths 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 among clinical teams.³¹ The multimodal training resources (face to face, online and by telephone and e-mail) also meant that the training to support use of the intervention was highly accessible. Both the clinical and QI interventions were also based on best-available evidence and the rationale underpinning the QI intervention, the programme theory, focused and guided the intervention design. However, although the trial interventions drew on learning and evidence from similar, smaller-scale work, the primary limitation was the lack of pilot trials of this particular set of interventions, in this particular context, to test both the efficacy of the pathway, under stricter trial conditions, and the feasibility of using the recommended QI intervention to implement the pathway in a more pragmatic pilot trial setting.

Chapter 3 Stepped-wedge cluster randomised trial

Introduction

Emergency abdominal surgery is associated with poor postoperative outcomes. Around 30,000 patients undergo this type of surgery each year in the UK NHS, with 30-day mortality rates in excess of 10% and wide variation in standards of care between hospitals. Several groups have studied the effect of QI initiatives to implement individual interventions or 'care bundles' of several treatments, and so improve care for these patients. Overall, the findings of these small studies suggest survival benefit, but most utilised uncontrolled cohort designs, which are associated with a high risk of bias. The feasibility and benefit of a national QI programme to implement a more extensive acute care pathway for this patient group remain uncertain. We conducted a stepped-wedge cluster randomised trial, with an embedded ethnographic evaluation, to evaluate the effect of implementing this pathway on survival following emergency abdominal surgery in NHS hospitals.

Methods

Trial design and participants

The EPOCH trial was a multicentre, stepped-wedge cluster randomised trial of a QI 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 (URL: www.epochtrial.org/protocol; accessed 18 July 2019). The trial was prospectively registered at isrctn.com on 27 February 2014, but a registration number was not issued until 7 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 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 that 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 aged ≥ 40 years and undergoing emergency open abdominal surgery in a participating hospital during the 85-week trial period (from 3 March 2014 to 19 October 2015). Patients were excluded from the analysis if they were undergoing a simple appendicectomy, 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.

Data collection

Trial data were collected through the NELA database (URL: www.nela.org.uk; accessed 6 September 2019) and then linked using unique patient identifiers to Hospital Episode Statistics (HES) and Office for National Statistics (ONS) in England and Wales, and the Information Services Division of NHS Scotland, to provide data describing mortality and hospital readmissions. The trial was approved by the East Midlands (Nottingham 1) Research Ethics Committee (reference number 13/EM/0415). Data were analysed without individual patient consent in accordance with section 251 of the National Health Services Act 2006.⁴⁵

Randomisation and masking

We planned to include 15 geographical clusters of five to seven hospitals. The QI intervention lasted 85 weeks, with one geographical cluster commencing the intervention each 5-week step from the 2nd to the 16th time period. Clusters were randomly assigned to 1 of 15 start dates for the QI intervention by an independent statistician, using a computer-generated random allocation sequence. As each geographical

area started in the usual care group and ended in the QI group, there were 17 time periods in total. Local investigators in each geographical area were notified 12 weeks in advance of activation of the QI programme at their hospital. As they were engaged in delivery of the intervention, it was not possible to mask hospital staff. Patients were masked to trial group allocation. The organisation of hospitals into geographical clusters minimised any contamination between sites due to natural workforce movements between hospitals.

Trial intervention

The EPOCH trial care pathway and QI methodology are described in *Chapter 2*.

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 readmission within 180 days of surgery. We selected 10 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) preoperative 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 (< 2 hours, 2–6 hours, 6–18 hours, or > 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.

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 and Hughes approach,⁴⁶ for an analysis with fixed time effects and random cluster effects, modified to exclude data collected during the 5-week period in which the intervention commenced in individual clusters. Using HES 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/5 weeks/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.⁴⁷ Patients who presented during the 5-week time period immediately after QI 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 that withdrew after the trial start date, or did not implement the intervention, were included in the analysis. Hospitals that 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%) of whom were censored at hospital discharge. All analyses included time period as a fixed effect using indicator variables, and were adjusted for age, sex and indication for surgery using fixed factors.⁴⁸ Age was included as a continuous covariate, assuming a linear association with outcome.⁴⁹ Missing baseline data for indication for surgery were handled using a missing indicator approach.⁵⁰ 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 (i.e. the time period in hospital). This allowed additional correlation between patients in the same hospital and the same period, compared with patients in other periods, as is recommended.^{51,52}

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 (HR) from this analysis measures the relative probability of hospital discharge between treatment arms, with a HR of < 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 of < 1 indicating a lower probability of readmission).

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 5-week period immediately after implementation of the intervention. We analysed patients according to the following four groups: (1) no QI implemented (usual care group); (2) QI implemented for < 5 weeks; (3) QI implemented for between 5 and 10 weeks; and (4) QI implemented for > 10 weeks. Our second analysis evaluated the intervention in other patient populations that 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. Owing to the small number of patients in this group, we summarised results descriptively rather than undertaking a formal statistical analysis.

Patient and public involvement

We actively involved patient and public representatives throughout this project. Our co-applicants on the original funding application included two patient representatives, one of whom has lived experience of emergency abdominal surgery. They have supported us in the oversight and conduct of the trial, advising on numerous issues from the preparation of patient-facing materials to advising on the patient perspective of registry data use without individual patient consent. Both are named authors of this report and have specifically helped in drafting the *Plain English summary*.

Results

Fifteen geographic areas underwent randomisation, including 97 NHS hospitals. Four hospitals withdrew before the start of the trial, leaving 93 hospitals participating. Between 3 March 2014 and 19 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 2* and *Table 2*). Baseline characteristics were similar between groups (*Table 3*).

Process measures

Ninety-one out of 93 (98%) hospitals were represented at the initial QI meeting for the relevant geographical cluster and 53 out of 93 (57%) hospitals were represented at the follow-up QI meeting. This representation included a named hospital QI lead for 89 out of 93 (96%) hospitals at the first meeting and 47 out of 93 (51%) hospitals at the second meeting. Most meetings (13/15) occurred within 2 weeks of the activation date. Patient-level process measures are described in *Table 4*. In accordance with our analysis plan, we did not test these for statistical significance.

Clinical outcomes

Complete primary outcome data were available for $> 99\%$ patients (see *Figure 2*). The primary outcome of 90-day mortality occurred in 1393 usual care group patients (16%) compared with 1210 QI group patients (16%) [QI vs. usual care HR 1.11, 95% confidence interval (CI) 0.96 to 1.28] (*Figure 3* and *Table 5*). Results were similar for 180-day mortality (HR 1.12, 95% CI 0.98 to 1.28) (*Figure 4*). Patients in the QI group had a lower probability of hospital discharge (HR for hospital discharge 0.90, 95% CI 0.83 to 0.97), leading to a marginally longer hospital stay {days in hospital: usual care 13 [interquartile range (IQR) 8–23] days vs. QI 13 (IQR 8–24) days}, although this difference was not clinically meaningful (*Figure 5* and *Table 6*). There was no difference between groups in hospital readmission within 180 days [usual care, $n = 1618$ (20%) vs. QI, $n = 1242$ (18%), HR for readmission 0.87, 95% CI 0.73 to 1.04] (*Figure 6* and *Table 7*). In a secondary

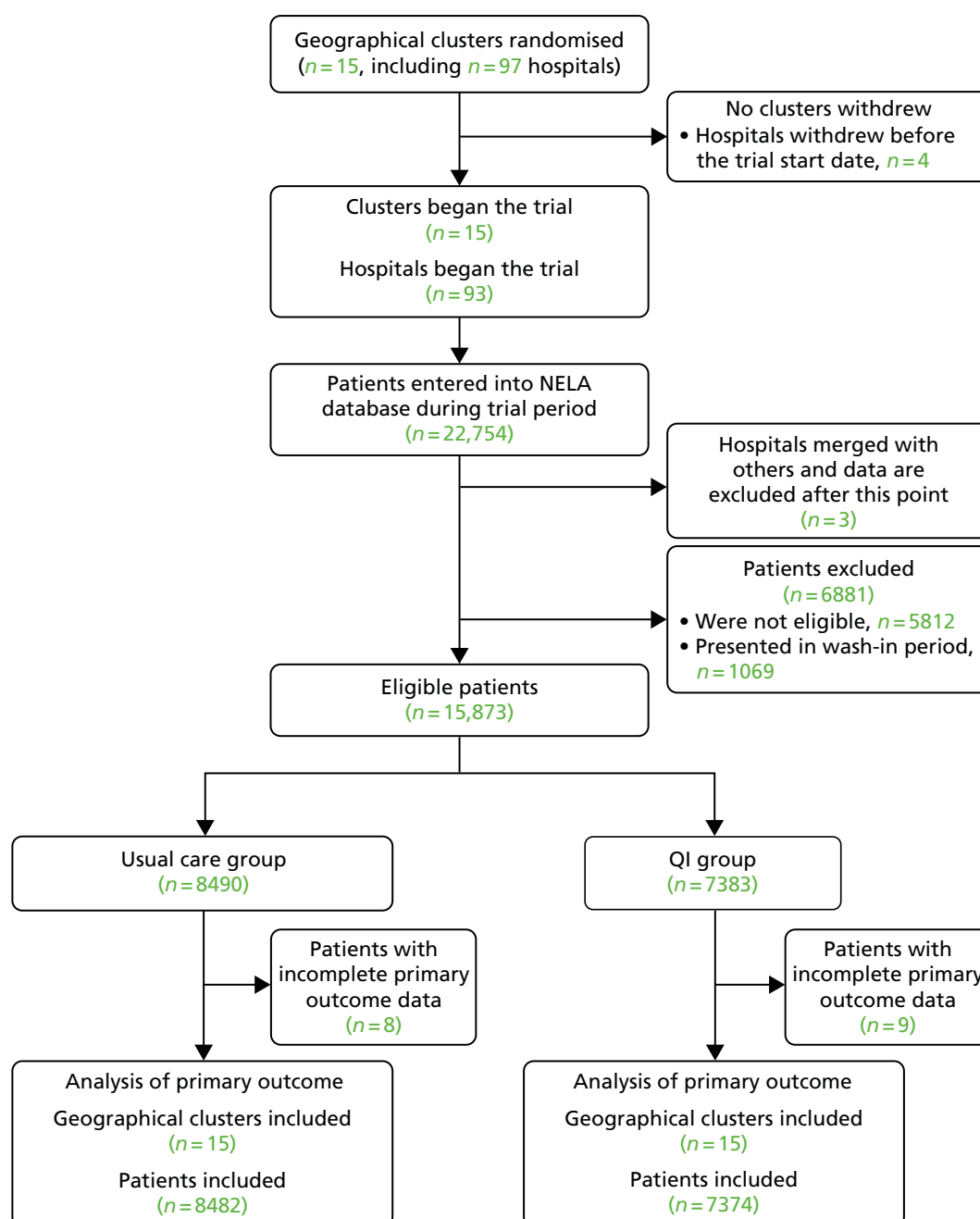


FIGURE 2 Inclusion of hospitals and patients in the trial.

analysis, we found no evidence that the QI strategy became more effective the longer it had been adopted (Table 8). To assess the impact of missing mortality data following hospital discharge from patients in Wales, we assessed the number of mortality events that 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 that few outcome events in Wales were missed. Analysis of the effect of the intervention over time is presented in Table 8 and of the inclusion of younger patients and those undergoing laparoscopic surgery in Table 9.

TABLE 2 Number of patients recruited in each cluster in each period

Geographical area (cluster)	Period																	Total
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
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	15,873
Notes Shaded cells denote periods after intervention implementation. Cells with '–' denote the 5-week period immediately after intervention implementation when participants were excluded. 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/Surrey/Sussex; (13) Wessex; (14) Scotland; and (15) West Midlands.																		

TABLE 3 Baseline patient characteristics

Characteristic	Number of patients with missing data		Summary measure	
	Usual care (N = 8490)	QI (N = 7383)	Usual care	QI
Baseline characteristics				
Female, n (%)	0 (0)	0 (0)	4550 (54)	3938 (53)
Age (years), mean (SD)	0 (0)	0 (0)	68 (13)	68 (13)
Indication for surgery, n (%)	13 (< 1)	5 (< 1)		
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)
Preoperative characteristics				
Estimated risk of death, n (%)	158 (2)	22 (< 1)		
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, n (%)	156 (2)	23 (< 1)		
I (no systemic disease)			615 (7)	533 (7)
II (mild systemic disease)			2815 (34)	2461 (33)
III (severe systemic disease)			3112 (37)	2745 (37)
IV (life-threatening systemic disease)			1605 (19)	1465 (20)
V (moribund patient)			187 (2)	156 (2)
P-POSSUM score, median (IQR)	152 (2)	13 (< 1)	7.6 (2.9–22.7)	7.4 (2.8–22.9)
Systolic blood pressure (mmHg), mean (SD)	255 (3)	147 (2)	128 (24)	128 (25)
Glasgow Coma Score, mean (SD)	221 (3)	72 (1)	14.8 (1.4)	14.7 (1.5)
Blood lactate (mmol/l), median (IQR)	4103 (48)	2870 (39)	1.6 (1.1–2.8)	1.5 (1.0–2.6)
ASA, American Society of Anaesthetists; IQR, interquartile range; P-POSSUM, Portsmouth Physiological and Operative Severity Score for the enumeration of Mortality and morbidity; SD, standard deviation.				

TABLE 4 Patient-level process measures

Process measure	Number of patients with missing data		Summary measure	
	Usual care (<i>N</i> = 8490), <i>n</i> (%)	QI (<i>N</i> = 7383), <i>n</i> (%)	Usual care, <i>n</i> (%) or median (IQR)	QI, <i>n</i> (%) or median (IQR)
Consultant decision to operate	184 (2)	72 (1)	7472 (90)	6589 (90)
Consultant reviewed patient at time of decision	448 (6)	334 (5)	5961 (85)	5271 (84)
Preoperative documentation of risk	158 (2)	22 (< 1)	4570 (55)	4893 (66)
Patient entered operating theatre within specified urgency time frame	1012 (12)	430 (6)	5636 (75)	5515 (79)
Consultant surgeon present in operating theatre	155 (2)	17 (< 1)	7117 (85)	6472 (88)
Consultant anaesthetist present in operating theatre	160 (2)	14 (< 1)	6313 (76)	5832 (79)
Goal-directed fluid therapy used during surgery	180 (2)	24 (< 1)	3942 (47)	4329 (59)
Serum lactate measured at end of surgery	171 (2)	24 (< 1)	4474 (54)	4431 (60)
Time (hours) from decision to operate to entry into operating theatre	630 (7)	417 (6)	5.0 (2.1–16.8)	4.3 (2.0–15.3)
Critical care admission immediately after surgery ^a	163 (2)	22 (< 1)	5395 (65)	5050 (69)

IQR, interquartile range.

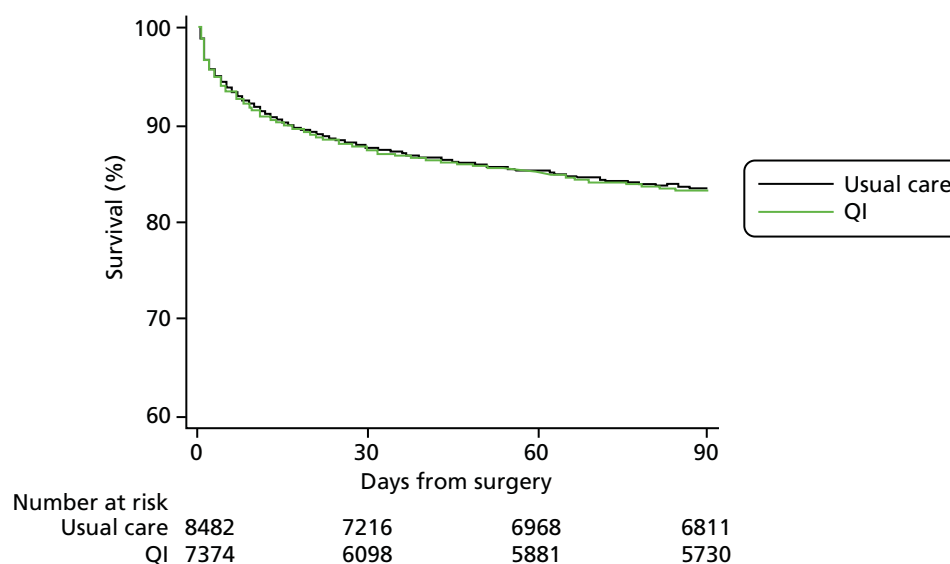
^a 29 patients in the usual care group and 27 patients in the QI group died during surgery.**FIGURE 3** Mortality within 90 days of emergency abdominal surgery.

TABLE 5 Patient outcomes

Outcome	Number of patients included in analysis		Summary outcome measure		
	Usual care (N = 8490), n (%)	QI (N = 7383), n (%)	Usual care, n (%) or median IQR	QI, n (%) or median IQR	HR (95% CI) (QI vs. usual care)
All-cause mortality within 90 days of surgery	8482 (> 99)	7374 (> 99)	1393 (16)	1210 (16)	1.11 (0.96 to 1.28)
All-cause mortality within 180 days of surgery	8482 (> 99)	7374 (> 99)	1698 (20)	1440 (20)	1.12 (0.98 to 1.28)
Duration of hospital stay (days)	8320 (98)	7353 (> 99)	8 (13–23)	8 (13–24)	0.90 (0.83 to 0.97)
Hospital readmission within 180 days of surgery	7969 (94)	6723 (91)	1618 (20)	1242 (18)	0.87 (0.73 to 1.04)

IQR, interquartile range.

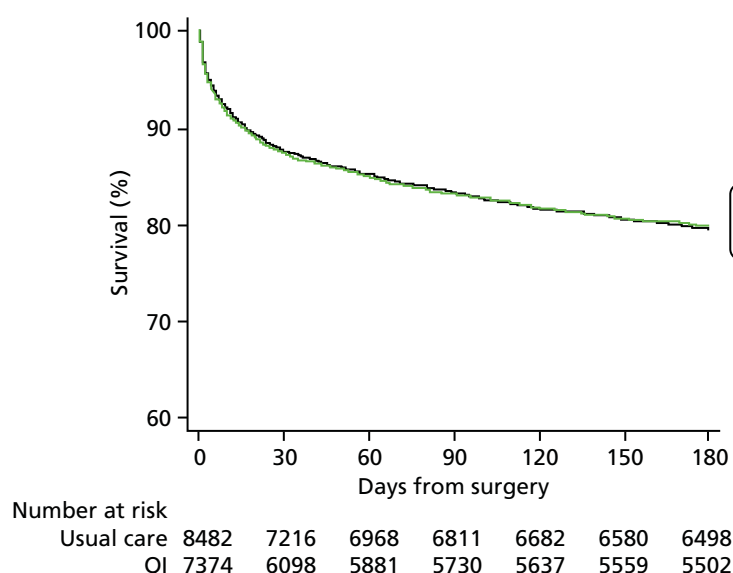


FIGURE 4 Mortality within 180 days.

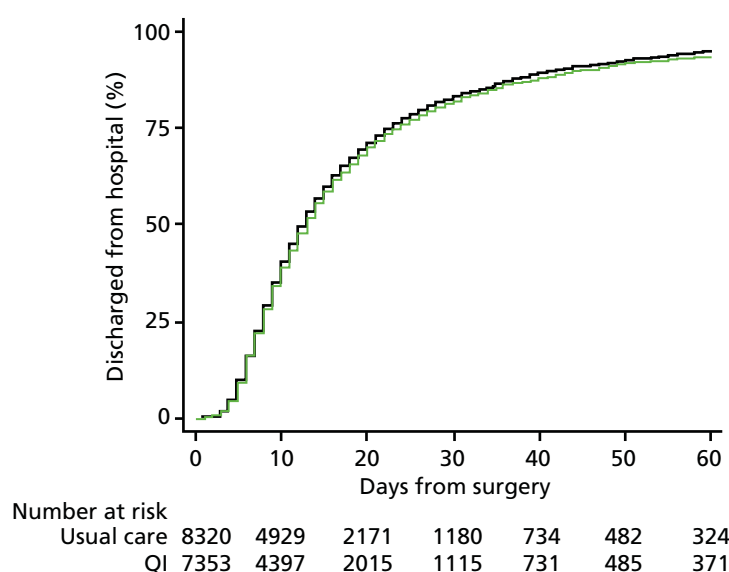
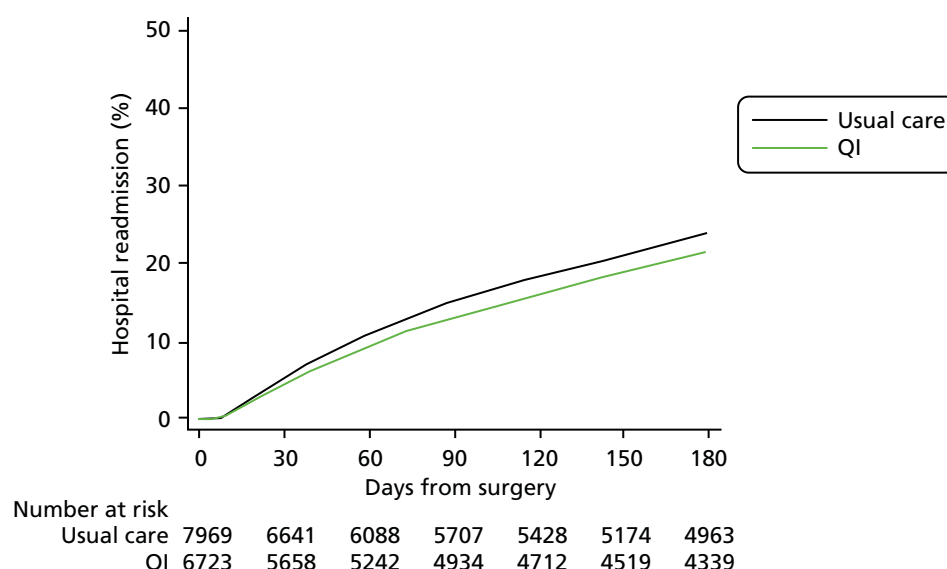


FIGURE 5 Duration of hospital stay after emergency abdominal surgery.

TABLE 6 Number of hospitals included in each analysis

Outcome	Hospitals (<i>N</i> = 93), <i>n</i> (%)
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 readmission within 180 days of surgery	87 (94)

**FIGURE 6** Time to hospital readmission.**TABLE 7** Summary statistics for duration of hospital stay and hospital readmission

Outcome	Usual care, <i>n</i> (%)	QI, <i>n</i> (%)
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 readmission 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 8 Summary effect of QI intervention over time

Duration	90-day mortality, <i>n/N</i> (%)	HR (95% CI)	<i>p</i> -value (overall)
No QI	1393/8482 (16)	Reference	0.15
QI for < 5 weeks	198/1069 (19)	1.11 (0.94 to 1.32)	
QI for 5–10 weeks	185/983 (19)	1.21 (1.01 to 1.44)	
QI for > 10 weeks	1025/6391 (16)	1.05 (0.90 to 1.23)	

TABLE 9 90-day mortality in patients who underwent laparoscopic surgery or were aged 18–40 years and met all other inclusion criteria

Outcome	No QI (<i>N</i> = 1503), <i>n</i> (%)	QI (<i>N</i> = 1459), <i>n</i> (%)
90-day mortality	101 (7)	66 (5)

Discussion

The principal finding of this trial was that there was no survival benefit associated with a national QI 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 among the 10 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.

The strengths of this trial include wide generalisability (a large number of consecutive patients were enrolled by many hospitals), robust trial design and the devolved leadership to local clinical QI teams. The evidence-based EPOCH trial care pathway was developed through a Delphi consensus process to update national professional guidelines.⁸ Partnership with the 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 HES 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% CI for our primary effect estimate was narrow, with a lower limit that indicates a maximum potential relative mortality reduction of 4%. Our findings are unlikely to change with a larger sample size. Because of the difficulty in obtaining post-discharge survival data in Wales, we changed our primary analysis from a binary to a time-to-event approach, allowing the inclusion of mortality events censored at hospital discharge. However, post-discharge data from England and Scotland suggest that few events were missed as a result of this approach.

Chapter 4 Process evaluation

Introduction

Quality improvement interventions, such as that delivered in the EPOCH trial, are complex due to their interacting components, and the multiple organisational and social levels at which they operate.⁵³ 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 in a trial setting, such complexity may mean that the target group is not actually exposed to the intervention as planned.²⁴ 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 (the TIDieR checklist²⁴ and the SQUIRE guidelines²⁵), but such detailed reporting is not common in the QI literature.⁵⁴

In this chapter we focus on the process evaluation data to describe how one of the largest trials of a QI intervention to date was delivered and received across 93 hospitals that offer emergency abdominal surgery in the NHS, and provide detailed analysis to facilitate a greater understanding of the main trial results. The ethnographic evaluation is described in *Chapter 5*.

Methods

We undertook a mixed-methods process evaluation based on recommended guidance for the evaluation of cluster trials,⁵⁵ which was 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 trial QI leads).

Data sources and data collection

Table 10 details the evaluation foci and the data sources used to investigate each. Following commencement of the trial, the variability of engagement with the QI 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, we collected a range of QI programme activity data (see *Table 10*) 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. Only one response was required per hospital, but QI leads were asked to complete the questionnaire with colleagues. QI programme data were collected by the programme co-ordinator (TS) and included data on participation in programme activities, such as meetings and use of the trial VLE.

Data analysis

The programme activity and questionnaire data were analysed and reported using descriptive statistics [frequency (%) for categorical data or median (range) for continuous data]. Free-text data in the exit questionnaire were coded by two investigators, using both inductive and deductive content analysis techniques. Findings were discussed until themes were agreed and draft results were discussed with the ethnographic researchers to enhance external validity.⁵⁶

TABLE 10 Data collected for process evaluation

Aspect of process evaluation	Data collection method	Data collected and data type
Recruitment of sites	Review of trial administrative records	Recruitment strategy, including inclusion and exclusion criteria Reasons given for non-participation (text in trial administrative documents)
Delivery to the clusters	Collation of registers from QI programme meetings (30 meetings) Collation of VLE usage logs	Names, roles and hospital of each of the attendees at the QI programme meetings (two meetings/cluster) The level of usage of the 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 or not a stakeholder meeting was held (QI strategy 1) Whether or not a QI team was formed and professional composition of any such team (QI strategy 2) If and how data feedback occurred (QI strategy 3) Whether or not run charts were used (QI strategy 4) Whether or not the patient pathway was segmented (QI strategy 5) Whether or not the PDSA approach was used (QI strategy 6)
Response of sites/ individuals	Online exit questionnaire	Free-text responses to two reflective questions: if you were to be involved in EPOCH trial again, (1) 'what would you continue doing' and (2) 'what would you do differently'?

Results

Programme activity data, as defined in *Table 10*, were available for all 93 hospitals. Eighty-three per cent (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 out of 77 (22%) responses included surgical input and 6 out of 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 trial QI leads).

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 ongoing improvement work focused on emergency laparotomy.

We have documented records of 14 hospitals expressing interest, but subsequently not participating due to existing improvement work in that site. For other hospitals that 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.

Delivery of the intervention at the cluster level

A total of 15 face-to-face QI educational meetings, planned to coincide with cluster activation and 15 follow-up meetings (one for each geographical cluster) were held as part of the QI programme. *Figure 7* summarises the EPOCH trial QI 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 8* 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 in a VLE, 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 trial. The site could be accessed only 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 VLE hits per hospital

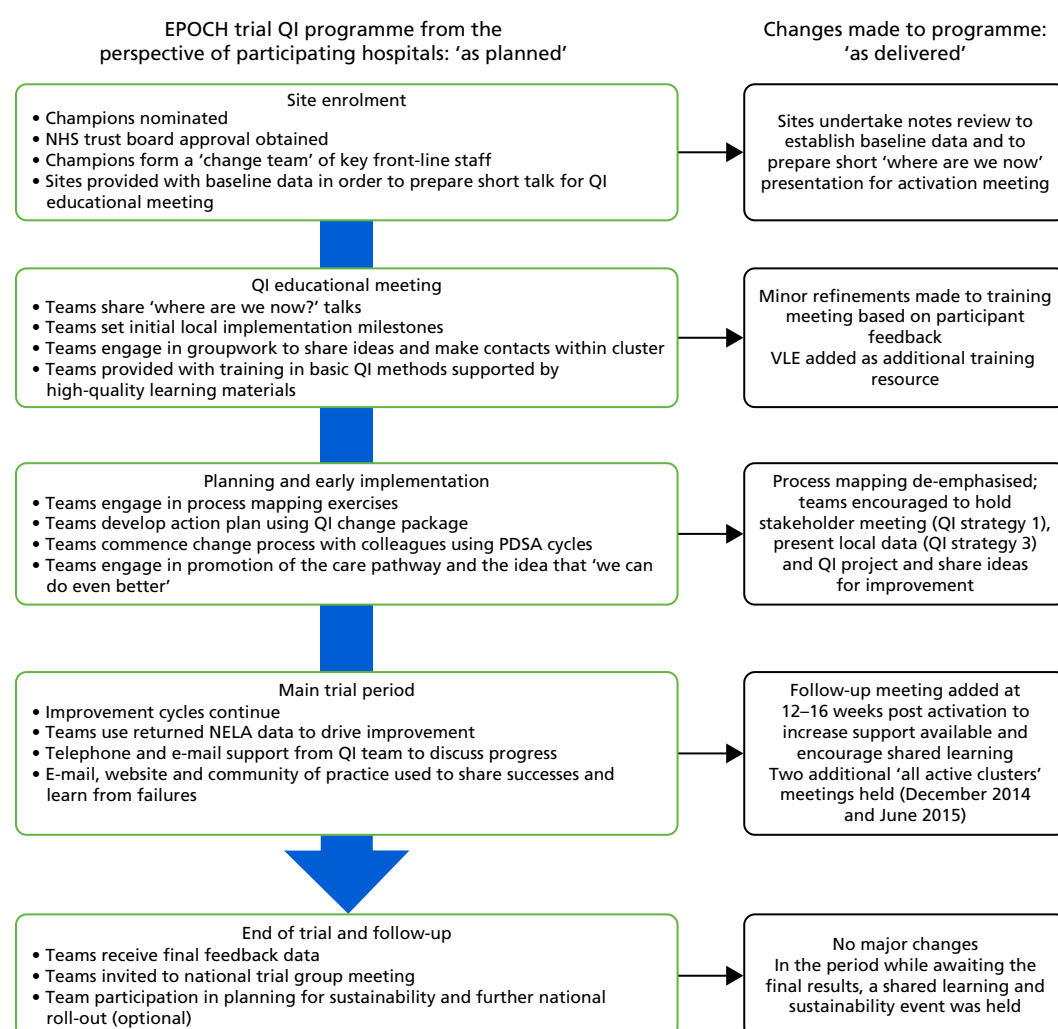


FIGURE 7 The QI programme as planned and as delivered. Reproduced from Stephens *et al.*²² This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated. The figure includes minor additions and formatting changes to the original figure.

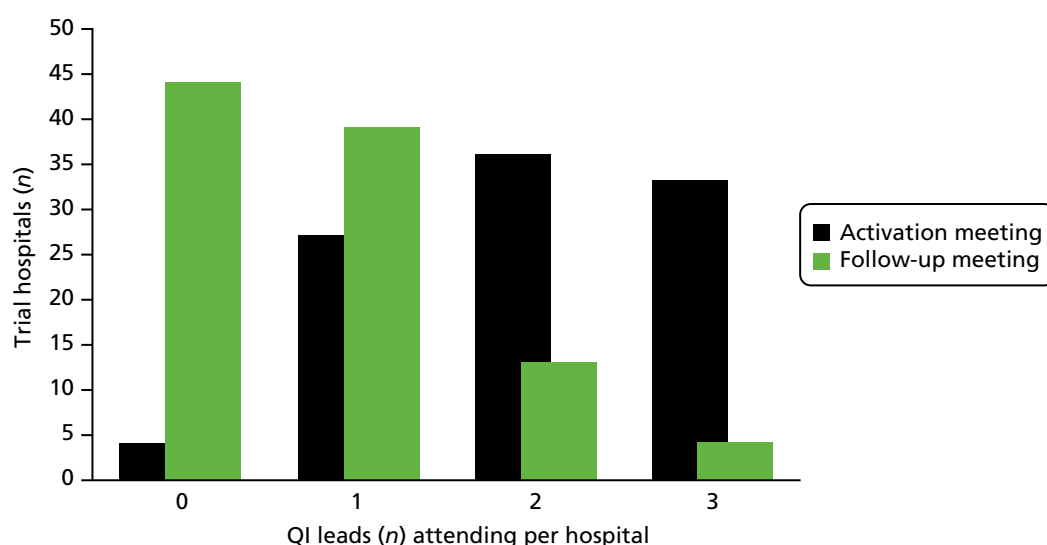


FIGURE 8 Quality improvement lead attendance at QI meetings.

was 136 (minimum 11, maximum 519, interquartile range 70–194). The number of users per hospital ranged from one to seven, with a median of three users, but as some site teams were small, and as online resources could be downloaded and/or printed by one user for colleagues to view, we feel the total hits per hospital is a more useful metric of VLE usage. Given the number of pages/resources available (84 by the trial end), these data suggest probable appropriate usage by much of the cohort, but with some variability and a significant minority of low users.

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 11*. 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 in, the intervention, and more meetings and input together with more time in the intervention period.

TABLE 11 Common themes identified from feedback regarding the QI programme

'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 (<i>n</i> = 14)	More clarity about the intervention and how to implement it (<i>n</i> = 10)
Networking with colleagues from other hospitals (facilitated by meetings) (<i>n</i> = 11)	More meetings and more input from the central team (<i>n</i> = 8)
Good communication and support (<i>n</i> = 12)	Better support/better run chart tool (<i>n</i> = 7)
The Excel tool to generate run charts from NELA data (<i>n</i> = 11)	A longer intervention period for those activated late (due to the stepped-wedge trial design) (<i>n</i> = 7)
Enthusiasm and motivation generated by the EPOCH trial team and project overall (<i>n</i> = 8)	Fewer components in the clinical pathway (<i>n</i> = 4)

Reproduced from Stephens *et al.*²² This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated. The table includes minor additions and formatting changes to the original table.

Quality improvement leads were also asked to rate the support they received from the QI programme team on a 5-point scale (very good to very poor). Of the 75 who responded to this question, 36 out of 75 (48%) QI leads rated support as very good, 30 out of 75 (40%) QI leads rated support as good and 9 out of 75 (12%) QI leads rated support as average.

Delivery to individuals at local site (quality improvement intervention fidelity)

The clinical intervention was a 37-component care pathway (see *Box 1*). Questionnaire data showed that, regarding the care pathway, only 11 care processes were the focus of improvement efforts in > 50% of responding hospitals; the remaining pathway components had more variable uptake (see *Segmenting the pathway* and *Figure 9*). The QI intervention comprised six strategies (see *Figure 1* and *Table 1*). Questionnaire data showed that 10 out of 77 (13%) QI leads responding said that all six strategies had been used, 23 out of 77 (30%) QI leads indicated five had been used, 21 out of 77 (27%) QI leads indicated four had been used, 8 out of 77 (10%) QI leads had used three strategies, 10 out of 77 (13%) QI leads had used two and 5 out of 77 (6%) QI leads had used just one. No QI lead reported zero QI strategy usage. *Table 12* shows the reported usage of each QI strategy and each is discussed briefly below.

TABLE 12 Reported usage of each QI strategy

Question related to QI strategy usage	Response (<i>n</i> = variable)
PDSA approach	
Did you or your colleagues use the PDSA cycle approach during your QI activities?	<ul style="list-style-type: none"> • Yes, sometimes: 61% (45/74) • Yes, often: 5% (4/74) • No: 34% (25/74)
QI team formation	
At your site, was a formal team created to work on QI activities related to the EPOCH trial? 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	<ul style="list-style-type: none"> • Yes: 60% (46/77) • No: 27% (21/77) • Other (comments indicated informal teams often existed): 13% (10/77)
Data collection and analysis	
After starting the EPOCH trial, did you or your colleagues download and analyse your local NELA data?	<ul style="list-style-type: none"> • Yes: 79% (61/77) • No: 21% (16/77)
If yes, how frequently did you do this?	<ul style="list-style-type: none"> • Analysing data monthly or bi-monthly: 43% (26/61) • Analysing data less frequently: 57% (35/61)
If yes, did you use run charts?	<ul style="list-style-type: none"> • Used run charts to analyse data: 92% (56/61)
Were systems set up to collect NELA data prospectively?	<ul style="list-style-type: none"> • Yes: 51% (38/74) • No: 49% (36/74)
Stakeholder meeting	
Did you hold a stakeholder meeting as one of your QI activities? For example, a meeting for all professionals involved in patient care	<ul style="list-style-type: none"> • Yes: 55% (41/75) • No: 45% (34/75)
Pathway segmentation	
Please indicate the statement that most closely fits your hospital's improvement or implementation activity during the EPOCH trial	<ul style="list-style-type: none"> • We introduced a single pathway of care (across pre-, intra- and postoperative phases): 22% (17/77) • We introduced separate pathways or care bundles for the perioperative phases: 32% (25/77) • We focused on introducing individual/separate interventions: 40% (31/77) • Other: 5% (4/77)

Reproduced from Stephens *et al.*²² This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated. The table includes minor additions and formatting changes to the original table.

Use of plan–do–study–act

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 a clear set of instructions and supporting tools for putting it into practice. The data in *Table 12* indicate that this approach was used, but perhaps not in the regular, methodical manner recommended.

Team approach

At the activation meetings, QI leads were strongly advised to recruit a formal team of ‘willing’ interprofessional colleagues to work with them on the local improvement activities. The data in *Table 12* indicates that just under two-thirds of sites had a formal team to work on this major project.

Use of data feedback and run charts

At the activation meetings, using NELA data as a driver for engaging colleagues and monitoring improvement was promoted and tools provided to do so. The data in *Table 12* shows that most, but not all, teams were analysing NELA data; far fewer were doing this on a regular (monthly/bi-monthly) basis. Many sites reported challenges in simply collecting the data, with only half of all questionnaire respondents indicating that systems had been set up to collect the audit data prospectively.

Engagement

At 5 weeks before activation to the intervention, sites were contacted and recommended to start planning a stakeholder meeting, to coincide with activation. Just over half of the respondents indicated that they had held such a meeting (see *Table 12*); when a meeting was held, most reported that it was successful (questionnaire data, not shown). The exit questionnaire also asked about senior support during the trial. Of the 71 who responded to this question, 15 (21%) described active executive board support for the QI work related to the EPOCH trial (e.g. funding staff time to support the project or making the project a board-level quality and safety priority).

Segmenting the pathway

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 step-wise implementation of the pathway that would be workable in QI leads’ local context. The data in *Table 12* suggests that the majority of sites appear to have followed this approach. However, the data in *Figure 9* also suggests that most sites did not progress beyond this to attempt to introduce most or all of the pathway components.

Response of quality improvement leads: reflections on the change process

Quality improvement leads reflected on ‘what would you continue doing?’ and ‘what would you do differently if you were to do the EPOCH trial again?’. Ninety-six per cent (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 (*Table 13*). Two clear themes, from responses to both questions, were related to the importance of effective engagement and involvement of colleagues (themes 2 and 6), and to data collection and feedback systems (themes 1 and 7). Other reflections on what QI leads ‘would continue doing’, related to QI methodology (themes 3 and 5) and the utility of specific, recommended clinical interventions, particularly mortality risk estimate scoring (theme 4). When considering ‘what they would do differently’, QI leads also highlighted some of the ‘real-life’ challenges of delivering QI at the frontline, developing leadership and project management skills (theme 9) and the need to obtain strong senior support (theme 8).

Discussion

The principal finding of this process evaluation was that the QI programme delivered the QI skills training and resources as intended and the programme was generally well received by QI leads. Local adaptation to both the QI and clinical interventions was actively encouraged, but the extent of variability and adaptation

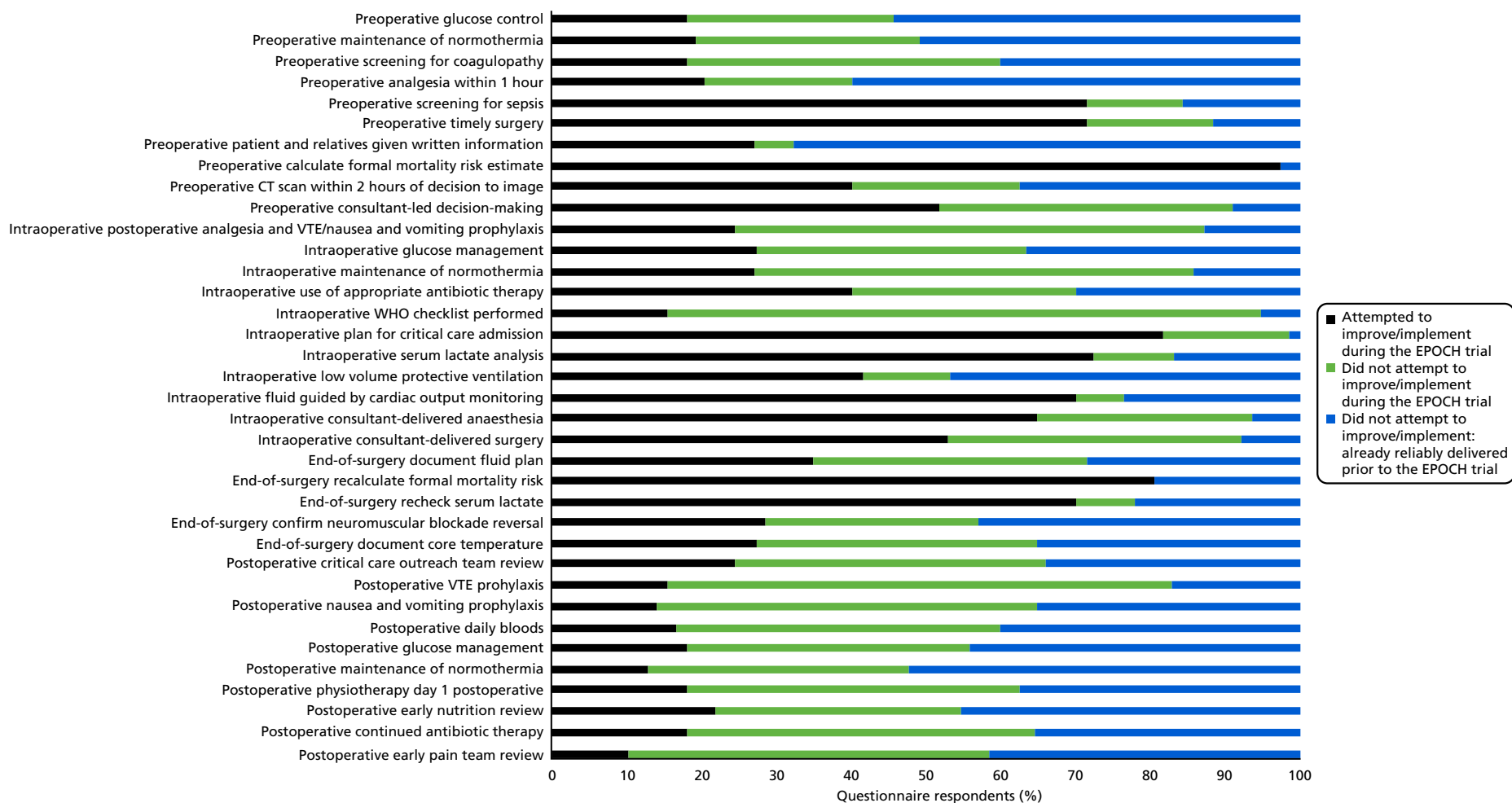


FIGURE 9 Clinical process change attempted during intervention period. CT, computerised tomography; VTE, venous thromboembolism; WHO, World Health Organization. Reproduced from Stephens *et al.*²² This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated. The figure includes minor additions and formatting changes to the original figure.

TABLE 13 Themes emerging from inductive content analysis of QI leads' responses

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

Reproduced from Stephens *et al.*²² This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated. The table includes minor additions and formatting changes to the original table.

in the implementation process was greater than anticipated, particularly in relation to the clinical processes. There were only 11 clinical processes that 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 (see *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 towards substantial improvement of care processes was seen across the whole cohort as a result of the EPOCH trial QI programme.

When testing clinical interventions in a clinical trial, it is important to make the distinction between the design of the intervention and the operational elements required for effective delivery.⁵⁷ Our 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 QI interventions can fit in different contexts, and this was built into the EPOCH trial intervention. However, fidelity to key parts of an intervention are also important to maximise the likelihood of success.⁵⁸ In this case, it may have been that an intervention design that focused on a smaller number of strategies 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 were also an operational challenge for many. NELA had commenced only 4 months before the start of the trial and, 20 months after the launch of NELA, at the end of this trial, only half of hospitals reported having prospective data collection systems in place. It is probable, therefore, that many QI leads were focused on collecting and inputting data to the detriment of other improvement activity. 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. Although data are central to any QI project, it is the use of this data through feedback, combined with other improvement strategies, that is likely to achieve more robust results.^{30,31,38} 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. Organisational challenges may have meant that many teams simply ran out of time to implement the pathway in the intervention period. Earlier, smaller studies have shown that marked improvement may take time and can continue after the intervention period.⁵⁹

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 improbable, given the evidence base for the clinical and QI 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 'socioadaptive' interventions can be low among health professionals.³² However, a key reflection of QI leads was that they would have liked to have spent more time engaging and involving colleagues. We would suggest, therefore, that more emphasis and training in socioadaptive interventions should be built into future programmes, together with a recognition that dedicated time is required to support front-line staff in prioritising such interventions.³⁹ Some leads reflected on their difficulties in engaging with senior or executive-level colleagues and only one-fifth of respondents indicated that they received active support from their board. Effective QI 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.⁶⁰ This is an important lesson; if the goodwill and motivation of front-line 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 QI interventions, such as the Institute for Healthcare Improvement breakthrough series collaborative model.¹³ 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.^{13,61} The EPOCH trial may have suffered from the lack of a pilot trial and perhaps future similar interventions should be piloted first,²³ or use a cluster trial design that allows for iterative intervention development within the trial period to enable ongoing intervention optimisation.⁶²

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 UK NHS. Following calls for better intervention reporting, we hope that we have provided insights into possible reasons why, ultimately, the trial was unsuccessful and learning for future studies of this nature.^{24,25} The evaluation was conducted by researchers both inside and outside the main trial team, offering both detailed, nuanced knowledge of the trial, with 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 trial programme, it is possible that some relevant factors are missing. Self-reported data may be subject to both recall and/or social desirability bias. To minimise recall bias, we started collecting data within 1 month of the completion of the trial. Although we cannot quantify the magnitude of potential social desirability bias, many respondents reported both positive and negative experiences and many reported not using several of the QI strategies.

Conclusions

Programmes designed to support clinician-led improvement may need to focus on developing the necessary QI capabilities, while 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 probable prerequisites for the successful delivery of complex interventions, such as implementing a care pathway for emergency surgery.

Chapter 5 Ethnographic evaluation

Introduction

A notable lesson from the last 15 years of efforts to improve the quality of health care in the UK and elsewhere is that the process is rarely a straightforward one.^{63,64} Studies worldwide have often found that even the most well-founded improvement interventions result in mixed or negative results, and that interventions that seem to work in one setting flounder when introduced in another.^{14,15,32} It is increasingly recognised that improvement is a complex endeavour, in which multiple variables interact in difficult-to-predict ways, such that improving health care is rarely a matter of implementing a single, simple intervention with a direct, causal impact on outcomes. Rather, successful improvement requires a complement of technical and social components, as it involves not only introducing proven clinical processes, but also ensuring that these are engaged in and adhered to reliably by the range of practitioners involved.³⁰ This also implies that context may be important, as although technical clinical interventions may be universally effective, social interventions may work better in some social, professional and organisational contexts than others, and therefore require adaptation in implementation.

The design of the EPOCH trial programmes explicitly acknowledged these challenges. The care pathway itself, comprising multiple technical interventions with an established association with improved outcomes, was accompanied by a set of social interventions designed to help secure and sustain change in the participating sites. Moreover, this set of social interventions was informed by a well-established approach to structuring health-care improvement, the QI collaborative,⁶⁵ and articulated in an explicit 'programme theory' that explained the key interventions to be undertaken, the theory underlying them and the logic by which they were expected to result in changed practice. The existence and articulation of such a programme theory is crucial to its evaluability: it shows the components of a programme and how they are expected to give rise to the intended outcomes, and therefore provides evaluators with an object of inquiry, for which the implementation and impact can be studied.^{66,67}

The EPOCH's programme theory, and the complex of interventions it involved, are described in detail in *Chapter 2*. The range of activities involved in implementing the EPOCH trial, from the activation of sites, through the ongoing support by the national EPOCH trial team, to the joint meetings to review and reinvigorate progress, were premised on the importance of four 'pillars': (1) the convening of a multidisciplinary team covering the three key specialties involved in the pathway; (2) the engagement of others around the team whose co-operation or active involvement would also be crucial to engendering change; (3) the use of small cycles of change (PDSA cycles) to develop, implement, test and refine improvements to the care pathway; and (4) the use of data from the NELA and from local data collection exercises to monitor progress, focus efforts and make the case for improvement. Accordingly, our analysis in this chapter – which draws on a qualitative substudy of the realisation of the EPOCH trial in a theoretically informed sample of six of the participating sites, recruited at different points in the course of the trial – focuses on these pillars, and on how they were understood and enacted in each site.

On a superficial level at least, all six sites adopted very similar approach, which was faithful to the EPOCH trial programme theory. They all used data, built a QI team and worked on engaging colleagues and other actors across organisations, and they all deployed small cycles of change. A closer analysis, however, shows notable differences, both among the sites (e.g. in terms of the composition of the teams involved), and between the approach advocated by the EPOCH trial team and that put into practice by the participating sites (e.g. in the way in which PDSA cycles were undertaken). Moreover, the organisational context of each site influenced both the approach taken by the local leads and the degree to which it appeared successful in engendering the kind of change intended. For example, some sites were better able

to enlist engagement from surgeons than others and ongoing organisational turbulence in some of the sites impeded or delayed efforts to introduce improvements to the care pathway.

The objective of this chapter is to provide an understanding of the activities undertaken under the auspices of the EPOCH trial in six of the participating sites, with a view to assessing the degree to which it followed the programme theory exactly, the nature of the adaptations made and the role of context in aiding or impeding progress. After accounting for our methods, we begin by drawing on observation of the activation meetings and initial interviews with local leads and other participants in the sites to discuss the reception, understanding and face validity of the EPOCH trial. We then consider the variable ways the four 'pillars' of the EPOCH trial were put into practice in the sites, before looking at each site individually, including the ways in which the diverse contexts appeared to affect implementation of the care pathway. Finally, we summarise and discuss our findings, and consider their implications for the EPOCH trial's approach to improvement and other, similar endeavours.

Methods

The qualitative study comprised two phases, with formative and summative objectives. The first phase was a telephone interview-based study, involving leads from hospitals (not included in the EPOCH trial) that had already made significant progress in introducing perioperative emergency laparotomy pathways similar to that to be implemented through the EPOCH trial. The objective here was to build on learning from these sites, to inform the development of the EPOCH trial approach, particularly the QI approach taken by the EPOCH trial team, and the interventions recommended to local leads. Phase 1 involved 10 interviews with such informants and resulted in an internal report that was used by the EPOCH trial QI team to hone their approach. The findings from phase 1 are not included in this chapter.

The second phase ran concurrently with the trial itself. It included ethnographic observation at seven regional activation meetings, and interviews and observations in six hospitals randomised to the intervention at various points in the trial, with a view to investigating the barriers to implementation and identify strategies to increase uptake of the care pathway. Findings from this phase were also fed back formatively to the QI team during the course of the trial to inform further finessing of their approach. However, the primary objective of this phase was to generate summative understanding of the 'fidelity' with which the care pathway has been implemented and the impact of this, and thereby offer insight into heterogeneity in outcomes between sites, understanding of aspects of the pathway that are more challenging to implement or sustain and recommendations for future similar interventions.

Sampling for the second phase of the qualitative study sought to secure variation in the sites included, with particular reference to a number of variables postulated as potentially important to the impact of the EPOCH trial at the hospital level: point of activation in the trial, status of hospital (teaching or district general hospital), professional affiliation of local lead, volume of annual surgical throughput and prior quality of care (as suggested by two indicators: summary hospital-level mortality indicator and mortality for surgery in general, based on figures from Dr Foster⁶⁸). (At the time of sampling, figures specific to emergency laparotomy were not available, as the first report from NELA had not yet been published.) The characteristics of the sites sampled according to these variables are presented in *Table 14*.

Of note, none of the included sites had a surgical lead as named local investigator (although in most cases a surgical lead was identified as part of the local team and often they were the named local leads for NELA). Surgical local investigators were in a minority across the trial and it did not prove possible to recruit a site that included one. This should be noted as a limitation of our trial, particularly in relation to its findings on engagement of different professional and disciplinary groups.

TABLE 14 Characteristics of the six sites sampled (note that exact figures are not given to protect anonymity)

Site	Activation	Status	Lead	Throughput ^a	SHMI ^a	Mortality rate ^a
1	Early	Teaching	Anaesthetist	High	Lower than expected	Higher than expected
2	Early	DGH	Critical care	Medium	Higher than expected	Within control limits
3	Early	Teaching	Anaesthetist	Medium	Within control limits	Within control limits
4	Late	Teaching	Anaesthetist	High	Lower than expected	Within control limits
5	Late	DGH	Critical care	Medium	Within control limits	Within control limits
6	Late	Teaching	Critical care	High	Higher than expected	Within control limits

DGH, district general hospital; SHMI, summary hospital-level mortality indicator.

a Surgical throughput, SHMI and mortality rates for all surgery: figures pertaining to 2013 were used, sourced from drfosterintelligence.co.uk.⁶⁸

For each site, initial contact was made by e-mail and assent to participate gained from the local lead, following which, appropriate research governance approvals were obtained. First face-to-face contact was usually made at the activation meeting for that site, after which arrangements were made for interviews with the core team, observation of planned EPOCH trial-related activities and contact with wider key stakeholders. Generally speaking, both the core teams and the groups of wider stakeholders were smaller in numbers than originally participated and, consequently, a lower number of interviews were undertaken than originally planned. This was most notable in site 3, where a very limited number of individuals were actively involved in EPOCH trial activities, for reasons that are explained in the site-by-site accounts (see *Site 3*). Data collection was undertaken by two researchers: David Kocman (sites 2–6) and Janet Willars (site 1).

The research strategy was to follow six participating sites from activation until the end of the trial, with a view to reconstructing trajectories of implementation capturing key nodes of decision-making, factors affecting implementation, actors involved and their understandings, and the implementation tools and strategies they chose to deploy. This approach was particularly attentive to how local QI teams dealt with obstacles and challenges over time. Across all sites, a total of 54 interviews was undertaken, along with > 200 hours of observation (including observation of activation events and other regional and national meetings, such as the 12-week follow-up events that were run for most clusters and the national EPOCH trial team meeting that occurred towards the end of the trial) (*Table 15*). Interviewees in each site included local EPOCH trial leads, other doctors from the three main disciplines at various levels of seniority (including registrars and foundation-level doctors, as well as consultants), a small number of doctors from other, related disciplines, such as radiology, nurses involved in the perioperative pathway and a small

TABLE 15 Data collection in all sites

Location	Interviews (n)	Observation (hours) (n)
1	11	72
2	20	54
3	5	18
4	7	18
5	6	12
6	4	18
National	1	24
Total	54	216

number of administrators. Observations focused primarily on formal meetings of stakeholders, including individuals from multiple professions and clinical disciplines. The number of interviews and observations conducted across sites varied. This variation depended mainly on the length of participation in the trial. Sites recruited from cohorts activated early in the trial afforded longer participation time than sites recruited from cohorts activated later. The lengths of their implementation journeys affected the intensity of ethnographic engagement. The exception was site 3, where levels of data collection were similar to sites recruited later in the trial. This was caused by local struggles with implementation and subsequent feeling by participants that there was 'not much to show'. The quantity of data collected was therefore determined, in part, by constraints in the field, and greater numbers of interviews and observations in the later sites may have enriched the study. However, overall, themes identified in the course of analysis were convergent and our analysis appeared to be at or close to theoretical saturation.

Topic guides for interviews and observation schedules focused on the quality of care locally, the aspects of the care pathway being addressed by local efforts, and the strategies and interventions that local teams were deploying to improve care. Interviews with clinicians were held at several points during the trial to monitor progress and reflect on what had been achieved, what had impeded progress and what might have been done differently. Final interviews conducted towards the end of the trial invited local leads to reflect on their improvement effort as a whole and asked about plans beyond the EPOCH trial. Some leads stated that they valued the interviews as a unique opportunity to step out of their busy schedules, sit down and reflect.

All interviews were digitally audio-recorded and field notes recorded in a diary at the time of observation, or as soon as possible afterwards. Interview recordings, fieldnotes and within-team debriefs discussing the data collected were then professionally transcribed. Analysis of phase 2 qualitative data was based on the constant comparative method,⁶⁹ but informed by sensitising concepts arising from the literature, from discussion within the team and from the findings of phase 1. In particular, coding for the purpose of this report was structured around the key components of the EPOCH trial theory of change (use of PDSA cycles, use of a team approach, use of data to inform and encourage improvement and engagement with wider stakeholders). Alongside this, a more inductive approach to coding, involving coding of concepts and ideas detected through a line-by-line reading of transcripts from interviews, fieldnotes and debriefs, proceeded concurrently, enriching understanding and giving rise to other lines of analysis not directly relevant to the research questions, which will be reported in future separate publications.

The EPOCH trial approach as received and understood

Activation meetings were the starting point for sites' improvement journeys in the trial. Activation was a central part of the EPOCH trial's approach to seeking to improve care. Activation of sites in a regional cluster was akin to the initial administration of a treatment to perioperative care practice within a region, enabling the trial to distinguish between before- and after-treatment periods for measurement purposes. In other words, it marked the formal commencement of the intervention in each cluster. In terms of QI, the EPOCH trial approach was to use the activation meetings for multiple purposes. They were meant to generate energy and build momentum among site QI leads, and to provide them with tools and strategies to take back and use in implementation. Some components of the activation meetings had more than one role. Opening presentations by the EPOCH trial team, for example, used data to generate buy-in among the audience. They showed the magnitude of the problem and crafted a pretext for proposing the EPOCH trial as a legitimate and necessary solution. At the same time, the presentations showed participants how to use data themselves, and enabled local QI leads to replicate these methods later in their own sites.

The EPOCH trial approach provided participants with a range of tools and QI 'knowledge bites' to take away. Key messages revolved around introducing local leads to key concepts in QI. Among other things, leads were prompted to realise that the care provided by their institutions was not perfect and that doing the right thing was about working differently in the way front-line care was provided, rather than necessarily requiring more resources. Participants were also told that improving care involved combining evidence-based clinical practice with thinking about 'softer skills' invested into persuading colleagues, understanding

variation and building up knowledge about how to do things differently. Data and measurement were said to occupy a central role in persuasion; and persuasion, engaging others and the so-called 'soft periphery' of activities around the clinical and process interventions proposed were to be placed at the heart of making change happen. Implementation framed as a social problem was at the heart of the theory of change. Therefore, the key message was that implementation was a social as much as a technical endeavour, and that social relationships and contexts were a precondition to successful implementation. Participants learned that generating a collective commitment among colleagues would enhance change, building commitment to the shared objectives was crucial in maintaining momentum and segmenting the care pathway into smaller, more achievable improvement projects delivered through rapid cycles of testing was likely to be more effective than lengthy negotiating of top-down implementation:

Just get on and do it, have these standing meetings, if people aren't available don't worry about them, if people are resisting forget about them, go with the grain and make it easy for yourselves. So I think that was the rationale underlying this whole thing, that basically OK yeah, we could spend a lot of time training them in the detail of QI methods, but that wouldn't be time well spent, and ultimately this project doesn't hinge on getting PDSA cycles and measurement just perfect, it hinges on enthusiasm and sustaining that enthusiasm.

Fieldwork debrief, cluster 2 activation meeting

The EPOCH trial approach was intended to be practical and relevant. To this end, the EPOCH trial team prompted local QI leads to start planning their first steps based on baseline overviews they had prepared before the meetings. To maintain the energy created during activation meetings, the EPOCH trial team asked participants what they were going to do 'by next Tuesday'. Importantly, the EPOCH trial team also provided participants with a set of tools. Most notable among these, participants reflected, were run chart templates developed by the EPOCH trial team for participants to populate with a selection of their local NELA data. The run charts allowed sites to monitor nine of the EPOCH trial's process measures and in-hospital mortality.

Participants perceived the EPOCH trial cluster activation meetings, as well as the 12-week follow-up meetings, positively. Unequivocally, they felt that the EPOCH trial team demonstrated the relevance of the project. They felt energised by them and expressed satisfaction about the useful tools and tips provided. They also reflected positively on the practical nature of the meetings and the opportunity to share ideas and learn from others:

It was quite useful actually because it was more sharing of ideas and lots of looking at each other's data and sharing ideas basically. It seems to be the way it has worked so far. Someone comes up with a good idea and everyone else shares or steals it [laughs]. The things we have used here have been stolen from other people.

Anaesthetic senior house officer, site 5

The impact of activation meetings was demonstrated in a number of ways, evidenced by our data:

- Leads in all sites adopted the EPOCH trial narrative of the nature of the problem and the solution.
- Among the tools, participants reflected particularly positively on run charts provided by the EPOCH trial team to support data monitoring and visualisation in local engagement exercises.
- Leads professed a shared commitment to change and agreed to propagate this among their colleagues.
- Sites started to use tools provided to them by the EPOCH trial team. They mainly adopted the use of data as a tool for building commitment among colleagues and a way of bringing them on board of the improvement project.
- Leads started discussing and planning change that was within their direct reach and easier to control.
- Leads shared resources on Moodle [release 2.6, URL: <https://moodle.org> (accessed 17 July 2019)], such as the checklist-based boarding card.

However, our observations and interviews showed that most of the local leads did not come to activation meetings as *tabulae rasae*, only to become enthused and educated. They were mostly senior clinicians who often had spent years as local champions, and sometimes lone enthusiasts, trying to improve perioperative care. Their buy-in was often already high and many had already achieved local improvements relevant to the EPOCH trial's mission, long before the activation meetings. Nonetheless, even in those cases, the activation meeting was an important place for learning and sharing experiences. It was important for local enthusiasts to see that they 'were not alone' in struggling to improve perioperative care and learn how other sites managed to change aspects of care (e.g. to break into elective lists):

I've always felt that the emergency laparotomy patients have been treated as a bit of an underclass, you know, we could be doing things a bit better for them, there's never been the interest or the resources thrown at it, so when NELA turned up I thought that's a fantastic idea and was keen to get involved. [. . .] And then EPOCH [. . .], it was just really looking at the NELA website and the EPOCH study, and [colleague] said 'oh yes, we enrolled onto this study as well'.

Anaesthetic lead, site 2

There was quite an interest from some of the other participants, in how is it that they achieved this change, how, how did they reorganise it. And then he went on explaining how they simply shifted the . . . well of course the mind, but also the structure, the organisational structure of operating, using theatres, and they are willing to send elective cases home, I think, that was the bottom line of it.

Fieldwork debrief, cluster 4 activation meeting

The EPOCH trial approach as enacted

On the face of it, the six sites followed a similar path in how they went about putting the EPOCH trial into practice. Site leads undertook education of their peers, focusing on showing colleagues how well or poorly they performed in terms of outcomes, record-keeping, compliance with evidence-based processes, introducing what they wanted to do [typically using the Portsmouth Physiological and Operative Severity Score for the enumeration of Mortality and morbidity (P-POSSUM) mortality risk assessment tool] and making the tools available to fellow clinicians. They tweaked existing systems (e.g. intervening to make the P-POSSUM part of existing theatre booking systems) and designed and introduced new tools (e.g. boarding cards that had to be completed prior to surgery) to support, monitor and enforce the pathway elements. They communicated within teams, monitored their plans, looked at NELA-based run charts and tried to devise alternatives when their plans were not materialising as planned. All of these activities had a strong basis in the activation meetings, in which options were discussed and interventions that had been successful elsewhere were advocated. For example, the 'boarding card' approach was usually presented at the activation meetings, and both this intervention and interventions to enforce P-POSSUM completion as part of theatre booking processes were highlighted as good practice in the first report of NELA.²

In terms of improvement planning, all sites used a combination of NELA data, five case reviews and older local audit exercises. Invariably across sites, they found that they were 'doing things right' in some aspects of the pathway and were not so strong in others. Indeed, this variation in clinical practice was exactly what was targeted by the EPOCH trial:

Some parts are quite good. There are a lot of parts that are very inconsistent and unfortunately a lot of those are very basic though in theory should be easy to change.

Consultant anaesthetist, site 6

Use of plan-do-study-act

As noted above, at activation meetings, the use of PDSA cycles was presented to participating teams explicitly as a model for experimentation and the planning of change. In addition to practical activities designed to demonstrate the usefulness of the approach – and its preferability to developing a blueprint for change without testing it in practice – the activation meetings provided references and reading materials that would help each team to follow the PDSA approach in practice. In this sense, PDSA was presented to QI leads in a 'textbook' format, with a clear set of instructions and supporting tools for putting it into practice.

No site, however, applied the formal PDSA methodology by the book. Instead, sites adopted a less formal planning approach, which included the general tenets trying out, reviewing and improving small changes, but which was otherwise much looser (e.g. typically excluding the setting of numerical goals against which to measure progress). This less formal planning was seen by site leads as more practical and intuitive. It followed a basic planning cycle of assessing the situation, planning what to change, preparing the approach to change, doing an intervention and evaluating their actions.

This more informal, less tight approach appeared, in part, to be a function of the fact that many parameters were predetermined by the EPOCH trial format: data collection methodology, process measures, as well as an idea about the kinds of changes needed. Site leads knew what shifts in indicators meant improvement and what improvements they wanted to achieve (higher compliance with process measures). This meant that less deliberation was needed about what kinds of changes to make; the focus was to a large extent determined by the trial. The leads did, in general, continue to use the language of 'cycles' of improvement, but these cycles did not follow the (quite specific and exacting) PDSA blueprint. Rather, typically, the starting point was a hunch or intuition around what the initial focus should be, informed by discussions in the activation meetings, rather than something systematically thought through and planned; the fact that they did not adhere fully to the formal PDSA cycle should not be taken to mean that sites failed to engage in creative experimentation. Local deliberation focused on which elements of the pathway to choose to implement, which tools to use to support implementation, how to adapt them or devise new ones and whether or not to change strategy when the tools, even after adaptation, were found not to be working. When boarding passes and other interventions were used as tools of improvement, they were thus iteratively modified, albeit not in the tight, systematic way prescribed by the PDSA methodology:

When we first started, we thought right, I know [person 7] and [person 5] are really keen than rather than try to introduce all the interventions together that we would do much smaller cycles which is why we started with [the boarding pass]. And we felt that the compliance was so poor because there were almost too many bits of paper, they were picking up a NELA document then they had to remember to get the end-of-care sticker and then they had to remember to get the boarding pass, and we just were going, 'It is too many things', and people were getting confused what they were doing. So we said well let's get rid of those two and incorporate it into one, stick it on the front and then they just pick up that bundle of paperwork and it is all there for them.

Consultant anaesthetist, site 2

This more intuitive approach to the cycle of improvement continued through the process.

Team approach

Developing a team approach was suggested by the national EPOCH trial team and all sites committed to it by formally nominating representatives of the three specialisms recognised as key in the EPOCH trial: (1) surgery, (2) anaesthesia and (3) critical care. However, the degree of continued involvement from the three specialisms varied significantly and some sites reported difficulties in securing ongoing engagement from surgery in particular. This typically meant that, in practice, project development was carried out by anaesthetists and/or critical care specialists.

When active surgical involvement in the core EPOCH trial team was lacking, participants saw this as a disadvantage for engaging surgical colleagues in improving the pathway (see *Engagement*). This appeared to be a particularly important barrier in site 3.

In contrast, three of the sites (sites 2, 5 and 6) secured ongoing, regular engagement from surgeons as part of the core EPOCH trial teams and they reported at least some success in interesting practitioners in the surgical department.

The experiences of site 1 are perhaps particularly instructive in demonstrating the value of multidisciplinary involvement in the core local team and, particularly, the importance of strategically positioned actors.

Here, activity relating to the EPOCH trial really took off only after the appointment of a clinical surgical lead in emergency surgery who took the view that NELA and the EPOCH trial were integral parts of improving the field. That single, but interested and strategically placed, actor dramatically improved access to others, which had not been achieved before by a handful of less strategically positioned EPOCH trial team members.

Use of data

All six sites tried hard to collect and use data in their improvement efforts. All used run charts based on NELA data to monitor care process and to attempt to mobilise support from colleagues. However, sites often struggled with using sampling batches of five sequential case notes suggested to monitor changes in compliance with key processes through time. Most used this technique, which had been recommended by the EPOCH trial team, to establish their baseline prior to the activation meetings. However, no site maintained this method of monitoring 6 months after activation.

Some sites (sites 4 and 6) were able to build on existing audit work that predated NELA and the EPOCH trial. The experience that these sites had in collecting and making use of data for improvement seemed advantageous, in that leads were aware of the power of such data and how to present them, and their colleagues were more immediately accepting of the value of data for identifying areas of weakness and targeting improvement efforts. However, this more mature 'data culture' also meant that areas for improvement apparent in the data had already been identified and sometimes addressed, and so the data were not so helpful in identifying 'low-hanging fruit' and swift potential wins. In site 6, for example, a series of local audits that had commenced several years before the EPOCH trial had already led to concerted efforts to improve care, and these had continued with active efforts to make use of learning from NELA, in collaboration with local commissioners. This meant that several key changes [to preoperative risk stratification, theatre capacity and intensive care unit (ICU) admission] had already been made, potentially reducing 'head room' for further improvements through involvement in the EPOCH trial:

We had looked at our data, on the basis of NELA we had looked at our data and it continued to be poor. [. . .] So a commissioner-led, unilateral decision was made to admit all patients with emergency laparotomies into ITU [intensive therapy unit]. So we admit about 92% of our patients now who go directly to ITU. The default is ITU and if there is a problem or someone is deemed to not be suitable for ITU that is the only reason they don't go there.

Engagement

Some sites had better experience of effectively engaging with colleagues and relevant departments than other sites. Notable in particular was the range of stakeholders engaged across the six sites, going far beyond doctors and nurses in the core three specialties of surgery, anaesthesia and critical care. Local teams drew on the local connections they enjoyed to their advantage, pulling in contacts in management, other disciplines such as radiology, and clinicians and administrators with particular responsibilities relevant to the pathway, for example sepsis identification and treatment:

Interviewer: What professional groups are involved, so the three main medical specialisms?

Participant (surgical lead, site 2): Yeah so in terms of, well obviously anaesthetics, critical care and surgery. Medicine are a group that need to be involved but up until now they haven't really had much involvement and when I e-mailed [colleague] her reply was 'I thought this was more anaesthetic, critical care, surgery type of thing'. So when I then wrote back and said about the care of the elderly role I think she probably thought 'oh OK that's quite a major thing'. So at the stakeholder meeting tonight I'm hoping that [colleagues] have had a talk about it already. Other professional groups, radiology in terms of the CT [computerised tomography] requests. We already have good back-up already in terms of 24-hour access to microbiology and CT reporting. And then we've got the allied care, so we've got the physio[therapist]s, the dietitians. Dietitian is a bit, we do have, there's not enough dietitians and they have got business cases already to try to get more dietitians, not on the back of this but on the request that we've got for critical care, 'cause at the minute they come in maybe there times a week and we're trying to get them on the ward rounds every day. So they've got a business case there.

Interviewer: Anyone from the board?

Participant (surgical lead, site 2): In terms of the board we've got yeah, we've got the business unit manager. I've tried to get our safety lead but he's away today, he can't come. We've got one of the execs [executives] and then we've got – who else is going to come? Head of theatres. Yeah so obviously there's the whole theatre team as well, so there's a lot of teams. Yeah.

The ability to engage colleagues successfully and encourage active involvement in improvement efforts, depended to a large extent on existing relationships and, more broadly, on the local organisational culture and its receptiveness to challenge and change. Similar efforts in sites to use, for example, existing data on process compliance and patient outcomes to convince colleagues of the need to change could result in very different degrees of engagement, with scepticism and hostility resulting in some sites, whereas in others, they brought new recruits to the cause. A strong network of existing relationships among supportive colleagues seemed crucial in making the difference between efforts that floundered and efforts that resulted in tangible change. The nature of engagement that was viable also depended on the attitudes towards local performance and improvement efforts that prevailed prior to the activation of the EPOCH trial: there were notable differences in the degree to which local stakeholders had examined their own performance and intervened to improve in the years prior to the EPOCH trial, which translated into differential enthusiasm for the EPOCH team itself, when it was introduced.

The six sites in detail

Site 1

Site 1 was a large urban hospital. At activation, the local team were rather pessimistic about the prospects for instigating change in response to the EPOCH trial, mainly due to ongoing organisational changes in the trust as a whole and in surgery in particular. Through time, they fostered closer links with other key stakeholders, notably the trust's research and development (R&D) team, which provided assistance with data monitoring. Towards the end of the trial, a recently appointed QI lead was hoping to achieve some momentum.

In terms of the EPOCH team pathway itself, site 1 focused largely on postoperative care. The intention was that emergency laparotomy patients be admitted to critical care by default. When critical care lacked the capacity to take the patients and they needed to be cared for on a ward instead, an ICU outreach team would provide their support. This was, indeed, mandated in the first months. They also planned to focus on improving recognition of emergency laparotomy patients, particularly those who risked 'disappearing' in the course of prolonged stays on medical wards, to 'make this group more overt to the whole hospital and make the whole hospital aware of what happens to these patients if a patient does badly' (consultant anaesthetist).

Participants in site 1 saw the size of their organisation as a key challenge, making it particularly complicated to achieve high consultant presence, speedy access to radiology and getting the patient into theatre within the target of 6 hours. Nevertheless, by the end of the trial, the local team felt that there had been focus on all three of these fronts. However, they played down the role of their own efforts in this, seeing them instead as byproducts of the wider reorganisation, along with a local, longer-established focus on improving care for trauma patients:

Getting to theatre is much better than it used to be because of the peer review stipulations for trauma and the need to simply crash into theatre. So we have Code Red cases quite often, and as a consequence when emergency laparotomies come through, we treat them in the same way. [. . .] I think they're seen as having the same degree of urgency as someone would have coming through with major trauma.

Consultant surgeon

We are doing all the work on that at the moment. I think it probably won't be fully operational by the time EPOCH finishes. It will be there but it won't have had any chance to impact.

Intensive care consultant

Use of plan–do–study–act

In common with the other sites, site 1 did not deploy PDSA cycles, at least by the formal, textbook definition. Instead, a more informal planning cycle was in evidence. This picked up pace as the focus on the EPOCH trial intensified following the appointment of a surgical QI lead in the second half of the trial. Especially, at first, the approach adopted in site 1 appeared to be focused on gaining 'quick wins' that could be delivered easily, and did not require detailed planning and testing. Through time, a more systematic approach to review and planning was adopted, as the team started to meet regularly with the trust's R&D team:

EPOCH is always part of it, and we discuss what we can do next, what is important at the moment.

Research nurse

Team

The local team started with two clinicians and a research nurse, who mainly helped with 'the approval side and with the R&D stuff and then co-ordinating meetings and just the communication side of things. Setting up a distribution list, inviting people to various things' (research nurse). Surgery was not represented among the team members initially and the team mainly focused on critical care aspects of the pathway. Other clinicians were ad hoc coming on board and picking up pieces of work, such as an intensive care specialist who developed her own pathway for emergency laparotomy patients admitted to critical care. In the first part of the trial, the team appeared to struggle with difficulties with co-ordination and uncertain leadership:

Organisation 1 has been a bit of a challenge in many ways because they've got a principal investigator who is very difficult to get hold of and therefore the team doesn't seem to meet as you would expect maybe a team who are developing a project to meet. And so they've got a principal investigator who sort of devises things, but it seems as though the main crux of the project is run by somebody who is not the PI [principal investigator]. So it's [person 2], he sort of runs it. The research nurse is [person 3]. She was my port of call to begin with, but every time I've gone to see [person 3] it's almost like she doesn't really know what her role is in EPOCH.

Ethnographic debrief

In the second half of the trial, two new appointments were made: a clinical lead responsible for wider improvements in emergency surgery, and a clinical lead in perioperative medicine who focused on NELA data collection and took on the role of educating junior doctors about the risk scoring. Both were newly created clinical positions with dedicated non-clinical time for improvement work. With these two new allies, the EPOCH trial received a boost. According to the surgical lead, their focus was not on the EPOCH trial as such, but a service-wide restructuring of trauma and emergency care. However, the impact of the EPOCH trial was evident. By the end of the trial period, the EPOCH trial had become a regular agenda item at weekly research meetings:

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

Use of data

At activation, the team in site 1 decided to do their own audit of emergency laparotomy care, rather than use the five case reviews method advocated at the activation meeting. This audit was designed to provide baseline data to inform further implementation. Site 1 struggled to collect NELA data. Data started to be collected systematically only later in the trial after new appointments were made and the new leads in emergency surgery and perioperative medicine could offer dedicated time to support data collection. The perioperative lead in particular undertook 'a lot of the emergency laparotomies [and] makes sure that the NELA data is filled in' (research nurse).

Things had shifted significantly by the end of the trial. A clear commitment emerged to using NELA data locally for improvement, along with a recognition that these were essential to addressing EPOCH's objectives in a systematic way. The EPOCH trial team understood that 'EPOCH only works with NELA working. That is [why] we have to shift towards supporting NELA to make EPOCH work' (research nurse). However, even at this point, the focus was more on ensuring consistent data collection, with activities to make active use of NELA data for improvement deferred until after the trial period:

We need to look at the recent outcome of the NELA. But we haven't, because we were concentrating on NELA and less on the EPOCH care pathway, we haven't been able to monitor that unfortunately.

Research nurse

For most of the trial, other sources of information were deployed to assess local quality of care in place of NELA data, such as experiential knowledge:

We're reasonably quick in doing CT [computerised tomography] scans. I'm talking about the patients I get involved as an intensive care consulting doctor. So if they call me because they are worried, and I go and see the patient and I request a CT scan, I get it done within 2 hours, but that's probably because it's an intensive care request. I couldn't comment on the patients that we don't get involved in.

Intensive care clinical fellow

Engagement

In site 1, with the initial focus on postoperative journey, the ICU was the prime target of engagement activity. As above, this seemed to represent a quick win that could be actioned rapidly, with immediate positive results: in seeking to build a consensus that intensive care should admit all emergency laparotomies by default, the team felt that they were pushing at an open door:

There's pretty much a consensus amongst all the intensive care consultants that we should be just admitting these patients without question. Just admit them because they are a high-risk group and that's what we try to do.

Intensive care clinical fellow

During the development of the postoperative pathway/protocol, the team sought advice from a microbiologist. When implementing it, they turned to the R&D team who were overseeing innovations developed as part of trials, and held further consultations with critical care consultants and anaesthetists. The team also targeted nursing staff with 'training and explanation' (intensive care clinical fellow).

The team also sought to engage, with mixed success, other groups pertinent to improving emergency surgery, from management who were repeatedly presented with business cases for large-scale organisational improvements, including staffing levels, to the consultant group to mandate the use of P-POSSUM scoring system. They also sought to engage with theatre managers to improve access to theatres, the radiology department and others, such as information technology (IT) services simultaneously implementing a new

electronic booking system. Again, there was a sense that although none of these groups was obstructive, larger-scale agenda were their primary preoccupation. If these aligned with the EPOCH trial, then the result was of mutual benefit, if not, then the EPOCH trial would have to be sidelined, at least temporarily:

What me and [person 3] and some nurses from around the trust are doing is developing a template for that form to be for the EPOCH patients and have it on the trust-wide electronic system so we have been involved with the EPR people, that is the Electronic Patient Record people, to help with the technical side of how that gets put on and everything and how we might be able to use data that is already in the system to populate the form.

Research nurse

Towards the end of the trial, the newly recruited surgical lead captured the intended effect of their engagement as:

. . . a culture change within surgery that's happened and the acceptance that emergency general surgery actually has a subspecialty role in its own right and actually trauma surgery exists even as an entity; and that's something that's only happened in the last year really within the [organisation].

Consultant surgeon

All in all, by the end of the trial, the feeling among the EPOCH trial team members was of unexpected improvement and a sense that wider organisational changes had finally aligned in a way that was conducive to the aims of the EPOCH trial. Thus, there was an upbeat mood about the prospect of a lasting commitment to continue improving emergency surgery, albeit according to a time scale that was much longer term than the EPOCH trial's time-limited intervention:

It's a huge amount of progress. Yeah. But what we need now is sustainability.

Consultant surgeon

I just feel so positive about it because [person 2] is so influential and keen to make it happen.

Research nurse

Site 2

In site 2, the team began the trial with a plan to implement a selection of pathway items using the boarding card and the end-of-surgery sticker as implementation tools:

We're not going to try to do all 37. [. . .] At the meeting we thought that some of the end-of-surgery bundle we could implement fairly easily and [name 1] felt that we could do the P-POSSUM score very easily at the start [. . .] and then I get them to do cardiac output monitoring and the admission to critical care. Yeah they're probably the first ones we'll start with. They seem the easier ones to do, I feel like we've got more control over them compared to trying to get a computed tomography [CT] within 2 hours and care of the elderly review there. They seem like a bigger challenge at the minute.

Consultant anaesthetist

Soon, however, they extended their scope. Along with the bundle logic of improvement, site 2 decided to implement the 'whole bloomin' lot', with the exception of the postoperative checklist:

[The] end point 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, very much in the way that the models were done for reducing central line sepsis in the States [where they] drastically reduced the amount of central line sepsis by just putting a whole sheet of measures in and say sod it, we don't matter which one does it, let's just do the whole lot.

Consultant surgeon

Another reason for changing the course was an attempt to integrate multiple tools into one:

... we kind of felt that the compliance was so poor because there were almost too many bits of paper, so they were picking up a NELA document then they had to remember to get the end-of-care sticker and then they had to remember to get the boarding pass, and we just were going it is too many things and people were getting confused what they were doing so we said well let's get rid of those two and incorporate it into one, stick it on the front and then they just pick up that bundle of paperwork and it is all there for them.

Consultant anaesthetist

Their main implementation tool was, thus, a checklist that brought all the EPOCH trial bundles together. But, by the end of the trial, they were still discussing the need to 'implement the checklist' as progress had not been as rapid as they had hoped. Nevertheless, as with other sites, there was partial success in increasing compliance with the process, evidenced through the NELA-based run charts produced by the team.

Use of plan-do-study-act

As elsewhere, site 2 did not adopt formal PDSA cycles. A less formal, though still cyclical, approach to planning was performed, reliant on informal review of progress at regular meetings of the local team. These typically covered the situation and status of interventions, discussion of what had worked and what had not, decision on what to do next and agreement on who would do it. At a high level, then, the format followed the general principles of planning, intervening, learning and adapting, but these represented flexible principles governing the format of meetings, rather than a tight methodology.

Team

The local QI team was bigger than in most other sites, as site 2 operated across two hospitals. It comprised two surgical consultants, two critical care consultants and an anaesthetic consultant. The team met formally several times over the course of the trial. Alongside this, more frequently, individual members liaised informally in relation to specific issues, such as undertaking five case reviews, interpreting run charts and preparing presentations, devising implementation tools and reviewing implementation tools:

... so she is looking at the anaesthetic charts and looking at the surgical notes to check whether what has been claimed has been done is getting done. So yes, [person 3] and [person 4] met last Wednesday and I am going to find out tomorrow exactly what the outcome of that was, but an e-mail did come round, following that meeting, saying that again people aren't completing them or they are not getting put in the right place in theatres, they are getting found in the notes and then you are having to try to chase them up and they are getting lost.

Consultant anaesthetist

Use of data

During the first months after activation, the five case reviews method advocated at activation meetings was used. However, they soon stopped using this method, as sustaining the monthly periodicity proved challenging and, more importantly, it was felt that auditing the checklist tool provided a suitable monitoring mechanism:

We are not really doing [the five case reviews], the reason being [. . .] is that when you look through those cases essentially it is that EPOCH form that we have designed so if they get completed that will then in theory should tell us, and [. . .] it is doing the same job twice.

Consultant anaesthetist

National Emergency Laparotomy Audit data, however, were used consistently almost from the outset. Run charts were reviewed periodically and presented on several occasions during the trial, not least because the checklist was found not to be completed and the only reliable data were the run charts

provided by NELA audit collection. To support the audit of their extended pathway checklist, the team in site 2 at one stage asked the national NELA team to add other elements of the pathway so that they could monitor process compliance through the run charts, rather than their locally devised checklist:

Adding the questions to the NELA data set will enable us to get a better return on our data, rather than [name 2] having to go and dig out all of these white forms – if all of the questions on [the checklist] can be taken from the NELA data set.

Consultant anaesthetist

[So person 7] went down to one of the London meetings, to see whether [NELA] could add on the other EPOCH questions. I think the bottom line was that NELA could do it. [But] you could still lock your case without those questions being completed. [So] we kind of said if you can't lock it until you complete the ones we want you to complete as well. But NELA couldn't do that, so I think people on the team felt that you may as well just keep it as a paper format.

Consultant anaesthetist

The commitment to making proactive use of NELA data continued beyond the trial, with the local team planning to keep on using the NELA-derived run charts after the trial had ended.

Engagement

The team in site 2 undertook extensive efforts to engage their colleagues. These included, in particular, junior doctors in surgery and anaesthesia, with whom the team had close contacts and was able to wield some influence through regular informal contact and dissemination. Other departmental colleagues in surgery and anaesthesia were also regularly engaged, through both formal presentations in numerous meetings and informal 'corridor chats'. The team also made early connections with the R&D department of the trust, where a research nurse had regular contact with the QI team, supporting them with specific audit tasks.

The QI team also identified other professional groups relevant to the pathway and considered various strategies of engagement:

Other professional groups, radiology in terms of the CT [computerised tomography] requests. We already have good back-up already in terms of 24-hour access to microbiology and CT reporting. And then we've got the allied care, so we've got the physio[therapist]s, the dietitians. Dietitian is a bit, we do have, there's not enough dietitians and they have got business cases already to try to get more dieticians, not on the back of this but on the request that we've got for critical care,' cause at the minute they come in maybe three times a week and we're trying to get them on the ward rounds every day. So they've got a business case there.

Consultant anaesthetist

At one point, the team discussed the possibility of attaching Commissioning for Quality and Innovation (CQUIN) to the pathway implementation. However, the idea later waned. Towards the end of the trial, even the QI lead could not remember whether or not this gave rise to any concrete action:

There was talk about trying to link it with the commissioners and CQUIN targets. [Person 4] was talking to the committees, in terms of what part of the pathway we could link. I need to go back to [person 4] and see what ultimately they decided to do.

Consultant anaesthetist

Site 3

The team in site 3 was characterised from the start by a degree of pessimism about the prospect of much change in response to the EPOCH trial. As in site 1, the team was acutely conscious of ongoing, large-scale changes at the trust level, including financial challenges and efforts to rationalise theatre space:

Taking it from the top, in terms of speaking to senior management, this hospital has been discussing for a year how to rebuild and where to rebuild and made a decision last week, so that's a year's worth of people's time taken up with that. Four months ago they decided to completely rejig the theatre schedules, and that has caused huge amounts of ructions. They want to change everyone's job plans and contracts for Saturday working, [. . .] half of the department doesn't know, half of us will not have too many changes, but we will have a new, they're extending the working day, so that affects everybody who works in the afternoon, and then they've got to decide if you're going to work extra lengths on some days what are you going to give up, and it has a knock-on effect on our on-call rotas and so many things. [. . .] Now, since everyone seems to think different things, worry about different, actually we don't know exactly how our job planning is going to be for the next, starting [in winter], so somehow, at least for me, there's no space, no room for new things.

Intensive care consultant

At the initial activation meeting, one team member even went as far as to suggest that site 3 might represent a 'control site', against which other sites, able to undertake more active work, might be compared. Nevertheless, some progress was evident in this site through time.

Use of plan–do–study–act

Given their pessimism about the prospects for change, the improvement efforts in site 3 evolved around engaging surgical colleagues and promoting risk stratification, as a relatively simple intervention that could be crucial to ensuring rapid movement through the care pathway for emergency surgery patients. This was not seen as requiring extensive development and testing through PDSA cycles, but rather as something to be achieved through a change in culture, which could be negotiated through engagement with surgeons. There was, however, a tension inherent in this strategy: surgical buy-in was to be improved via the P-POSSUM score, but the introduction of P-POSSUM, the team knew, was dependent on surgical input. Thus, they faced something of a 'Catch 22':

The P-POSSUM scoring, it's a surgical scoring system. I mean there are anaesthetists who know P-POSSUM and use it, but at the end of the day we would like that to come from the surgeons, as part of their risk scoring for when they discuss with the patients' potential problems. I mean we, we do know, we do use it sometimes, but I do think that it's something that actually, as part of the risk assessment prior to getting consent, they should do.

Intensive care consultant

At 10 months, with progress limited, the team decided instead on a more modest objective, to seek to maintain levels of contribution to the NELA audit, in a context when surgeons were seen as unco-operative, and the NELA audit had to be retrofilled by the anaesthetist. Although it was acknowledged that this would not directly affect the quality of care provided to the patients, it was at least seen as securing a foundation for future improvement efforts:

I actually think the fact that the NELA audit is happening is one of the things that changes the situation anyway but that is the same for everybody, and we have managed to feed back some of the data so I think that is probably as much we are going to do.

Intensive care consultant

At the end of the trial, the picture remained the same, the primary focus was on NELA completion, and the local lead was (still) hoping to start implementing the P-POSSUM, but with no definite time scale for

change, and the approach premised more on wider developments that might make it easier to comply with risk assessment, than with a specific project plan, PDSA-based or otherwise:

We are doing the data collection, because that, nobody goes back and completes the data sets. [And] we have [a plan to implement the P-POSSUM]. It is dead easy to have it as smartphone app [application]. So in terms of actually people accessing it, it is not a problem and we have got about half a dozen people who have actually bothered to do it.

Intensive care consultant

At the end of the trial there was a feeling that nothing had moved, despite having a team at the start, having tried to engage the wider community through presentations and even having done one-to-one tutorials. Consequently, the team felt that there was little to celebrate.

Team

Initially, the team comprised a surgical consultant, an anaesthetic consultant and a critical care consultant. However, within a few months the team disintegrated. The EPOCH trial lead took up a dual role and became lead for NELA as well. At the same time, she kept looking for other colleagues to create a new team. The struggle to build a new team lasted for most of the trial. Just before the end of the trial, the lead managed to form a small team with a senior registrar anaesthetist. An anaesthetist herself, the QI lead continued to look for a dedicated partner in surgery, but as of the end of the trial, was not successful. This failure to recruit and maintain a multidisciplinary team was seen as a major impediment to progress:

I think with the lack of . . . forward motion the surgeon fell by the wayside because there were huge repercussions on the surgical side that have caused a lot of ill feeling that has lasted for over 2 years. So there has been a big bun fight over there. On our side from the point of view of [person 3] I think he found things a bit becalmed and he . . . as a, a new consultant sort of got, at one point I think felt a bit adrift because he had actually, coming in from a different environment had took longer to find his feet.

Intensive care consultant

Use of data

A five-case review was completed for the activation meeting, but the team did not use this technique again over the course of the trial. Rather, the lead focused her efforts on regularly translating NELA data into run charts. The team used these data in around six presentations to colleagues over the course of the trial:

I have done several but the most recent one we were looking at the trends, of whether we were improving with risk scoring, risk stratification, delay to theatre time, and that was, came from the EPOCH data set up until the end of, was it January or December I can't remember now.

Intensive care consultant

The lead recognised the importance of a firm grounding in credible data to achieve buy-in and change among colleagues, 'it is the bread and butter that if you haven't got the basic data sets there, you can't actually build on anything'. However, in the absence of surgical engagement, she spent a lot of time trying to engage colleagues in data that they had an interest neither in collecting nor in using for improvement:

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

Engagement

Multidisciplinary engagement remained an unmet goal; throughout the trial, the lead reiterated how problematic the lack of surgical engagement had proved:

... as I say, the issue right from the very start has been the lack of engagement from certain of the senior surgeons, not all of them but some of them. An unwillingness to encourage their trainees to engage, which has been another issue. In practice the, you know you go and present at a meeting, you show your data, we look at, you know let's say the dead simple one of P-POSSUM scoring, which is so easy to do, and the clinical director of surgery is dead against it, he is too busy, he can't do this, why would he bother. And my response to him is you audit your own subspecialty which has a very small mortality rate, 1% to 3% for the average patient, obviously there is a higher-risk group, I said, and you refuse to contribute towards an audit to a much higher-risk subgroup. And he has no answer to that, it is all, you know, poo-pooed aside. So that is the sort of lack of engagement when you turn up at a meeting, and you have got the senior staff there, with one or two honourable exceptions, not encouraging their trainees to implement this despite the fact it has been implemented successfully elsewhere and the trainees are aware of this.

Intensive care consultant

In part, the issue seemed to be one of problem definition. The relatively low mortality rates at baseline meant that the lead struggled to convince colleagues that this was a problem worthy of their attention. Despite various engagement exercises, the lead reflected on how little success there was in terms of generating a sense of urgency through (1) presenting the problem and (2) presenting local mortality figures. The result, as she put it, was that surgeons and some anaesthetists did not see the problem proposed by the EPOCH trial as a problem. The EPOCH trial's programme theory saw data presentation as a key means of making the problem of emergency surgery morbidity and mortality visible, and shocking colleagues out of complacency and into action. In site 3, however, some presentations were received in the opposite way, as confirmation of relatively strong performance and as reassurance that intervention was not required. Colleagues were pleased that the hospital was doing better than the national average:

[Person 1] volunteered to deliver the presentation. What did the people say, I asked? Nothing much after the session itself, he said. They were pleased with the results. The run charts showed average mortality around 10%, and other measures were also quite good: that is on par or better than the national averages. For example, consultant presence was good and others. Risk assessment was not good. More can be done about that. And about a third of people go to ITU [intensive therapy unit]. I asked whether he thought it would be difficult to generate energy for change now when the results were received positively and the colleagues felt the hospital was doing well. [. . .] [person 1] said 'well, that's interesting, now you made me think about a similar case. This hospital had a problem with high mortality of patients following [a specific type of] surgery. The numbers were 13%. They invited a consultant who managed to get it down to 5%. But of course, people are still dying and you could improve your care even more. However, the feel is that this is an excellent result and it feels like for now we have a job done. So yes, continued [person 1], people being happy about relatively good outcomes may be a problem. Where do we go from mortality as low as 10% compared to the national average when there are others pressures, and when there is resistance from senior surgeons? We will see.

Notes from ethnographic observation

The failure to engage surgical colleagues led to a generalised sense of project failure for the team. Their P-POSSUM completion rates, as shown by NELA data, increased from zero to 10%. Despite the sense of failure, many of site 3's measures were good (mortality was better than the national average, consultant presence was at almost 100%, as were the figure, for time to theatre, time to scan and high-dependency unit admissions). However, these measures had been good before the EPOCH trial had started. The lead was a local enthusiast who has dedicated years to improving emergency surgery in site 3. She hoped to use the EPOCH trial to focus on some new elements, the P-POSSUM and, with it, the not-so-sick patients.

Within the project time frame, they did not succeed. It is, therefore, probable that the sense of failure reflected a sense of failed opportunity to improve (this time) rather than a sense of failing standards of care.

Site 4

Site 4 was a large teaching hospital, in which efforts to improve the emergency surgery pathway were already well under way. The site had developed a local version of the emergency laparotomy pathway that might be seen as a 'cousin' to the EPOCH trial pathway, rather than directly deriving from it. The pathway was developed at the same time as the EPOCH trial and both took the 2011 guidance as their reference point. Three years in design, site 4's pathway was intended to be easy 'for the end user to buy into' (intensive care consultant). It was more concise than the full EPOCH trial pathway and the focus on a smaller number of key components was seen as advantageous. The pathway had been fully designed by the time of activation, together with a dummy version ready to be integrated into a new electronic patient record system that was shortly to be launched.

Use of plan-do-study-act

No formal PDSA was used. However, the local team was continually planning and reviewing their journey on a more informal basis. The plan in site 4 was simpler than in other sites and did not involve the use of techniques, such as stickers or boarding passes, advocated at the activation meetings and popular elsewhere. Rather, site 4 planned to publicise the pathway through posters in theatres, visualising patient categories and actions, and to integrate risk categorisation into an electronic patient record system. Monitoring and audit were to be done through the NELA data:

So we will pull the NELA data. So if we pull the NELA data for the 2 months, we will have our performance and we know what we should do because we have that in our pathway so we need to find out what parts of the pathway are working well and which parts of the pathway are not working well.

Clinical fellow

In other words, pathway implementation here was reliant on the 'forcing function' provided by the new system and much less on securing enthusiasm and buy-in from colleagues. At 12 weeks, however, the electronic patient record system was yet to go live and little other activity had taken place. Consequently, the lead started thinking about 'doing one or two items' of the pathway in a more traditional, educational way. They initiated two projects: (1) using junior colleagues, whom they supervised, to educate others in using fluid monitoring, with the aim of improving its use and (2) improving risk scoring by, again, informing and educating others about use of a risk scoring mobile application, in lieu of the not-yet-functional electronic system and pending its eventual introduction. The project lead was not very optimistic about the prospects for short-term change, he realised that the elegant pathway was dependent on integration into an electronic system that could still take a long time to be introduced (and probably beyond the time frame of the trial):

Participant (anaesthetic lead, site 4): The idea was that the moment you book an emergency laparotomy, the navigator will pop up and it will tell you what to do, on the electronic system, it will tell you, you know, well fill in this data and this is what the risk is going to be. So that is now left to the subjectivity of the anaesthetist.

Interviewer: Which is the way it's been.

Participant (anaesthetic lead, site 4): It's always been, yeah. So we haven't been able to make that transition with any sort of certainty.

Interviewer: OK. So does that mean that this pathway is not really happening? Or does it mean that parts of it are not happening?

Participant (anaesthetic lead, site 4): I don't know. I think, it was always . . . The thing is that it's very difficult to police it. [. . .] The problem has always been, with any pathway, is the consistency with which it's done.

Ultimately, at the end of the trial, the local improvement project remained stuck in a post-launch, pre-implementation period. The IT department could not make the pathway operational through embedding risk stratification in the electronic record system. Alongside this, as in site 3, the lead felt that the efforts he had made to develop interest in the EPOCH trial had been impeded by difficulties in generating a sense of clinical problem. Mortality rates were below the national average and failed to generate the shock effect.

Team

The local pathway was a team creation, involving mainly clinicians from critical care and surgery, although the NELA lead was its main architect. Reliance on the electronic patient record system meant that the local EPOCH trial lead was content not to form a multidisciplinary team for the purposes of the EPOCH trial, but he did become a part of the team that had developed, and was seeking to implement, the local pathway. The team maintained mainly informal communications, with few specific meetings to take the EPOCH trial or the pathway forward:

So last Wednesday's meeting started off with [person 2] sending me a text on the morning saying can we meet up please. I said yes we could meet up at this time and we met up in the library and in the meantime [person 2] texted [person 4] or he spoke to [person 4] and he sent a message off to [person 3] and by the time we met, you know all four of us were meeting. So we are very good at those sorts of things. [. . .] that level of communication is very good.

Consultant anaesthetist

Use of data

The creation of the local pathway was data intensive. It was grounded in a local audit exercise, which took place in 2012, and reliant on NELA data:

We went to our theatre system and we looked at all the emergency laparotomies that had taken place in that calendar year. So I think it was from June 2011 to June 2012 and then from that we pulled out all patients over the age of 16 who had had emergency abdominal surgery of any type so the inclusion criteria was a little bit broader than NELA. So we included vascular, AAA [abdominal aortic aneurysm] ruptures. Then we reviewed the patient notes individually to make sure people were eligible and we cut it down to about 453 operations in that year and we did an audit of time to surgery from booking, seniority of consultant anaesthetist or surgeon and whether it was day or night, the proportion of patients who went to intensive care afterward and how long they went to intensive care for, or whether they went to a high dependency unit instead so a level 2 or level 3 post-op [postoperative], overall mortality and overall outcome.

Clinical fellow

Again, due to the expectation that pathway adherence would be secured through the forcing function of the electronic patient record system, there was limited use of data in site 4 under the auspices of the EPOCH trial. Prior to EPOCH trial activation, during the development of the local pathway between 2013 and 2015, those involved in developing the pathway had presented to their colleagues on four occasions, 'lobbying' them, as they put it, to generate buy-in for the pathway in preparation for its launch:

We then presented that to the anaesthetic team and the anaesthetic directorate, the intensive care team and the surgical team with a view to, I think, highlighting the issue mainly with our colleagues about this is actually a very high-risk group of patients.

Clinical fellow

There was no sense, however, that this had resulted in a groundswell of concern or enthusiasm to address emergency abdominal surgery as a major problem.

Engagement

The main efforts to engage colleagues in site 4 were related to the development of the local pathway prior to EPOCH trial activation. The team postponed formally launching the pathway until after activation. The key engagement activity after activation was presenting the pathway at a joint surgical and anaesthetic meeting. Interestingly, when the latest version of the pathway was introduced at the joint meeting, no mention of the EPOCH trial specifically was deemed necessary by the pathway designers. Nevertheless, the team felt that the EPOCH trial could be beneficial for their local improvement effort:

I think EPOCH helped us get a pathway in place because we were then part of a project. It gave it extra drive.

Clinical fellow

The local pathway was interdisciplinary, involving surgical and anaesthetic specialties, and it also required negotiating with other departments, such as microbiology, radiology, critical care, operating theatres, wards and accident & emergency (A&E):

The other important thing is by having a pathway we have then also got all the different teams who are involved in the patient's care to communicate and we have all got together to draw up the pathway and I actually think that has really helped as well. So it has been a collaboration more than us saying as a group that this is what we want. We talked to radiology and we talked to theatres and to A&E.

Clinical fellow

When the pathway was formally launched, it met with a favourable reception. There were voices of agreement and support for the introduction of the pathway. According to one of the designers, colleagues across departments praised the pathway for what was seen as concise and simple design (intensive care consultant). However, at the end of the trial, the team were still negotiating with IT services making their electronic pathway operational:

We haven't yet formalised a scoring system on the electronic health record system [. . .] That is what we were hoping but it is not as easy to do as you think it is and I think, in a way, we fantasised about an electronic health record of the Apple standard.

Consultant anaesthetist

Site 5

As with other sites, site 5 had some experience of improving emergency laparotomy before the EPOCH trial. Local champions among surgeons and critical care specialists carried out an audit of care in 2011 and initiated changes to care consistent with the pathway recommendations, namely consultant presence, access to radiology and access to intensive care. The mortality figure at activation was similar to the national average and the team saw joining the EPOCH trial as an opportunity to further local improvement. Formal risk scoring and other pathway elements clustered in the Surgical Survival Six were explored as a good starting point, with the view of implementing the whole pathway. The team also appropriated tools suggested by the national team, such as the 'boarding card' and NELA-based run charts. They translated the end-of-surgery bundle into a formal 'landing card':

The ideal that we are aiming for would be to have all of the 37 points done consistently for everybody and, although the way that I think we have approached it is to cater for the ones that are perhaps easier to understand and implement, I think in my mind I was hoping that what we do is start with the easy ones, that are more obvious to people who haven't looked into it and easier to implement and then, on the back of those, introduce the rest of them.

Anaesthetic senior house officer

According to the local lead, in this site over the course of the trial, the team achieved most of what it had set out to do. The list of achievements referred to areas that had been targeted before the trial, as well as through it. In the post-EPOCH trial period, the ongoing plan was to focus on implementing the postoperative bundle, which had not been achieved during the trial period:

Our major achievements were HDU [high-dependency unit] beds; we're getting most of our patients to HDU. Consultant review, CT [computerised tomography] scanning, [. . .] an improvement in risk assessment. Our challenges were, the major challenges were cardiac output monitoring, whether or not we believe that is something that we need to work on more is still out for debate, although we probably feel we should be better. [. . .] The other challenge we had was the postoperative review by care of the elderly physician [. . .] well vast majority of trusts aren't doing it. I think in 80, 90 per cent of places, most hospitals, over 70s aren't being seen by care of the elderly physicians. So that was the other major challenge, and [person 1] is working on that.

Intensive care consultant

Use of plan–do–study–act

Similar to other sites, site 5 did not use PDSA cycles as such, but local team members engaged in what the team saw as following the spirit if not the letter of the approach:

. . . everyone doing little bits and bobs, very much in the PDSA ethos.

Intensive care consultant

In practice, this meant moving along a planning cycle that was fairly systematic, but which lacked strict time-limited, measurable objectives:

See how it pans out, if it changes we'll come back and review it. And then we had a little, we had a meeting a few weeks ago, which was useful. We all just sat down, had a bit of a chinwag basically, thrashed out what we thought, where we thought we were falling short, and what else we needed to introduce, and just firmed up what we were doing, and obviously was lots of things that weren't working, if changes needed making to the forms. Again it's all based on fairly informal feedback, and we know where we need to improve from our run charts from last year from NELA anyway, and where we're falling short, you know, sepsis management, things like that, which I mentioned before.

Intensive care consultant

Team

The local team was led by an intensive care consultant and a surgical consultant. They maintained joint leadership throughout the project. The two teamed up with three junior doctors who were given responsibilities for developing specific tools (a boarding card, a landing card and a booklet for the comprehensive pathway). The consultants were instrumental in setting improvement strategy and provided continuous oversight and support for the junior doctors' tasks. They also handled engagement with other departments:

[Person 3], [person 2], [person 1], are doing all this stuff, and you know, they're keen and eager, we're just going to make sure we're just addressing this as we go along really, and just making sure things are tidied up and the forms are getting filled in and people are just paying attention to the quality, to the bundles.

Intensive care consultant

Use of data

The team made considerable use of data, mainly drawn from NELA, in presentations to anaesthetic staff and at an audit meeting to which wider colleagues were also invited. The process indicators suggested by the national EPOCH trial team were translated into run charts to demonstrate progress. Towards the end

of the trial, they also countenanced undertaking further, more specific analyses, but struggled to find the time to do so:

It was interesting, because we did have a mortality spike in October [. . .] operating by surgeons dipped in that month, it's usually quite high, it's usually 9% plus, and mortality went up, so we need to look at that data as well. Again, it's finding the time to do all this stuff, you look at every single patient subset, and, you know, without, the trust hasn't given anyone any time for this, so people doing it, you know, because they want to. So, you know, it would help if it had time, funded time for it, but you know that's never going to happen in the NHS, is it, so, not at the moment.

Intensive care consultant

Engagement

The team in site 5 were conscious of the importance of engaging their colleagues from the start, 'putting it out, because you know, as always, it's better to get many minds think about it than just yours' (intensive care consultant). In presenting the data from NELA to their colleagues, the local team were pleasantly surprised at the positive reception. Colleagues appeared ready to recognise the validity of the data, the need to improve and, moreover, the solutions put forward:

[At the audit meeting] I just basically put it up, the numbers speak for themselves. So patients aren't getting mortality scoring, they're not getting cardiac output monitoring, some patients aren't going to high-dependency units, we're not doing lactates, we're not doing risk scorings, we're not doing risk scorings at the beginning, at the end of theatre. Everyone accepted that, the numbers speak for themselves, here's what we want to do about it: boarding card, possibly landing card. Any questions? And everyone was fine with it. There were one or two questions . . . But usually just quite benign questions wanting a bit more information. So there was complete acceptance actually, I was quite surprised, I was ready for a bit of a contretemps, but you know, I know that people, a lot of people don't believe there's strong evidence for cardiac output monitoring, so I was expecting some comeback on that. [But] there wasn't anything.

Intensive care consultant

This experience reflected what the local lead saw as a relatively cohesive organisational culture across the professional groups, leaving it much more open to challenge and change than some of the other sites:

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.

Intensive care consultant

Correspondingly, there was strong engagement from anaesthetists and surgeons, perhaps assisted by the fact that both groups were actively represented in the local team, in contrast to some of the other sites:

. . . to be honest we haven't had many hurdles I don't feel. I mean I feel that, you know, [person 1] engaging the surgeons, I know certainly in some trusts they've had problems with some specialties engaging, other, other specialties not. We've had no problem with the surgical engagement, and have had no problem with the anaesthetic engagement either.

Intensive care consultant

Site 6

Efforts to improve emergency laparotomy care in site 6 had started several years before the EPOCH trial. In 2011 and 2012 a group from the hospital assessed their service against the then-published standards and initiated changes after reporting to the hospital management:

. . . at least from 2011 to now, there have been several clinicians who are at the front end of delivering this service, highlighting the issues and the concerns about whether or not we hit the targets set by

the college and that have been accepted by the Department of Health so they are essentially national guidelines.

Consultant surgeon

The main achievement stemming from these early initiatives was an arrangement to admit all emergency laparotomy patients to an intensive therapy unit. At activation, 9 in 10 emergency laparotomy patients were admitted to ICU, compared with 5 in 10 in 2012. Staff in site 6 also had worked on improving access to screening and access to theatres. This led to halving mortality in emergency laparotomies. In contrast to most other sites, where staff responded to an invitation from the EPOCH trial to participate, the team in site 6 proactively contacted the national EPOCH trial team to request to be included:

And then EPOCH was happening so I . . . we hadn't actually heard about it and I am not quite sure how. I mean, we nearly missed the boat so I ended up contacting [national EPOCH lead] directly and just saying 'can we join in?'.

Intensive care consultant

When comparing itself to the national standards at activation, the site was not surprised to be doing well in ICU admissions, where much of its prior efforts had focused. Participants expressed more surprise at their levels of goal-directed fluid therapy, which also stood at 90% compliance. Overall, though, process compliance at activation in the EPOCH trial was similar to the other sites: in some areas their compliance was consistently high, in other areas they were 'not doing so good'. The local team in site 6 planned to approach the issue of high-risk patients differently to the EPOCH trial. Rather than relying on P-POSSUM scoring immediately before and after surgery, they planned to introduce a risk assessment earlier in the patient journey, which they hoped would speed up the process of patient identification, risk stratification and time to theatre.

Use of plan-do-study-act

Once again, in site 6, the EPOCH trial team did not use the formal PDSA methodology. Instead, they consciously opted for an iterative, informal process that followed the basic planning cycle:

The only thing is we are not being particularly good at the PDSA cycle but then again . . . Well I suppose we are. We are just not doing it formally. I am just doing it in a way that makes sense to me and I think [person 3] and I think in a relatively similar way and so therefore I suppose we are studying and acting on it. I felt that the presentation at EPOCH was very – you need to do it like this – and part of me thought, 'As a study, I am going to have to do it like that?', and part of me thought, 'That won't work with my brain', so I have carried on and done it in a way that works and makes sense to me.

Intensive care consultant

This meant, for example, that explicit measurement and specific goals were not included in site 6's improvement cycles:

I understand defining how you will know it has improved. I can see that but . . . it doesn't become a continuous process, does it? It becomes finite. I think there is a danger of if you don't set a target that you do it too slowly but actually most people who lose weight do it gradually and sustained. They don't set themselves a target for 2 weeks and then stop. I just found it a bit didactic and that is not me.

Intensive care consultant

Nonetheless, site 6, like other sites, ventured on various experimentation journeys that were planned and controlled, even if not strictly following the PDSA prescription.

Team

A surgical consultant and an anaesthetic consultant constituted the core team. Both had led efforts to improve emergency surgery prior to the EPOCH project. They mobilised help for specific tasks from others: research nurses to identify emergency laparotomies and ensure process compliance and data completion; the local sepsis lead to co-ordinate implementation of Sepsis 6 (a bundle of six interventions that may improve outcomes for patients with sepsis) and amending surgical documentation; and the local anaesthetic lead to conduct case reviews for feeding back to colleagues about their performance. In notable contrast to, for example, site 1, where EPOCH trial-specific efforts were a part of wider initiatives to improve surgery, or site 4, where pathways for various conditions were to be 'hardwired' into processes through the introduction of the electronic patient record, here efforts were highly focused on the narrow patient group whose data were included in the trial, with these patients picked out to ensure both process compliance and data entry.

Use of data

Throughout the project, the EPOCH trial team mainly used NELA data translated into run charts provided by the national EPOCH trial team. They used these data to generate buy-in at presentations and also to target individual clinicians more specifically about their performance, through e-mails. During the implementation process the EPOCH trial team also used other sources of data, such as existing theatre logbooks, patient case notes and everyday conversations with colleagues, to inform their actions:

[We asked theatre co-ordinators] please remember to ask for the P-POSSUM score [before booking theatres] and they have done it and we haven't really had to remind them very much at all. [We know it is happening] because it is in the book. Our theatre booking form says laparotomy and next to it there is an 'any other information' box, which there is for everything, and people are just writing P-POSSUM underneath it and putting the percentage in. I think when it started a lot of the time the registrars were phoning up and saying I have a laparotomy to do and the surgical nurses were saying what is the P-POSSUM score and actually the information I get from talking to them and just feedback is now they are phoning up and saying I have got a laparotomy and the P-POSSUM score is . . . it seems to be happening. That seems to have worked so that is the first bit.

Intensive care consultant

Engagement

At activation, the local team felt that there was a recognition in the hospital that emergency surgery needed improvement. As in site 5, there was a sense that the local organisational and professional culture was conducive to change and the work already achieved through pre-EPOCH trial audits and process changes had readied the way for further improvement:

It has been recognised for some time in the trust that it is not done very well and the emergency patient is a bit of a second-class citizen compared with the elective cases.

Consultant surgeon

Correspondingly, in the course of the trial, the local team were able to actively engage with numerous groups, including anaesthetic and surgical colleagues at joint meetings, which saw colleagues in both groups coming forward to make offers of support and help and also prompted active work from junior colleagues. More broadly, they engaged with the local sepsis lead, and worked closely on amending surgical documentation to incorporate sepsis identification and treatment; with nursing and administrative colleagues responsible for surgical documentation and for theatre booking, to incorporate P-POSSUM scoring into the process; and with radiology colleagues, to discuss how to improve access for emergency laparotomy patients to computerised tomography scans, with a view, ultimately, to negotiating a service-level agreement with radiology. Unusually, the approach adopted in site 6 also involved composing a feedback e-mail to individual clinicians, informing them about their compliance with the pathway and 'stating the bits they did well and suggesting areas that potentially could be improved, in a fairly soft way'

(intensive care consultant). This was meant to be an extension of an existing practice of chasing people about inputting the NELA data. Data for the feedback were collected from NELA and patient case notes:

... it is almost a bit of an audit and a lot of it is to just nudge people in the right way and ensure that they are aware of what we are not quite achieving and the last bit really is feedback on how the patient is doing because I think sometimes people don't realise how long these patients spend on intensive care in hospital afterwards or what complications they get.

Intensive care consultant

However, the team in site 6 was also judicious in how it sought to engage participants. Central to the approach was also strategically avoiding some arguments that the team felt would be hard to have because they might slow their efforts to engage others:

There are some parts of it that I have no intention of trying to enforce. To be honest, the goal-directed therapy was one of them, but we are doing it anyway, but I was never going to have that argument because I felt it was the wrong argument to have because some people have very strong views on it and it would become a negative association with EPOCH and laparotomies. So that was one I wasn't going to have. [Also] things like protective ventilation, I am not sure I am going to have that argument. I am going to try and do that in a slightly different way.

Intensive care consultant

All in all, the team in site 6 were notable for the breadth and intensiveness of the approaches they used to attempt to put the EPOCH trial pathway into practice. They were self-critical of their efforts, 'Now whether we could have got more people involved to speed that through, possibly'(intensive care consultant), but appeared to have further built on the foundations provided by prior work to encourage and cajole improved compliance with carefully selected elements of the pathway.

Discussion

Although covering only a small subsample of the sites that participated in the EPOCH trial, our ethnographic study highlights the range of background conditions, team composition and approaches to change taken by participating sites. In particular, it demonstrates that superficial similarities of approach could hide major differences in practice. On the face of it, all six sites adopted the four key pillars of the EPOCH trial's programme theory faithfully. However, for each pillar, there were important divergences. The approaches taken in all six sites differed from each other and from some of the prescriptions provided by the core EPOCH trial team. Differences in team composition, most notably the lack of surgical engagement in sites 1 and 3, could significantly diminish teams' sense of ability to improve, and the groups they were able to engage. The EPOCH trial's insistence on active engagement from anaesthesia, critical care and surgery thus appears valid, even if not all participating sites were able to achieve it. The ways in which data were used varied substantially, with some teams taking time to process run charts from their NELA entries, whereas others were more reliant on less formal and perhaps less valid, though sometimes highly persuasive, sources of knowledge, such as informal intelligence on the key deficits in existing routines. Site leads engaged with a wide variety of colleagues, often going far beyond the key stakeholders in surgery, anaesthesia and critical care suggested by the EPOCH trial team. Who was engaged depended far more on existing relationships and what was realistic locally than on systematic efforts to recruit the same groups of stakeholders across sites. Most notably, no site followed the PDSA approach to the letter, with most seeing it instead as a guiding framework for their improvement efforts. Even in the published literature, faithful adherence to the PDSA approach has been shown to be the exception rather than the rule.⁴³ The fact that participating teams in the EPOCH trial seemed to have struggled to adhere to the approach appears to reflect, in part, a difficulty in apprehending the detail of the approach, as reflected, for example, in site 6 lead's confession that PDSA cycles were something that 'will not work with my brain', and in part, to a lack of time to dedicate to detailed planning, review and adaptation, with several sites using the PDSA

'ethos' as a rubric to guide the structure of irregular team meetings to review progress. Apparent in several sites was a rush to quick solutions, or at least to pluck the low-hanging fruit that might result in immediate improvement, and this seemed to reflect a concern that more a more tempered, sustained approach to improvement would be difficult to maintain (as, for example, in site 1) and, thus, a reluctance to engage in time-consuming activities, such as PDSA cycles and bespoke data collection and analysis.

The EPOCH trial's enactment, then, was far from consistent in participating sites, aligning with the general principles behind the programme (around use of data, deploying small cycles of change, engaging widely and making use of the team available), but straying away from some of the more specific methods prescribed. Of course, much of this might be seen as legitimate adaptation to local context and, indeed, slavishly following what was suggested by the core EPOCH trial team would certainly not have been appropriate: implicit in the theory of change was a reliance on the local knowledge and industry of the local leads and their teams. As noted in *Chapter 4*, however, adaptation was greater than anticipated and, more broadly, these local leads and teams started from starkly different positions in terms of the organisational contexts they faced. Some appeared well placed, with evident engagement from the crucial groups and prior efforts that had galvanised action. However, starting point in terms of compliance with process measures and mortality rates, on their own, did not seem to be strongly associated with sites' readiness for change. Relatively good performance on these measures could act as a deterrent that foiled teams' efforts to engage their colleagues, as for example in sites 3 and 4, where leads found it difficult to generate the 'burning platforms' that might convince their peers of the need for improvement.³¹ What seemed more important was the coherence and receptivity of the local groups of professionals. In sites 5 and 6, leads were pleasantly surprised with the warmth of the reception they received when introducing the EPOCH trial, characterised by neither scepticism towards the deficiencies identified, nor hostility towards the improvements proposed. This gave leads in these sites a strong platform from which they launched a wide range of efforts to improve the pathway.

Of course, the tenacity and hard work of the teams themselves were also important in this process. It was clear, for example, that leads in sites 5 and 6 had more time to devote to the EPOCH trial than their counterparts in sites 1 and 3. However, this itself also seemed to be, in part, a product of local context and culture. In the absence of surgical engagement in site 3, it was difficult for the lead to even get improvement activities started. In contrast, in site 6, the individualised approach to feedback developed by leads could only have worked in a setting where this kind of constructively critical input was accepted and seen as legitimate. As the lead in site 3 put it:

I am sure that there are some places, that do a lot better than us and they are driven by very strong team of sort of core clinicians who are, who just push it through, who must have some time at some point in the week to go round and keep on top of it, as opposed to playing catch-up.

What was apparent was that, in the absence of a conducive local culture, no amount of effort on the part of leads was likely to have much of an impact. The impact of wider organisational change was also critical, as recognised in the existing literature on QI.^{31,70} Sometimes it could be positive, adding weight to work undertaken specifically for the EPOCH trial, as in site 1. Other times, it could be an impediment, as in site 3, and perhaps most notably in site 4. Here, the ongoing efforts to hardwire a new emergency surgery pathway into organisational routine through the introduction of an electronic patient record system led to inertia on the part of the local leads, who felt assured that this would, in time, generate improvements in process, in a way that no amount of encouragement and engagement could hope to. Yet the new system did not materialise within the time frame of the EPOCH trial and, as in site 3, the situation at the end of the trial appeared to have changed little from the start. Dependence on system-based fixes alone for improvement is ill-advised.⁷¹

Yet, this also points to another important finding from the ethnographic study, again not necessarily visible from the numbers alone, with a perhaps more positive message. In all six sites, the EPOCH trial was just one chapter in a longer-term narrative of change. The sites were at different stages in this narrative and it

was not a simple, linear story of gradual improvement, indeed for some sites (notably site 3), merely holding onto the achievements of the past was a large enough task in itself. Likewise, although from the trial perspective, the EPOCH trial represented a single, major concerted intervention, with fixed objectives and time scales for change, for the sites themselves, it was one input among many, and often not an especially large-scale or significant one, in the context of other past, concurrent and future changes. Sometimes the timing of the EPOCH trial activation was propitious; other times, it was unfortunate. However, in several sites, there was a sense from observation and interviews that the EPOCH trial had provided a foundation for further work in the future (e.g. by helping to consolidate data collection activities or by initiating projects that would not quite come to fruition within the time frame of the trial). Inevitably, given the stepped-wedge design, some sites had longer to instigate changes than others. A consistent message from studies of health-care QI is that it is not a rapid process and will inevitably involve setbacks, as well as wins.^{31,63} Regardless of the population-level achievements against the outcome measures defined in the trial, what was evident at the local level in many sites was a longer-term picture of improvement, to which the EPOCH trial had contributed, sometimes minimally, sometimes importantly.

Chapter 6 Health economic analysis

Introduction

More than one million adult patients undergo inpatient non-cardiac surgery in the NHS each year, with an estimated mortality of between 1.6% and 3.6%.^{72,73} However, patients undergoing emergency surgery are exposed to a much greater risk of death.^{74,75} The 90-day mortality of patients who undergo emergency laparotomy, a major surgical procedure to treat an acute and often life-threatening problem with the gut or other abdominal organ, was 25%.³ There is an urgent need to consider interventions that might improve survival for patients receiving emergency laparotomy.

An integrated care pathway was proposed to improve the quality of care for these patients, representing a best-practice standard of perioperative care deliverable in all NHS hospitals.⁸ To implement this integrated pathway, the evidence-based QI approach was used to change the current practice and culture of care for the patient group. The QI method has been shown to be associated with improved clinical outcomes for surgical patients.^{9,76} A QI intervention has been designed to be implemented into the integrated care pathway to improve patient outcomes following emergency laparotomy; it included QI training and support for the local QI leads nominated by each participating hospital to develop and implement the action plan tailored for each hospital's needs. The effectiveness of this QI intervention was evaluated by the enhanced perioperative care for high-risk patients (the EPOCH trial).

The intervention was expected to improve patients' health-related quality of life (HRQoL) and reduce mortality, but also to lead to increased costs in the short-term perspective when compared with current clinical practice. In order to establish whether or not this intervention has the potential for widespread implementation, its cost-effectiveness needs to be assessed to determine if the gain in health outcomes justifies any increased costs. Therefore, the aim of this study was to assess the cost-effectiveness of the QI intervention for patients undergoing emergency laparotomy.

Methods

Overview

The analysis was undertaken from a UK health service perspective and costs are expressed in Great British pounds at 2016/17 prices. Health outcomes were estimated in terms of quality-adjusted life-years (QALYs). As the intervention was expected to have long-term effects on quality and duration of life, its cost-effectiveness was assessed both within the trial period and over the lifetime of patients.

Data sources

To assess cost-effectiveness within the trial period, survival outcome, health-care resource use and HRQoL were estimated using individual patient data on a subsample of patients from the trial. In eligible patients in 8 of 93 participating hospitals, data on hospital readmissions after surgery, outpatient appointments and primary care health-care resource use were collected using a questionnaire at the end of the trial period (i.e. 180 days after surgery). HRQoL data were collected using the EuroQol-5 Dimensions, three-level version (EQ-5D-3L) before surgery (baseline) and at 90 days and 180 days after surgery. Assumptions were employed to extrapolate costs and health outcomes beyond the 180-day period of the trial. Patient-level data were collected and collated by the NELA in all participant hospitals from the start of the trial. These data were then linked to the ONS, HES and the Information Services Division of NHS Scotland databases, using patient identifiers to allow collation of outcome data, including mortality and hospital readmissions.

Resource use and costs

Resource use associated with inpatient stay after surgery, during the trial period, were extracted from the NELA database and then combined with NHS reference unit costs to calculate the costs of inpatient stay after surgery.⁷⁷ Costs of readmissions and costs of outpatient and primary care were calculated using resource use reported in the questionnaire and the *NHS Reference Costs 2015 to 2016*⁷⁷ and Personal Social Services Research Unit *Unit Costs of Health and Social Care 2016*.⁷⁸ Unit costs are presented in Table 16. Costs of the QI intervention were estimated by the clinical trial team, by considering the resource use associated with the QI intervention, and is assumed to be the same across all hospitals.

Health-related quality of life

The EQ-5D-3L responses were transformed into HRQoL weights using the UK tariff,⁷⁹ on a scale in which 0 represents death and 1 represents full health. The HRQoL weights between measurement points were multiplied by the time between those points, along with mortality data from the ONS, to estimate total QALYs gained over the trial period, using the area under the curve method and linear interpolation between time points,⁸⁰ and then adjusted for baseline EQ-5D-3L, to calculate the incremental QALYs associated with the QI intervention.

Missing data

Multiple imputation was performed to replace each missing observation with a set of plausible imputed values, including EQ-5D-3L scores and three cost variables, following the method recommended by Faria *et al.*⁸¹ The imputation was implemented separately by randomised treatment allocation. Prediction mean matching was used to ensure imputed values were in the appropriate range (e.g. no negative costs or EQ-5D-3L scores of > 1). We used multiple imputation by chained equations in Stata® (StataCorp LP, College Station, TX, USA), using the command 'mi impute chained', and Rubin's rules were implemented using the command 'mi estimate', also in Stata. The multiply imputed data sets were then used in cost-effectiveness analysis.

TABLE 16 Unit costs of resource use

Resource use item	Unit	Unit cost (£) ^a	Source
Inpatient care			
Inpatient stay after surgery	Day	359	NHS NEL_XS: major general abdominal procedures, 19 years and over, with CC Score 0 [FF51E]
Hospital readmission	Event	3434	NHS EL: major general abdominal procedures, 19 years and over, with CC Score 0 [FF51E]
Outpatient and primary care			
Outpatient appointment	Visit	130	NHS Total Outpatient Attendances [service code: 100]
A&E visit	Visit	148	NHS Total Outpatient Attendances [service code: 180]
GP surgery visit (GP)	Visit	38	<i>Unit Costs of Health and Social Care 2016</i> ⁷⁸
GP surgery visit (nurse)	Visit	21	<i>Unit Costs of Health and Social Care 2016</i> ⁷⁸
GP home visit	Visit	120	<i>Unit Costs of Health and Social Care 2016</i> ⁷⁸
Stoma nurse visit	Visit	55	<i>Unit Costs of Health and Social Care 2016</i> ⁷⁸
Occupational therapy	Visit	45	<i>Unit Costs of Health and Social Care 2016</i> ⁷⁸
Physiotherapy	Visit	49	NHS Total Outpatient Attendances [service code: 650]
Psychotherapy	Visit	194	NHS Total Outpatient Attendances [service code: 713]
Dietetics	Visit	71	NHS Total Outpatient Attendances [service code: 654]

CC, complication and comorbidity; GP, general practitioner.

^a Rounded up to nearest Great British pound.

Analysis

For the within-trial analysis, costs and QALYs were calculated per patient and then analysed with the seemingly unrelated regression model to estimate the incremental mean costs and QALYs.⁸² If the QI intervention was associated with QALY gain and increased costs, the incremental cost-effectiveness ratio (ICER) was generated and then compared against the threshold values of the National Institute for Health and Care Excellence, which is ranged from £20,000 to £30,000 per QALY gained.⁸³ The threshold of £13,000 per QALY, representing the marginal productivity of the NHS estimated by Claxton *et al.*,⁸⁴ was also used here to consider whether or not the QI intervention represents a cost-effective use of resources. The probability of the QI intervention being cost-effective at different thresholds of willingness to pay was calculated^{85,86} and represented visually on a cost-effectiveness acceptability curve.^{87,88}

Cost-effectiveness over the lifetime horizon was estimated by incorporating assumptions of HRQoL after the trial period, general population survival data from the ONS and health-care costs from the literature. HRQoL of patients alive at the end of the trial period were assumed to return from their EQ-5D-3L at day 180 to the population average for someone of their age and sex over the next 3 years, and subsequently have the population age- and sex-adjusted EQ-5D-3L scores until death.⁸⁹ QALYs over the 3 years after the trial were estimated based on linear interpolation, which is standard practice.⁹⁰ Patients whose life expectancy was < 3 years were assumed to have the HRQoL at day 180 until death. Life expectancy was estimated to match the population average based on their age and sex, reported in the ONS *National Life Tables, United Kingdom: 2012–14*.⁹¹ We assumed that patients alive were attributed average health-care costs by age and sex over the rest of their modelled lives.⁹² The expected lifetime costs and QALYs were discounted using 3.5% per annum in line with UK guidelines.⁸³ Cost-effectiveness acceptability curves were also presented to reflect decision uncertainty.

Scenario analyses were conducted to explore the cost-effectiveness of the intervention in the whole trial population. As data on inpatient stay after surgery were available for the EPOCH trial population, costs of inpatient stay were calculated using the data directly from the trial, and costs of readmissions and costs of outpatient and primary care were estimated using generalised linear model (GLM) regressions based on the subsample. Baseline HRQoL and HRQoL change from randomisation were estimated using regression models developed based on the subsample and then combined with the mortality data for the EPOCH trial population, to calculate QALYs over the trial period using area under curve and linear interpolation method. A GLM with gamma distribution and log-link was employed to estimate the HRQoL decrement due to the emergency laparotomy before surgery, accounting for relevant covariates and the distribution of the data.⁹³ A panel data approach was then employed to estimate changes in HRQoL after randomisation, differentiating between intervention and usual care.⁹⁴ The assumptions of long-term effects used in primary analysis were also applied to estimate costs and QALYs over the lifetime horizon.

Sensitivity analyses were conducted to assess the impact of different assumptions of HRQoL beyond trial period on the trial results:

- Patients alive at the end of trial followed the group mean EQ-5D-3L scores until death.
- Patients alive at the end of trial had their individual HRQoL at day 180 until death.

Complete-case analyses were also conducted as part of sensitivity analyses, using data on patients with complete data on all variables at all follow-up points.

Subgroup analyses were performed to establish whether or not cost-effectiveness varies between high- and low-risk patients, defined by P-POSSUM score (risk for morbidity and mortality). All statistical analyses were performed using Stata 14.

Results

Patient characteristics

The characteristics of patients in the subsample and the EPOCH trial population are summarised in *Table 17*. A total of 680 eligible patients were included in the subsample, with 265 patients in the QI intervention group and 415 patients in the usual care group. The EPOCH trial population included 15,856 eligible patients (8482 in the usual care group and 7374 in the QI group). Overall, there were no marked differences in age, sex, indication for surgery and P-POSSUM scores between the intervention and usual care groups in both the subsample and the EPOCH trial sample, but there were more older patients, females and patients at high risk of morbidity and mortality defined by P-POSSUM score in the full EPOCH trial population than in the subsample. In the subsample, patients in the QI group had higher baseline EQ-5D-3L scores than the usual care group (0.262 vs. 0.168).

Missing data

The missingness of data on health-care resource use and EQ-5D-3L collected in the subsample is summarised in *Table 18*. Missing data on inpatient care ranged from < 1% to 4.15% and there were more missing values in outpatient and primary care (23.4% in the usual care group and 26.4% in the intervention group). For EQ-5D-3L scores, there were more missing values at follow-ups than baseline (ranging from 19.0% to 30.6%).

TABLE 17 Baseline characteristics of patients in subsample and EPOCH trial population

Patient characteristics	Subsample (N = 680)		EPOCH trial population (N = 15,856)	
	Usual care (N = 415)	QI (N = 265)	Usual care (N = 8482)	QI (N = 7374)
Age (years), mean (SD)	66.7 (13.3)	66.4 (12.5)	68.5 (13.1)	68.2 (13.1)
Female, n (%)	204 (49.2)	131 (49.4)	4547 (53.6)	3935 (53.4)
Indication for surgery, n (%)				
Peritonitis	20 (4.82)	7 (2.64)	352 (4.15)	251 (3.40)
Perforation	39 (9.40)	27 (10.19)	765 (9.02)	691 (9.37)
Intestinal obstruction	189 (45.54)	134 (50.57)	3834 (45.20)	3374 (45.76)
Haemorrhage	15 (3.61)	3 (1.13)	213 (2.51)	149 (2.02)
Ischaemia	18 (4.34)	5 (1.89)	364 (4.29)	332 (4.50)
Abdominal infection	8 (1.93)	8 (3.02)	296 (3.49)	239 (3.24)
Other	35 (8.43)	13 (4.91)	523 (6.17)	471 (6.39)
Multiple indications	89 (21.45)	68 (25.66)	2122 (25.02)	1862 (25.25)
Missing	2 (0.48)	0	13 (0.15)	5 (0.07)
P-POSSUM score, median (IQR)	4.9 (2.0–14.1)	5.8 (2.7–18.1)	7.6 (2.9–22.7)	7.4 (2.9–22.9)
Baseline EQ-5D-3L score, mean (SD)	0.168 (0.448)	0.262 (0.466)		

IQR, interquartile range; SD, standard deviation.

TABLE 18 Missing data in resource use and HRQoL collected in the subsample

Resource use and HRQoL	Usual care (N = 415), n (%)	QI (N = 265), n (%)
Inpatient stay after surgery	1 (< 1)	0
Number of hospital readmissions	17 (4.10)	11 (4.15)
Outpatient and primary care	97 (23.4)	70 (26.4)
Baseline EQ-5D-3L score	6 (1.45)	4 (1.51)
90-day EQ-5D-3L score	79 (19.0)	51 (19.2)
180-day EQ-5D-3L score	101 (24.3)	81 (30.6)

Health-care resource use

The health-care resources used by this subsample over the trial period are summarised in *Table 19*.

Costs

Table 20 reports the mean costs of health-care resource use, costs of intervention and total costs per patient in the QI intervention and the usual care groups in the subsample. For the usual care group, the mean costs of inpatient stay after surgery, of readmissions and of outpatient and primary care were £6486, £934 and £797, respectively, and the corresponding costs for the QI group were £7001, £881 and £760. The cost of the QI intervention was estimated to be £32 per patient. The total cost for a patient in the usual care group was £8216, whereas the cost for a patient in QI group was higher, at £8675. During the trial period, the QI intervention was associated with £458 higher costs (95% CI –£812 to £1728). The cost of inpatient stay after surgery is the principal costs driver, accounting for around 80% of the mean total costs.

TABLE 19 Resources used by the subsample over the trial period

Resource use	Usual care (N = 415)			QI (N = 265)		
	Mean (SD)	Median (IQR)	Used by (%)	Mean (SD)	Median (IQR)	Used by (%)
Inpatient care						
Inpatient days	18.1 (20.5)	12 (7–20)	100	19.5 (19.5)	13 (9–21)	100
Number of readmissions	0.27 (1.11)	0 (0–0)	12.8	0.26 (0.79)	0 (0–0)	14.6
Outpatient and primary care						
Outpatient appointment	4.14 (6.19)	2 (1–5)	84.0	3.49 (4.78)	1 (1–4)	83.1
A&E visit	0.41 (1.54)	0 (0–0)	19.8	0.47 (1.78)	0 (0–0)	19.5
GP surgery visit (GP)	2.47 (4.01)	2 (0–3)	67.9	2.35 (3.27)	1 (0–3)	62.1
GP surgery visit (nurse)	1.90 (4.67)	0 (0–1)	44.0	1.41 (2.88)	0 (0–2)	42.6
GP home visit	0	0	0	0.01 (0.10)	0 (0–0)	1.03
Stoma nurse visit	0.33 (1.38)	0 (0–0)	9.1	0.45 (1.59)	0 (0–0)	11.3
Occupational therapy	0.19 (0.73)	0 (0–0)	9.4	0.37 (1.97)	0 (0–0)	9.2
Physiotherapy	0.38 (2.41)	0 (0–0)	8.5	1.10 (6.73)	0 (0–0)	7.2
Psychotherapy	0.14 (1.47)	0 (0–0)	2.2	0.12 (0.87)	0 (0–0)	3.1
Dietetics	0.04 (0.52)	0 (0–0)	0.9	0.12 (1.65)	0 (0–0)	0.5

GP, general practitioner; IQR, interquartile range; SD, standard deviation.

TABLE 20 Costs of resources used by the subsample over the trial period, costs of intervention and total costs

Cost (£) ^a	Usual care (N = 415)		QI (N = 265)	
	Mean (SE)	95% CI	Mean (SE)	95% CI
Cost of inpatient stay after surgery	6486 (360)	5778 to 7194	7001 (431)	6152 to 7850
Cost of readmissions	934 (189)	561 to 1305	881 (170)	547 to 1215
Cost of outpatient and primary care	797 (58)	682 to 912	760 (60)	642 to 878
QI intervention	0		32	
Total costs (£)	8216 (411)	7409 to 9024	8675 (497)	7696 to 9684
Incremental cost (£)			458 (–812 to 1728)	

SE, standard error.

^a Rounded up to nearest Great British pound.

Health-related quality of life

The EQ-5D-3L scores and QALYs are summarised in *Table 21*. At baseline, the QI patients had higher EQ-5D-3L scores than those in the usual care group, but the follow-up EQ-5D-3L scores were similar between the two groups. The mean QALYs over the 180 days after surgery were 0.274 (95% CI 0.260 to 0.288) per patient in the usual care group and 0.287 (95% CI 0.269 to 0.306) per patient in the QI intervention group. After adjusting for baseline EQ-5D-3L scores, the difference in QALYs between the two groups was minimal, with fewer QALYs in the QI intervention group (mean -0.002 , 95% CI -0.021 to 0.017).

Base-case analysis

The cost-effectiveness results within the trial period and over the lifetime horizon in the subsample are presented in *Table 22*. Within-trial analyses showed that the QI intervention was associated with higher total costs, but the result was not significant (mean difference £458, 95% CI $-£812$ to $£1728$).

TABLE 21 EuroQol-5 Dimensions, three-level version scores of the subsample and QALYs over the trial period

	Usual care (N = 415)		QI (N = 265)	
HRQoL	Mean (SE)	95% CI	Mean (SE)	95% CI
Baseline EQ-5D-3L	0.168 (0.022)	0.124 to 0.211	0.262 (0.029)	0.205 to 0.319
90 days EQ-5D-3L	0.649 (0.017)	0.615 to 0.684	0.669 (0.023)	0.623 to 0.715
180 days EQ-5D-3L	0.706 (0.018)	0.670 to 0.741	0.677 (0.025)	0.628 to 0.725
QALYs	0.274 (0.007)	0.260 to 0.288	0.287 (0.009)	0.269 to 0.306
Adjusted incremental QALYs ^a			−0.002 (−0.021 to 0.017)	
SE, standard error.				
a Adjusted for baseline EQ-5D-3L scores.				

TABLE 22 Cost-effectiveness analysis in the subsample over trial period

Cost outcome	Usual care (N = 415)	QI (N = 265)
180 days' costs (£), mean (95% CI)	8216 (7409 to 9024)	8675 (7696 to 9684)
Incremental costs (£), mean (95% CI)		458 (-812 to 1728)
180 days' QALYs, mean (95% CI)	0.274 (0.260 to 0.288)	0.287 (0.269 to 0.306)
Adjusted incremental QALYs, mean (95% CI)		-0.002 (-0.021 to 0.017)
ICER (£/QALY)		Dominated
Probability (%) of cost-effectiveness at		
£13,000/QALY		24.0
£20,000/QALY		24.3
£30,000/QALY		24.9
Lifetime costs (£), mean (95% CI)	47,749 (46,063 to 49,436)	49,258 (47,294 to 51,221)
Incremental costs (£), mean (95% CI)		1508 (-1108 to 4124)
Lifetime QALYs, mean (95% CI)	9.434 (8.938 to 9.929)	9.564 (8.986 to 10.143)
Adjusted incremental QALYs, mean (95% CI)		0.131 (-0.640 to 0.903)
ICER (£/QALY)		11,511
Probability (%) of cost-effectiveness at		
£13,000/QALY		51.8
£20,000/QALY		56.3
£30,000/QALY		58.7

After adjusting for baseline EQ-5D-3L scores, the QI and usual care groups showed similar QALYs gained over the trial period (mean difference -0.002 , 95% CI -0.021 to 0.017), although QALYs were fewer in the QI group. The probability of being cost-effective at £13,000, £20,000 and £30,000 per QALY threshold was 24.0%, 24.3% and 24.9%, respectively. The cost-effectiveness acceptability curve is shown in *Figure 10*. Therefore, the QI intervention was dominated by the current practice and it does not appear cost-effective in the subsample over the 180 days of the trial period.

In the lifetime perspective, the QI intervention was associated with higher costs (mean difference £1508, 95% CI $-\text{£}1108$ to $\text{£}4124$) and more QALYs (mean difference 0.131 , 95% CI -0.640 to 0.903), which led to an ICER of £11,511 per QALY, lower than the three cost-effectiveness threshold values considered. The probabilities of being cost-effective at £13,000, £20,000 and £30,000 per QALY threshold were 51.8%, 56.3% and 58.7%, respectively (see *Figure 10*). Therefore, the QI intervention may be cost-effective for patients in the subsample over the longer term, but this is subject to substantial uncertainty.

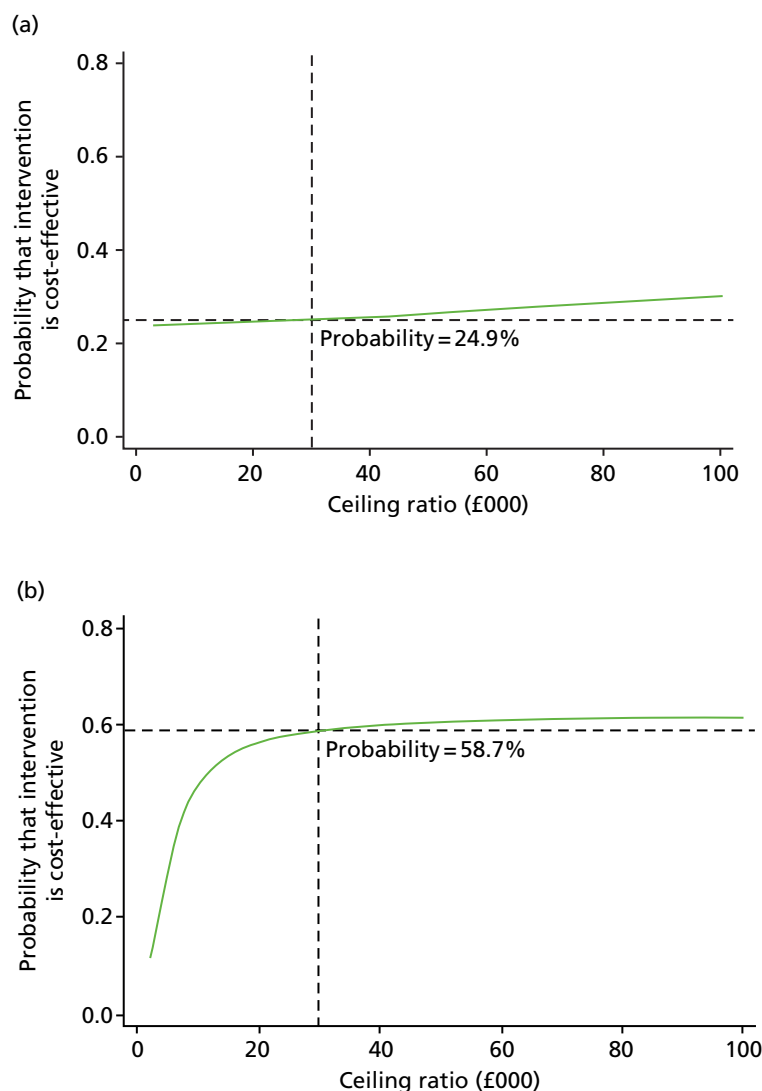


FIGURE 10 Cost-effectiveness acceptability curve for the QI intervention over (a) trial period; and (b) lifetime horizon.

Regression

Cost and health-related quality-of-life regression

The results of marginal effect using GLM regression, with a gamma family and log-link on costs of readmissions and costs of outpatient and primary care, are presented in *Table 23*. Only patients with hospital readmissions were included in the regression model for costs of readmissions, which resulted in only 133 observations. The marginal effect using GLM regression with a gamma family and log-link on HRQoL decrement at baseline is summarised in *Table 24*. At baseline, HRQoL decrement was higher for females, resulting in worse baseline HRQoL.

In the model of HRQoL change from randomisation to follow-up, binary covariates were included to represent whether the EQ-5D-3L scores were taken at 90 days or 180 days and an interaction term for treatment group. The model was fitted using generalised least squared random-effects estimators (Stata command 'xtreg'). The number of patients with EQ-5D-3L data in the follow-up period was 630 and the number of observations was 1244. The results of HRQoL from randomisation are presented in *Table 25*. At 90 days after surgery, the HRQoL of patients in the QI intervention and usual care groups improved, but it did not improve as much in the QI intervention group as it did in the usual care group (0.675 vs. 0.682). At 180 days after surgery, the HRQoL of both groups improved further, although it improved more in the usual care group than in the QI intervention group (0.760 vs. 0.702). The baseline EQ-5D-3L scores were negatively associated with change in HRQoL, showing that the higher baseline EQ-5D-3L scores are associated with a smaller HRQoL gain.

TABLE 23 Marginal effect on costs of readmission and costs of outpatient and primary care

Using GLM with a gamma family and log-link	Cost of readmissions (£)	Cost of outpatient and primary care (£)
QI intervention	240	-4
Age (years)	-119*	-7*
Female	2025	-23
Intervention time period (days)	55	-8
* $p < 0.05$.		

TABLE 24 Marginal effect on HRQoL decrement at baseline

Using GLM with a gamma family and log-link	HRQoL decrement
Age (years)	-0.001
Female	0.086*
Indication for surgery	
Peritonitis	Ref
Perforation	-0.149
Intestinal obstruction	-0.238*
Haemorrhage	-0.212
Ischaemia	-0.102
Abdominal infection	-0.261*
Other	-0.207*
Multiple indications	-0.148*
* $p < 0.05$. Ref, reference.	

TABLE 25 Changes in HRQoL from randomisation based on the subsample

Covariate	HRQoL change
Randomised to usual care group (90 days of follow-up)	0.682
Randomised to QI group (90 days of follow-up)	0.675
Randomised to usual care group (180 days of follow-up)	0.760
Randomised to QI group (180 days of follow-up)	0.702
Baseline EQ-5D-3L score	-0.857

Scenario analysis

Using the regression models, the estimated costs, EQ-5D-3L scores and QALYs for the EPOCH trial population are presented in *Table 26*. Within the trial period, the QI intervention was associated with higher costs, at £296, but fewer QALYs, at -0.013. Over the lifetime horizon, the QI intervention was also associated with higher costs, at £601 and more QALYs, at 0.019, resulting in an ICER of £31,632 per QALY, which is higher than the highest cost-effectiveness threshold value of £30,000 per QALY. Therefore, the intervention was probably not cost-effective for the EPOCH trial population.

TABLE 26 Estimated costs, EQ-5D-3L scores and QALYs for the EPOCH trial population over the trial period and over the lifetime horizon

Outcome	Usual care (N = 8482)		QI (N = 7374)	
	Mean (SE)	95% CI	Mean (SE)	95% CI
Within-trial period				
Costs of inpatient stay after surgery	6226 (75)	6079 to 6372	6348 (79)	6194 to 6502
Costs of readmissions	1116 (23)	1071 to 1160	1309 (28)	1254 to 1364
Costs of outpatient and primary care	789 (1)	787 to 791	738 (1)	736 to 740
QI intervention	0		32	
Total costs (£)	8130 (79)	7975 to 8284	8426 (85)	8259 to 8592
Incremental costs (£)			296	
Baseline EQ-5D-3L score	0.200 (0.001)	0.200 to 0.203	0.203 (0.001)	0.201 to 0.204
90 days EQ-5D-3L score	0.595 (0.003)	0.589 to 0.600	0.589 (0.003)	0.583 to 0.595
180 days EQ-5D-3L score	0.631 (0.003)	0.624 to 0.638	0.589 (0.003)	0.582 to 0.595
QALYs	0.239 (0.001)	0.236 to 0.241	0.226 (0.001)	0.223 to 0.228
Adjusted incremental QALYs			-0.013	
ICER (£/QALY)			Dominated	
Over lifetime horizon				
Lifetime costs (£)	42,779 (221)	42,345 to 43,213	43,380 (237)	42,915 to 43,845
Incremental costs (£)			601	
Lifetime QALYs	8.118 (0.059)	8.002 to 8.234	8.131 (0.063)	8.007 to 8.255
Adjusted incremental QALYs			0.019	
ICER (£/QALY)			31,632	

SE, standard error.

Sensitivity analysis

Assumptions of health-related quality of life after trial

Assuming patients alive at the end of the trial followed the group mean EQ-5D-3L scores or individual EQ-5D-3L scores at day 180 after surgery until death, the QI intervention was associated with fewer QALYs gained and dominated by the current practice (Table 27).

Complete-case analysis

Table 28 presents the characteristics of patients in the subsample and those with complete data, who were alive at the end of the trial. A total of 472 patients were included in the complete-case analysis, with 176 of these in the QI group. The cost-effectiveness results in the complete-case analysis are summarised Table 29. Within the trial period, the QI intervention was associated with lower total costs (mean –£165, 95% CI –£1603 to £1274) and fewer QALYs (mean –0.005, 95% CI –0.023 to 0.013) than the usual care group, resulting in an ICER of £33,000 per QALY. In the long-term perspective, QI intervention was associated with higher costs (mean £123, 95% CI –£1840 to £2087) and fewer QALYs (mean –0.004, 95% CI –0.752 to 0.744) than the usual care group. For both the within-trial and the lifetime horizon analyses, the cost-effectiveness probability at £13,000, £20,000 and £30,000 per QALY was around 50%, indicating substantial uncertainty.

TABLE 27 Cost-effectiveness analysis in the subsample over the trial period using different assumptions on HRQoL

	Usual care (N = 415)	QI (N = 265)
Lifetime costs (£), mean (95% CI)	47,749 (46,063 to 49,436)	49,258 (47,294 to 51,221)
Assumption 1: mean HRQoL at end of trial		
Incremental costs (£), mean (95% CI)		1508 (–1108 to 4124)
Lifetime QALYs, mean (95% CI)	8.668 (8.235 to 9.100)	8.504 (8.016 to 8.992)
Adjusted incremental QALYs, mean (95% CI)		–0.152 (–0.818 to 0.514)
ICER (£/QALY)		Dominated
Probability (%) of cost-effectiveness at		
£13,000/QALY		17.1
£20,000/QALY		22.4
£30,000/QALY		25.8
Assumption 2: individual HRQoL at end of trial		
Incremental costs (£), mean (95% CI)		1508 (–1108 to 4124)
Lifetime QALYs, mean (95% CI)	9.763 (9.126 to 10.399)	9.203 (8.411 to 9.996)
Adjusted incremental QALYs, mean (95% CI)		–0.694 (–1.707 to 0.320)
ICER (£/QALY)		Dominated
Probability (%) of cost-effectiveness at		
£13,000/QALY		4.7
£20,000/QALY		6.0
£30,000/QALY		6.9

TABLE 28 Baseline characteristics of patients in subsample and complete cases

Patient characteristics	Subsample (N = 680)		Complete case (N = 472)	
	Usual care (n = 415)	QI (n = 265)	Usual care (n = 296)	QI (n = 176)
Age (years), mean (SD)	66.7 (13.3)	66.4 (12.5)	66.3 (12.7)	66.4 (12.1)
Female, n (%)	204 (49.2)	131 (49.4)	134 (48.3)	90 (51.1)
Indication for surgery, n (%)				
Peritonitis	20 (4.82)	7 (2.64)	12 (4.05)	5 (2.84)
Perforation	39 (9.40)	27 (10.19)	31 (10.47)	12 (6.82)
Intestinal obstruction	189 (45.54)	134 (50.57)	135 (45.61)	93 (52.84)
Haemorrhage	15 (3.61)	3 (1.13)	14 (4.73)	1 (0.57)
Ischaemia	18 (4.34)	5 (1.89)	9 (3.04)	4 (2.27)
Abdominal infection	8 (1.93)	8 (3.02)	7 (2.36)	5 (2.84)
Other	35 (8.43)	13 (4.91)	24 (8.11)	9 (5.11)
Multiple indications	89 (21.45)	68 (25.66)	62 (20.95)	47 (26.70)
Missing	2 (0.48)	0	2 (0.68)	0
P-POSSUM, median (IQR)	4.9 (2.0–14.1)	5.8 (2.7–18.1)	4.5 (1.9–13.8)	5.3 (2.8–13.7)
Death within 90 days, n (%)	34 (8.19)	16 (6.04)	0	0
Death within 180 days, n (%)	43 (10.36)	23 (8.68)	0	0

IQR, interquartile range; SD, standard deviation.

TABLE 29 Cost-effectiveness analysis in the complete-case analysis over trial period

	Usual care (N = 296)	QI (N = 176)
180 days' costs (£), mean (95% CI)	8050 (7121 to 8979)	7885 (6841 to 8929)
Incremental costs (£), mean (95% CI)		–165 (–1603 to 1274)
180 days' QALYs, mean (95% CI)	0.302 (0.288 to 0.316)	0.312 (0.292 to 0.332)
Adjusted incremental QALYs, mean (95% CI)		–0.005 (–0.023 to 0.013)
ICER (£/QALY)		33,000
Probability (%) of cost-effectiveness at		
£13,000/QALY		55.0
£20,000/QALY		53.1
£30,000/QALY		50.5
Lifetime costs (£), mean (95% CI)	52,297 (51,029 to 53,566)	52,421 (50,998 to 53,843)
Incremental costs (£), mean (95% CI)		123 (–1840 to 2087)
Lifetime QALYs, mean (95% CI)	10.372 (9.903 to 10.841)	10.336 (9.763 to 10.909)
Adjusted incremental QALYs, mean (95% CI)		–0.004 (–0.752 to 0.744)
ICER (£/QALY)		Dominated
Probability (%) of cost-effectiveness at		
£13,000/QALY		48.6
£20,000/QALY		48.9
£30,000/QALY		49.2

Subgroup analysis

High-risk patients

According to P-POSSUM score, patients were categorised into high- and low-risk groups, using the median value. For the high-risk patients, the QI intervention was associated with increased costs and more QALYs within the trial period and the ICER was £119,400 per QALY, which is much higher than the highest cost-effectiveness threshold value of £30,000 per QALY, so it was unlikely to be cost-effective over the trial period. However, over the lifetime horizon, the QI intervention was associated with more QALYs at an increased cost, generating an ICER of £3839 per QALY. The probabilities of being cost-effective at the threshold of £13,000, £20,000 and £30,000 per QALY were around 99%, showing it as very likely to be cost-effective in the high-risk patients. The cost-effectiveness acceptability curves are shown in *Figure 11*. Sensitivity analyses were also conducted for this subgroup and results were similar, but when assumption 2 was used, there was substantial uncertainty around the decision. Therefore, the QI intervention is likely to be cost-effective for patients at high risk of morbidity and mortality before surgery (*Table 30*).

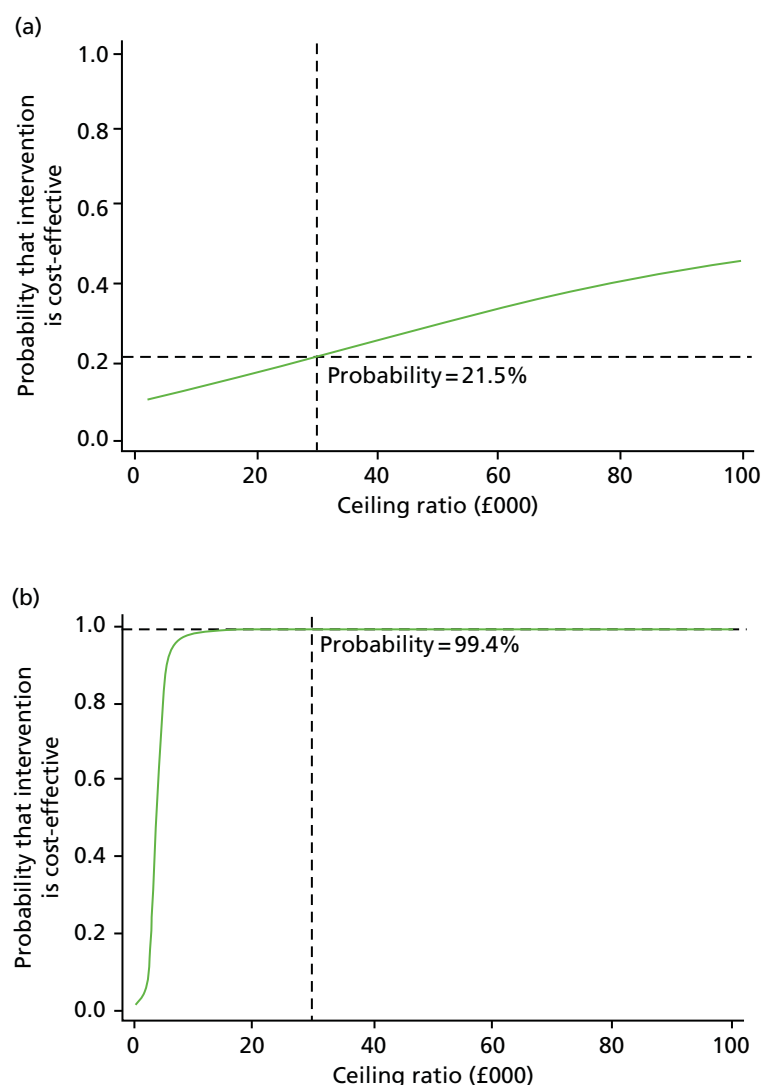


FIGURE 11 Cost-effectiveness acceptability curve for the QI intervention in the high-risk group (a) within the trial period; and (b) over the lifetime horizon.

TABLE 30 Cost-effectiveness analysis in the subgroups defined by P-POSSUM score: high-risk group

High-risk group	Usual care (N = 198)	QI (N = 140)
180 days' costs (£), mean (95% CI)	9322 (8156 to 10,487)	10,515 (9076 to 11,954)
Incremental costs (£), mean (95% CI)		1194 (–627 to 3014)
180 days' QALYs, mean (95% CI)	0.246 (0.224 to 0.267)	0.273 (0.246 to 0.300)
Adjusted incremental QALYs, mean (95% CI)		0.010 (–0.018 to 0.039)
ICER (£/QALY)		119,400
Probability (%) of cost-effectiveness at		
£13,000/QALY		14.3
£20,000/QALY		17.1
£30,000/QALY		21.5
Lifetime costs (£), mean (95% CI)	44,458 (41,609 to 47,307)	49,487 (46,413 to 52,561)
Incremental costs (£), mean (95% CI)		5029 (811 to 9248)
Lifetime QALYs, mean (95% CI)	7.153 (6.509 to 7.796)	8.444 (7.676 to 9.211)
Adjusted incremental QALYs, mean (95% CI)		1.310 (0.317 to 2.303)
ICER (£/QALY)		3839
Probability (%) of cost-effectiveness at		
£13,000/QALY		98.9
£20,000/QALY		99.3
£30,000/QALY		99.4
Assumption 1		
Adjusted incremental QALYs, mean (95% CI)		0.916 (0.038 to 1.793)
ICER		5490
Probability (%) of cost-effectiveness at		
£13,000/QALY		94.1
£20,000/QALY		96.3
£30,000/QALY		97.1
Assumption 2		
Adjusted incremental QALYs, mean (95% CI)		0.396 (–0.919 to 1.712)
ICER (£/QALY)		12,699
Probability (%) of cost-effectiveness at		
£13,000/QALY		50.6
£20,000/QALY		59.1
£30,000/QALY		63.9

Low-risk patients

For the low-risk patients, the QI intervention was associated with lower costs and fewer QALYs within the trial period and over the lifetime horizon, and is unlikely to be considered a cost-effective intervention (*Table 31*).

TABLE 31 Cost-effectiveness analysis in the subgroups defined by P-POSSUM score: low-risk group

Low-risk group	Usual care (N = 217)	QI (N = 125)
180 days' costs (£), mean (95% CI)	7208 (6096 to 8320)	6613 (5385 to 7841)
Incremental costs (£), mean (95% CI)		–595 (–2309 to 1120)
180 days' QALYs, mean (95% CI)	0.299 (0.282 to 0.317)	0.304 (0.279 to 0.328)
Adjusted incremental QALYs, mean (95% CI)		–0.010 (–0.035 to 0.014)
ICER (£/QALY)		59,500
Probability (%) of cost-effectiveness at		
£13,000/QALY		69.1
£20,000/QALY		65.7
£30,000/QALY		61.0
Lifetime costs (£), mean (95% CI)	50,752 (48,909 to 52,596)	49,001 (46,615 to 51,386)
Incremental costs (£), mean (95% CI)		–1752 (–4748 to 1245)
Lifetime QALYs, mean (95% CI)	11.515 (10.885 to 12.145)	10.819 (9.988 to 11.651)
Adjusted incremental QALYs, mean (95% CI)		–0.724 (–1.759 to 0.312)
ICER (£/QALY) (moving from QI to usual care)		2420
Probability (%) of cost-effectiveness at		
£13,000/QALY		10.9
£20,000/QALY		9.8
£30,000/QALY		9.3

Discussion

The analysis shows that, in patients undergoing emergency laparotomy, a QI intervention strategy aiming to improve the quality of care for these patients is unlikely to be cost-effective within the 180 days after the surgery, but may be cost-effective over lifetime horizon, when compared with usual care, particularly in high-risk patients.

The primary analysis of this intervention showed no survival benefit from the QI intervention at either 90 or 180 days after surgery, consistent with our results that the intervention was unlikely to be cost-effective within the trial period. This may be due to the QI programme not being implemented as successfully as expected in the EPOCH trial. According to the ethnographic evaluation of this intervention (see *Chapter 5*), hospital staff often had little or no additional time to implement the change to improve patient care and the QI programme was not delivered as intended due to its complexity and the need for local adaptations in each hospital (see *Chapter 4*). Previous research of QI projects to improve outcomes for patients undergoing emergency abdominal surgery, or in other clinical areas, have delivered mixed results, suggesting that more focused discrete clinical interventions may be more successfully implemented than interventions that include larger numbers of care processes and clinician behaviours. When a lifetime horizon was adopted, this intervention has the potential to be cost-effective at accepted cost-effectiveness thresholds, but there was substantial decision uncertainty around the decision. In additional, sensitivity analyses showed that the QI intervention was not associated with increased QALYs, using different assumptions of HRQoL beyond the trial period. It should be noted that the cost of the QI intervention was very low at £32. The intervention was implemented at the hospital level and the intervention costs included the costs of designing, planning and delivering the QI programme across all participating hospitals. As mentioned previously, the lack of time to implement the change to improve patient care could result in less staff time associated with the

intervention, and therefore underestimate the intervention costs and favour the cost-effectiveness of the intervention. Therefore, the results presented here should be seen as indicative; more evidence is needed when determining whether or not this intervention should be recommended for patients undergoing emergency laparotomy.

It is worth noting that the intervention appears to be cost-effective for patients at high risk of mortality and morbidity before surgery, as defined by P-POSSUM score. High-risk patients seem to benefit more from and be more responsive to this intervention than the low-risk patients. Possibly in clinical practice, these high-risk patients called more attention from the hospital staff and, thus, the objective of the QI intervention was more readily achieved for this subgroup, which lead to higher QALYs within the trial period and better long-term effects. Such findings would have important implications for future research. A similar QI intervention targeting the high-risk group may be more successfully implemented than the intervention for all patients undergoing emergency laparotomy, and may have the potential for widespread implementation. Another issue worth noting is that the intervention was not cost-effective in the whole EPOCH trial population when using the regression models to predict costs and QALYs, as shown in the scenario analysis. In the subsample, the QI group had higher baseline EQ-5D-3L scores than the usual care group. As the regression results showed the strong relationship between baseline HRQoL and HRQoL gain, the HRQoL gain due to the intervention may be masked by the high baseline EQ-5D-3L scores and, consequently, the effect of the intervention may be underestimated when extrapolating results to the full EPOCH trial population. Furthermore, the intervention was not cost-effective. Likewise, the subsample may be not representative of the EPOCH trial population, although the baseline demographic characteristics were similar. Owing to the small sample sizes of the subsample ($n = 680$) and the high-risk patients ($n = 338$), caution should be taken when generalising these results to the EPOCH trial population.

A limitation of the present trial is that the comprehensive data on HRQoL was collected only up to 180 days after randomisation. As the intervention was expected to have long-term effects, it is possible that patients in the QI group may increase HRQoL to population norms sooner than the usual care group. When calculating costs, we took the assumption that the costs of the QI intervention were the same across all hospitals. As many hospitals in the UK were involved in the EPOCH trial, it is unlikely that the costs of the intervention are the same, but the results showed that the intervention costs were only a very small proportion of the total costs, so this assumption is unlikely to substantially affect results. Last, the cost-effectiveness of the intervention in the EPOCH trial population was dependent on the regression models developed based on data from the subsample. The prediction bias may apply when generalising results to the EPOCH trial population.

Conclusion

Using data from a large number of patients, enrolled by many hospitals, with an efficient trial design, the national QI programme to implement an enhanced pathway of care for patients undergoing emergency abdominal surgery is unlikely to be cost-effective within the trial period, but has the potential to be cost-effective when the long-term perspective was adopted, particularly for patients at higher risk of mortality and morbidity.

Chapter 7 Discussion

The principal finding of this trial was that there was no survival benefit associated with a national QI 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 among the 10 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 when existing relationships within and beyond the perioperative team were weaker and so QI leads had to spend time developing relationships with stakeholders. The cost-effectiveness analysis did not demonstrate any benefits within the trial period. There may have been some benefit over lifetime for a high-risk subgroup of patients. Together, the results of the cluster trial, ethnographic study and process evaluation suggest that QI 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 on the impact of smaller-scale QI projects to improve outcomes for patients undergoing emergency abdominal surgery. In the UK, the emergency laparotomy pathway quality improvement care (ELPQuiC) group examined the implementation of a care bundle in four NHS hospitals using a cohort study.⁴ Huddart *et al.*⁴ reported a reduction in mortality (risk ratio 0.61) among 726 patients. Although this study design is more prone to bias than a stepped-wedge cluster randomised trial, it seems more probable that the difference in findings relates to the intervention itself and the strong existing relationships between staff leading implementation in these early adopter hospitals. The EPOCH trial QI team were implementing an extensive care pathway with 37 components, whereas the ELPQuiC team were implementing a more discrete care bundle of five interventions. This 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 QI interventions for patients undergoing emergency abdominal surgery. The Peptic Ulcer Perforation (PULP) trial group used a cohort design to study the effect of a 'multidisciplinary perioperative care protocol' in seven hospitals and reported a considerable reduction in 30-day mortality in comparison with historical controls.⁶ 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) among 286 patients undergoing emergency abdominal surgery in seven hospitals.⁵ 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 cohort study with historical controls, finding a more modest reduction in 30-day mortality from 22% among 600 control patients to 16% among 600 intervention patients.⁷ It is possible that a background trend to improved mortality may explain the findings of these previous cohort studies, especially given the growing international focus on poor patient outcomes following emergency abdominal surgery.

Although our analysis accounts for temporal trends during the trial, it is possible that a decreasing mortality before the trial may explain why the mortality rate was lower than that predicted from NHS registry data. Meanwhile, recent studies of QI in other clinical areas have delivered mixed results.^{32,61,95–97} 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. The COM-B model ('capability', 'opportunity', 'motivation' and 'behaviour') offers a useful structure for behaviour change,⁹⁸ but does not address the need for institutional support or protected leadership time. In the EPOCH trial, interventions such as pre-operative documentation of patient risk were easier to implement than postoperative admission to critical care, but may have had less impact on

patient outcomes. It is important to note that the NELA was launched only 3 months before the EPOCH trial commenced. Our ethnographic findings suggest that the task of collecting and entering data into the NELA database was more time-consuming than expected, leaving some QI leads with little time to focus on change. We allowed a 5-week period for the transition between usual care and the launch of the QI 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 among hospitals exposed to the QI programme for longer than 10 weeks, which included hospitals exposed for up to 80 weeks.

The strengths of this trial include wide generalisability (a large number of consecutive patients enrolled by many hospitals), robust trial design and the devolved leadership to local clinical QI teams. The evidence-based EPOCH trial care pathway was developed through a Delphi consensus process to update national professional guidelines. Partnership with the 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 HES 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% CI for our primary effect estimate was narrow, with a lower limit that indicates a maximum potential mortality reduction of 4%. Our findings are unlikely to change with a larger sample size. Owing 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 that 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.

Implications for clinical practice

In this stepped-wedge cluster randomised trial, we did not identify any survival benefit from a national QI programme to implement an enhanced pathway of care for patients undergoing emergency abdominal surgery. This is probably because of variation between hospitals in fidelity of implementation and prioritisation of pathway components, and the time required to achieve effective change. Our findings also point to notable differences in approaches taken by participating sites in implementing the components of the pathway, the influence of differential levels of engagement from key groups, notably hospital managers and surgical colleagues, and the ways in which wider organisational turbulence could detract from, or occasionally complement, efforts at change. These findings suggest that future QI programmes should implement fewer, more discrete changes and ensure that clinical leaders have adequate time to achieve sustained improvements in patient care.

Recommendations for future research

It is essential that scientific reports of QI projects comply with the same high standards of scientific reporting as other forms of clinical research. Many projects are designed as QI projects and then reported as if they had been prospectively designed as a clinical investigation. This leads to a high risk of bias in favour of QI interventions, which may actually be ineffective. The QI community needs to accept that concepts of good science are as relevant to this field as to all others.

Further research specific to the EPOCH trial should include a more detailed process evaluation and ethnographic study to understand aspects in which the trial intervention did and did not work, to enable NHS organisations to consider how to improve patient care for this patient group.

Acknowledgements

The EPOCH trial was an investigator-initiated trial funded by the National Institute for Health Research (12/5005/10). The trial was sponsored by Queen Mary University, London. The EPOCH trial investigators were entirely responsible for trial design, conduct and data analysis.

We wish to thank all members of the EPOCH trial group.

Contributions of authors

Carol J Peden contributed to protocol development and design of the EPOCH trial; designed and delivered the EPOCH QI programme, including the programme theory; led the programme evaluation; provided input to the ethnographic study; and was responsible for the conduct of the trial.

Tim Stephens contributed to protocol development and design of the EPOCH trial; designed and delivered the EPOCH QI programme, including the programme theory; led the programme evaluation; provided input to the ethnographic study; and was responsible for the conduct of the trial.

Graham Martin contributed to protocol development and design of the EPOCH trial; provided input to the programme evaluation; and led the ethnographic study.

Brennan C Kahan contributed to protocol development and design of the EPOCH trial and was responsible for the conduct of the trial.

Ann Thomson contributed to protocol development and design of the EPOCH trial and was responsible for the conduct of the trial.

Kirsty Everingham contributed to protocol development and design of the EPOCH trial and was responsible for the conduct of the trial.

David Kocman provided input to the programme evaluation and led the ethnographic study.

Jose Lourtie supported the trial through the NELA database.

Sharon Drake supported the trial through the NELA database.

Alan Girling contributed to protocol development and design of the EPOCH trial.

Richard Lilford contributed to protocol development and design of the EPOCH trial.

Kate Rivett contributed to protocol development and design of the EPOCH trial.

Duncan Wells contributed to protocol development and design of the EPOCH trial.

Ravi Mahajan contributed to protocol development and design of the EPOCH trial.

Peter Holt contributed to protocol development and design of the EPOCH trial.

Fan Yang led the health economic evaluation.

Simon Walker led the health economic evaluation.

Gerry Richardson contributed to protocol development and design of the EPOCH trial and led the health economic evaluation.

Sally Kerry contributed to protocol development and design of the EPOCH trial and was responsible for conduct of the trial.

Iain Anderson supported the trial through the NELA database.

Dave Murray supported the trial through the NELA database.

David Cromwell supported the trial through the NELA database.

Mandeep Phull supported data acquisition during the EPOCH trial and assisted the running of the EPOCH QI programme. She reviewed the manuscript and contributed to the final draft.

Mike PW Grocott supported the trial through the NELA database.

Julian Bion contributed to protocol development and design of the EPOCH trial.

Rupert M Pearse contributed to protocol development and design of the EPOCH trial; led the programme evaluation; provided input to the ethnographic study; provided input to the health economic evaluation; and was responsible for the conduct of the trial.

All authors read and approved the final manuscript.

Publications

Martin GP, Kocman D, Stephens T, Peden CJ, Pearse RM. Pathways to professionalism? Quality improvement, care pathways, and the interplay of standardisation and clinical autonomy. *Sociol Health Ill* 2017;**39**:1314–29.

Stephens TJ, Peden CJ, Pearse RM, Shaw SE, Abbott TEF, Jones EL, *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* 2018;**13**:148. [Erratum published in *Implement Sci* 2018;**13**:148.]

Murray D, Pearse RM, Quiney N. Towards a coordinated national strategy to improve survival after emergency laparotomy. *Br J Anaesth* 2019;**123**:399–401.

Peden CJ, Stephens T, Martin G, Kahan BC, Thomson A, Rivett K, *et al*. Effectiveness of a national quality improvement programme to improve survival after emergency abdominal surgery (EPOCH): a stepped-wedge cluster-randomised trial. *Lancet* 2019;**393**:2213–21.

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. *Anaesth Crit Care Pain Med* 2019;**38**:321–2.

Stephens TJ, Peden CJ, Haines R, Grocott MPW, Murray D, Cromwell D, *et al*. Enhanced Perioperative Care for High-risk patients (EPOCH) trial group. Hospital-level evaluation of the effect of a national quality improvement programme: time-series analysis of registry data [published online ahead of print 12 September 2019]. *BMJ Qual Saf* 2019.

Data-sharing statement

This research used data subject to a confidentiality agreement with NHS Digital. The agreement in place for these data does not permit further distribution or sharing. Requests for the relevant data sets must be made directly to NHS Digital. Further information can be obtained from the corresponding author.

Patient data

This work uses data provided by patients and collected by the NHS as part of their care and support. Using patient data is vital to improve health and care for everyone. There is huge potential to make better use of information from people's patient records, to understand more about disease, develop new treatments, monitor safety, and plan NHS services. Patient data should be kept safe and secure, to protect everyone's privacy, and it's important that there are safeguards to make sure that it is stored and used responsibly. Everyone should be able to find out about how patient data are used. #datasaveslives You can find out more about the background to this citation here: <https://understandingpatientdata.org.uk/data-citation>.

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* 2017;**119**:249–57. <https://doi.org/10.1093/bja/aex137>
2. National Emergency Laparotomy Audit project team. *First patient report of the National Emergency Laparotomy Audit*. London: Royal College of Anaesthetists; 2015.
3. Saunders DI, Murray D, Pichel AC, Varley S, Peden CJ, UK Emergency Laparotomy Network. Variations in mortality after emergency laparotomy: the first report of the UK Emergency Laparotomy Network. *Br J Anaesth* 2012;**109**:368–75. <https://doi.org/10.1093/bja/aes165>
4. Huddart S, Peden CJ, Swart M, McCormick B, Dickinson M, Mohammed MA, Quiney N, ELPQuIC Collaborator Group. Use of a pathway quality improvement care bundle to reduce mortality after emergency laparotomy. *Br J Surg* 2015;**102**:57–66. <https://doi.org/10.1002/bjs.9658>
5. Vester-Andersen M, Lundstrøm LH, Møller MH, Waldau T, Rosenberg J, Møller AM, Danish Anaesthesia Database. Mortality and postoperative care pathways after emergency gastrointestinal surgery in 2904 patients: a population-based cohort study. *Br J Anaesth* 2014;**112**:860–70. <https://doi.org/10.1093/bja/aet487>
6. Møller MH, Adamsen S, Thomsen RW, Møller AM, Peptic Ulcer Perforation (PULP) trial group. Multicentre trial of a perioperative protocol to reduce mortality in patients with peptic ulcer perforation. *Br J Surg* 2011;**98**:802–10. <https://doi.org/10.1002/bjs.7429>
7. Tengberg LT, Bay-Nielsen M, Bisgaard T, Cihoric M, Lauritsen ML, Foss NB, AHA study group. Multidisciplinary perioperative protocol in patients undergoing acute high-risk abdominal surgery. *Br J Surg* 2017;**104**:463–71. <https://doi.org/10.1002/bjs.10427>
8. Anderson I, Eddlestone J, Lees N. *The Higher Risk General Surgical Patient – Towards Improved Care for a Forgotten Group*. London: Royal College of Surgeons; 2011.
9. Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat AH, Dellinger EP, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009;**360**:491–9. <https://doi.org/10.1056/NEJMsa0810119>
10. Kuper M, Gold SJ, Callow C, Quraishi T, King S, Mulreany A, et al. Intraoperative fluid management guided by oesophageal Doppler monitoring. *BMJ* 2011;**342**:d3016. <https://doi.org/10.1136/bmj.d3016>
11. Anthony T, Murray BW, Sum-Ping JT, Lenkovsky F, Vornik VD, Parker BJ, et al. Evaluating an evidence-based bundle for preventing surgical site infection: a randomized trial. *Arch Surg* 2011;**146**:263–9. <https://doi.org/10.1001/archsurg.2010.249>
12. Moraros J, Lemstra M, Nwankwo C. Lean interventions in healthcare: do they actually work? A systematic literature review. *Int J Qual Health Care* 2016;**28**:150–65. <https://doi.org/10.1093/intqhc/mzv123>
13. Schouten LM, Hulscher ME, van Everdingen JJ, Huijsman R, Grol RP. Evidence for the impact of quality improvement collaboratives: systematic review. *BMJ* 2008;**336**:1491–4. <https://doi.org/10.1136/bmj.39570.749884.BE>
14. Benning A, Dixon-Woods M, Nwulu U, Ghaleb M, Dawson J, Barber N, et al. Multiple component patient safety intervention in English hospitals: controlled evaluation of second phase. *BMJ* 2011;**342**:d199. <https://doi.org/10.1136/bmj.d199>

15. Benning A, Ghaleb M, Suokas A, Dixon-Woods M, Dawson J, Barber N, *et al.* Large scale organisational intervention to improve patient safety in four UK hospitals: mixed method evaluation. *BMJ* 2011;**342**:d195. <https://doi.org/10.1136/bmj.d195>
16. Polle SW, Wind J, Fuhring JW, Hofland J, Gouma DJ, Bemelman WA. Implementation of a fast-track perioperative care program: what are the difficulties? *Dig Surg* 2007;**24**:441–9. <https://doi.org/10.1159/000108327>
17. 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* 2012;**14**:1045–51. <https://doi.org/10.1111/j.1463-1318.2011.02856.x>
18. Batalden PB, Davidoff F. What is ‘quality improvement’ and how can it transform healthcare? *Qual Saf Health Care* 2007;**16**:2–3. <https://doi.org/10.1136/qshc.2006.022046>
19. Dixon-Woods M, Martin GP. Does quality improvement improve quality? *Future Hosp J* 2016;**3**:191–4. <https://doi.org/10.7861/futurehosp.3-3-191>
20. Auerbach AD, Landefeld CS, Shojania KG. The tension between needing to improve care and knowing how to do it. *N Engl J Med* 2007;**357**:608–13. <https://doi.org/10.1056/NEJMs070738>
21. Leatherman S, Sutherland K. *The Quest for Quality in the NHS: Refining the NHS Reforms. Policy Analysis and Chartbook*. London: Nuffield Trust; 2008.
22. Stephens TJ, Peden CJ, Pearse RM, Shaw SE, Abbott TEF, Jones EL, *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* 2018;**13**:148. [Erratum published in *Implement Sci* 2018;**13**:148.] <https://doi.org/10.1186/s132012-018-0823-9>
23. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008;**337**:a1655. <https://doi.org/10.1136/bmj.a1655>
24. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, *et al.* Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ* 2014;**348**:g1687. <https://doi.org/10.1136/bmj.g1687>
25. 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* 2016;**25**:986–92. <https://doi.org/10.1136/bmjqs-2015-004411>
26. Langley A, Denis JL. Beyond evidence: the micropolitics of improvement. *BMJ Qual Saf* 2011;**20**(Suppl. 1):i43–6. <https://doi.org/10.1136/bmjqs.2010.046482>
27. Langley G, Moen R, Nolan K. *The Improvement Guide: A Practical Approach to Enhancing Organizational Performance*. 2nd edn. San Francisco, CA: Jossey-Bass; 2009.
28. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ, GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;**336**:924–6. <https://doi.org/10.1136/bmj.39489.470347.AD>
29. Quiney N, Huddart S, Peden C, Dickinson M. Use of a care bundle to reduce mortality following emergency laparotomy. *Br J Hosp Med* 2015;**76**:358–62. <https://doi.org/10.12968/hmed.2015.76.6.358>
30. 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* 2011;**89**:167–205. <https://doi.org/10.1111/j.1468-0009.2011.00625.x>

31. 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* 2012;**21**:876–84. <https://doi.org/10.1136/bmjqs-2011-000760>
32. Dixon-Woods M, Leslie M, Tarrant C, Bion J. Explaining Matching Michigan: an ethnographic study of a patient safety program. *Implement Sci* 2013;**8**:70. <https://doi.org/10.1186/1748-5908-8-70>
33. 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* 2015;**10**:99. <https://doi.org/10.1186/s13012-015-0289-y>
34. Hysong SJ. Meta-analysis: audit and feedback features impact effectiveness on care quality. *Med Care* 2009;**47**:356–63. <https://doi.org/10.1097/MLR.0b013e3181893f6b>
35. Mills PD, Weeks WB. Characteristics of successful quality improvement teams: lessons from five collaborative projects in the VHA. *Jt Comm J Qual Saf* 2004;**30**:152–62. [https://doi.org/10.1016/S1549-3741\(04\)30017-1](https://doi.org/10.1016/S1549-3741(04)30017-1)
36. 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* 2012;**21**:13–20. <https://doi.org/10.1136/bmjqs-2011-000010>
37. 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* 2016;**31**:197–202. <https://doi.org/10.1177/1062860614568647>
38. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, *et al.* Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev* 2012;**6**:CD000259. <https://doi.org/10.1002/14651858.CD000259.pub3>
39. Gabbay J, le May A, Connell CJK. *Skilled for Improvement?* London: The Health Foundation; 2014.
40. Alexander JA, Hearld LR. The science of quality improvement implementation: developing capacity to make a difference. *Med Care* 2011;**49**:S6–20. <https://doi.org/10.1097/MLR.0b013e3181e1709c>
41. Perla RJ, Provost LP, Murray SK. The run chart: a simple analytical tool for learning from variation in healthcare processes. *BMJ Qual Saf* 2011;**20**:46–51. <https://doi.org/10.1136/bmjqs.2009.037895>
42. Greenhalgh T, Robert G, Macfarlane F, Bate P, Kyriakidou O. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q* 2004;**82**:581–629. <https://doi.org/10.1111/j.0887-378X.2004.00325.x>
43. 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* 2014;**23**:290–8. <https://doi.org/10.1136/bmjqs-2013-001862>
44. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ* 1996;**312**:71–2. <https://doi.org/10.1136/bmj.312.7023.71>
45. Great Britain. *National Health Services Act 2006*. London: The Stationery Office; 2006.
46. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* 2007;**28**:182–91. <https://doi.org/10.1016/j.cct.2006.05.007>
47. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ* 2011;**342**:d40. <https://doi.org/10.1136/bmj.d40>
48. Kahan BC, Jairath V, Doré CJ, Morris TP. The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies. *Trials* 2014;**15**:139. <https://doi.org/10.1186/1745-6215-15-139>

49. 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* 2016;**16**:42. <https://doi.org/10.1186/s12874-016-0141-3>
50. White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. *Stat Med* 2005;**24**:993–1007. <https://doi.org/10.1002/sim.1981>
51. 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* 2017;**36**:318–33. <https://doi.org/10.1002/sim.7137>
52. Hemming K, Taljaard M, Forbes A. Analysis of cluster randomised stepped wedge trials with repeated cross-sectional samples. *Trials* 2017;**18**:101. <https://doi.org/10.1186/s13063-017-1833-7>
53. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, *et al.* Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 2015;**350**:h1258. <https://doi.org/10.1136/bmj.h1258>
54. 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* 2016;**42**:196–206. [https://doi.org/10.1016/S1553-7250\(16\)42025-8](https://doi.org/10.1016/S1553-7250(16)42025-8)
55. 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* 2013;**14**:15. <https://doi.org/10.1186/1745-6215-14-15>
56. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs* 2008;**62**:107–15. <https://doi.org/10.1111/j.1365-2648.2007.04569.x>
57. Blencowe NS, Brown JM, Cook JA, Metcalfe C, Morton DG, Nicholl J, *et al.* Interventions in randomised controlled trials in surgery: issues to consider during trial design. *Trials* 2015;**16**:392. <https://doi.org/10.1186/s13063-015-0918-4>
58. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci* 2007;**2**:40. <https://doi.org/10.1186/1748-5908-2-40>
59. Eveleigh MO, Howes TE, Peden CJ, Cook TM. Estimated costs before, during and after the introduction of the emergency laparotomy pathway quality improvement care (ELPQulC) bundle. *Anaesthesia* 2016;**71**:1291–5. <https://doi.org/10.1111/anae.13623>
60. Spurgeon P, Mazelan PM, Barwell F. Medical engagement: a crucial underpinning to organizational performance. *Health Serv Manage Res* 2011;**24**:114–20. <https://doi.org/10.1258/hsmr.2011.011006>
61. 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* 2017;**7**:e014333. <https://doi.org/10.1136/bmjopen-2016-014333>
62. Burke RE, Shojania KG. Rigorous evaluations of evolving interventions: can we have our cake and eat it too? *BMJ Qual Saf* 2018;**27**:254–7. <https://doi.org/10.1136/bmjqs-2017-007554>
63. Dixon-Woods M. Why is patient safety so hard? A selective review of ethnographic studies. *J Health Serv Res Policy* 2010;**15**(Suppl. 1):11–16. <https://doi.org/10.1258/jhsrp.2009.009041>
64. Wachter RM. Patient safety at ten: unmistakable progress, troubling gaps. *Health Aff* 2010;**29**:165–73. <https://doi.org/10.1377/hlthaff.2009.0785>
65. Nadeem E, Olin SS, Hill LC, Hoagwood KE, Horwitz SM. Understanding the components of quality improvement collaboratives: a systematic literature review. *Milbank Q* 2013;**91**:354–94. <https://doi.org/10.1111/milq.12016>

66. Weiss C. *Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families*. New York, NY: Aspen Institute; 1995.
67. Davidoff F, Dixon-Woods M, Leviton L, Michie S. Demystifying theory and its use in improvement. *BMJ Qual Saf* 2015;**24**:228–38. <https://doi.org/10.1136/bmjqs-2014-003627>
68. Dr Foster Limited. *Dr Foster*. URL: drfosterintelligence.co.uk (accessed 11 July 2019).
69. Charmaz K. *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. London: Sage; 2007.
70. Waring J, Rowley E, Dingwall R, Palmer C, Murcott T. Narrative review of the UK Patient Safety Research Portfolio. *J Health Serv Res Policy* 2010;**15**(Suppl. 1):26–32. <https://doi.org/10.1258/jhsrp.2009.009042>
71. Dixon-Woods M, Martin G, Tarrant C, Bion J, Goeschel C, Pronovost P. *Safer Clinical Systems: Evaluation Findings*. London: The Health Foundation; 2014.
72. Weiser TG, Regenbogen SE, Thompson KD, Haynes AB, Lipsitz SR, Berry WR, Gawande AA. An estimation of the global volume of surgery: a modelling strategy based on available data. *Lancet* 2008;**372**:139–44. [https://doi.org/10.1016/S0140-6736\(08\)60878-8](https://doi.org/10.1016/S0140-6736(08)60878-8)
73. Pearse RM, Moreno RP, Bauer P, Pelosi P, Metnitz P, Spies C, *et al*. Mortality after surgery in Europe: a 7 day cohort study. *Lancet* 2012;**380**:1059–65. [https://doi.org/10.1016/S0140-6736\(12\)61148-9](https://doi.org/10.1016/S0140-6736(12)61148-9)
74. Findley GP. *Knowing the Risk: A Review of the Peri-Operative Care of Surgical Patients: Summary: National Confidential Enquiry into Perioperative Deaths*. London: National Confidential Enquiry into Patient Outcome and Death; 2011.
75. Pearse RM, Harrison DA, James P, Watson D, Hinds C, Rhodes A, *et al*. Identification and characterisation of the high-risk surgical population in the United Kingdom. *Crit Care* 2006;**10**:R81. <https://doi.org/10.1186/cc4928>
76. de Vries EN, Prins HA, Crolla RM, den Outer AJ, van Andel G, van Helden SH, *et al*. Effect of a comprehensive surgical safety system on patient outcomes. *N Engl J Med* 2010;**363**:1928–37. <https://doi.org/10.1056/NEJMsa0911535>
77. Department of Health and Social Care. *NHS Reference Costs 2015 to 2016*. 2016. URL: www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016 (accessed 18 July 2019).
78. Curtis L. *Unit Costs of Health and Social Care 2016*. Canterbury: PSSRU, University of Kent; 2016.
79. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;**35**:1095–108. <https://doi.org/10.1097/00005650-199711000-00002>
80. Matthews JN, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. *BMJ* 1990;**300**:230–5. <https://doi.org/10.1136/bmj.300.6719.230>
81. Faria R, Gomes M, Epstein D, White IR. A guide to handling missing data in cost-effectiveness analysis conducted within randomised controlled trials. *Pharmacoeconomics* 2014;**32**:1157–70. <https://doi.org/10.1007/s40273-014-0193-3>
82. Gomes M, Ng ES, Grieve R, Nixon R, Carpenter J, Thompson SG. Developing appropriate methods for cost-effectiveness analysis of cluster randomized trials. *Med Decis Making* 2012;**32**:350–61. <https://doi.org/10.1177/0272989X11418372>
83. National Institute for Health and Care Excellence (NICE). *Guide to the Methods of Technology Appraisal 2013*. London: NICE; 2013.

84. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, *et al.* Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technol Assess* 2015;**19**(14). <https://doi.org/10.3310/hta19140>
85. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. 4th edn. Oxford :Oxford University Press; 2015.
86. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Econ* 2001;**10**:779–87. <https://doi.org/10.1002/hec.635>
87. van Hout BA, Al MJ, Gordon GS, Rutten FF. Costs, effects and C/E-ratios alongside a clinical trial. *Health Econ* 1994;**3**:309–19. <https://doi.org/10.1002/hec.4730030505>
88. Fenwick E, O'Brien BJ, Briggs A. Cost-effectiveness acceptability curves – facts, fallacies and frequently asked questions. *Health Econ* 2004;**13**:405–15. <https://doi.org/10.1002/hec.903>
89. Janssen B, Szende A. Population Norms for the EQ-5D. In Szende A, Janssen B, Cabases J, editors. *Self-Reported Population Health: An International Perspective Based on EQ-5D*. Dordrecht: Springer Netherlands; 2014. pp. 19–30. https://doi.org/10.1007/978-94-007-7596-1_3
90. Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the Economic Evaluation of Health Care Programmes*. 3rd edn. Oxford: Oxford University Press; 2005.
91. Office for National Statistics. *National Life Tables, United Kingdom: 2012–14*. 2015. URL: www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/bulletins/nationallifetablesunitedkingdom/2015-09-23 (accessed 20 June 2018).
92. Asaria M. *Health Care Costs in the English NHS: Reference Tables for Average Annual NHS Spend by Age, Sex and Deprivation Group*. York: Centre for Health Economics University of York; 2017.
93. Danyliv A, Gillespie P, O'Neill C, Noctor E, O'Dea A, Tierney M, *et al.* Health related quality of life two to five years after gestational diabetes mellitus: cross-sectional comparative study in the ATLANTIC DIP cohort. *BMC Pregnancy Childbirth* 2015;**15**:274. <https://doi.org/10.1186/s12884-015-0705-y>
94. Henriksson M, Epstein DM, Palmer SJ, Sculpher MJ, Clayton TC, Pocock SJ, *et al.* The cost-effectiveness of an early interventional strategy in non-ST-elevation acute coronary syndrome based on the RITA 3 trial. *Heart* 2008;**94**:717–23. <https://doi.org/10.1136/hrt.2007.127340>
95. Bion J, Richardson A, Hibbert P, Beer J, Abrusci T, McCutcheon M, *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* 2013;**22**:110–23. <https://doi.org/10.1136/bmjqs-2012-001325>
96. Presseau J, Mackintosh J, Hawthorne G, Francis JJ, Johnston M, Grimshaw JM, *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* 2018;**13**:65. <https://doi.org/10.1186/s13012-018-0754-5>
97. Williams L, Daggett V, Slaven JE, Yu Z, Sager D, Myers J, *et al.* A cluster-randomised quality improvement study to improve two inpatient stroke quality indicators. *BMJ Qual Saf* 2016;**25**:257–64. <https://doi.org/10.1136/bmjqs-2015-004188>
98. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011;**6**:42. <https://doi.org/10.1186/1748-5908-6-42>

Appendix 1 Members of the EPOCH trial group

Writing Committee

Rupert Pearse (chief investigator), Carol Peden (QI lead), Tim Stephens, Graham Martin, Julian Bion, Ann Thomson, Brennan Kahan, Sally Kerry, Kate Rivett (patient representative), Duncan Wells (patient representative) and Gerry Richardson.

Trial Steering Committee

Stephen Brett (chairperson), Gareth Ackland (independent member), Julian Bion, Mike Grocott, Peter Holt, Sally Kerry, Graham Martin, Rupert Pearse, Carol Peden, Gerry Richardson, Kate Rivett (patient representative), Glenn Robert (independent member), Tim Stephens, Obioha Ukoumunne (independent member), Justin Waring (independent member) and Duncan Wells (patient representative).

Trial Management Group

Kirsty Everingham, Brennan Kahan, Sally Kerry, Rupert Pearse, Tim Stephens, Ann Thomson and Mandeep Phull.

National Emergency Laparotomy Audit Team

Jose Lourtie, Sharon Drake, David Cromwell and David Murray.

Methodology Team

Miqdad Asaria, David Cromwell, Rachel Evley, Omar Faiz, Gordon Forbes, Alan Girling, Brennan Kahan, Sally Kerry, David Kocman, Richard Lilford, Ravi Mahajan, Carolyn Tarrant, Gerry Richardson, Fan Yang and Janet Willars.

EPOCH trial cluster leads

Ajit Abraham (North East London), Pieter Bothma (East Anglia), Daniel Conway (Manchester, Merseyside and Yorkshire), Mark Edwards (Wessex), Omar Faiz (North West London), Michael Gillies (Scotland), Paul Hayden (Kent, Surrey and Sussex), Peter Holt (South London), Gary Minto (West Country), Tom Owen (North Lancashire and Cumbria), David Saunders (north-east of England), Clare Stapleton (Thames Valley), Robert Sutcliffe (West Midlands), Tamas Szakmany (Wales) and Nick Watson (East Midlands).

EPOCH trial Advisory Group

Iain Anderson (chairperson), Anthony Gordon, Simon Howell, Justin Kirk-Bayley, Martin Leuwer, Dileep Lobo, David Murray, Monty Mythen, Kate Rivett (patient representative), Neil Smith, Sean Tighe and Carl Waldmann.

Local EPOCH trial investigators

East Anglia

Queen Elizabeth Hospital, King's Lynn

Mark Blunt (principal investigator), Surjait Singh, Alistair Steel, Kate Wong and Leilani Cabrerros.

James Paget University Hospital, Norfolk

Pieter Bothma (principal investigator), Vivek Chitre, Ayodele Obideyi, Dhiraj Ali, Karl Blenk, Dan Broad, Andreas Brodbeck, Rajesh Dumpala, Arnth Engel, Ranjit Ganepola, Sudha Garg, Mike Gay, Michael Karlikowsk, Edward Lams, Dean Millican, Inga Misane, Ajaya Mull, Veena Naik, Nathan Pushpa, Chris Nutt, Ayodele Obideyi, Saravanna Sagadai, Hazel Stuart, Paul Noble, Niko Van De Velde, Liam Hudson, Raoul Benlloch, Satish Singh, Karan Verma, Damian Laba, Jack Carmichael, Peter Richardson, Graham Wilson, Ricky Lewis, Karthik Surendran, Essam El-Damatty, Sarada Gurung, Ilona Raulusaite, Nabua Gerstina, Chloe Rochester, Rai Kuldip, Andrew Lindner, Therese Murray, Chitre Vivek, Roshan Lal, Sarah Downey, Vamsi Velchuru, Kamal Aryal, Raman Guruswamy and Kirosh Shankar.

Colchester Hospital, Colchester

Helen Porter (principal investigator), Matthew Tutton and Helen Agostini.

Norfolk and Norwich University Hospital, Norwich

Simon Fletcher (principal investigator), Richard Wharton, Steve Hutchinson and Bala Maiya.

Ipswich Hospital, Ipswich

Richard Howard-Griffin (principal investigator), Michael Crabtree, Vlad Kushakovsky, Abdel Omer, Senthil Nadarajavan and Stephanie Bell.

Addenbrooke's Hospital, Cambridge

Vishal Patil (principal investigator), Asif Jah and Razeen Mahroof.

East Midlands

King's Mill Hospital, Nottingham

Nicholas Watson (principal investigator), John Tansley and Gareth Moncaster.

Leicester Royal Infirmary, Leicester

Neil Flint (principal investigator), Andrew Miller, Marcus Wood, Andreou Prematie, Sally Roth, Sarah Bowery and Dawn Hales.

Royal Derby Hospital, Derby

Tanuja Shah (principal investigator), Gill Tierney, Craig Morris and Syed Iftikhar.

Lincoln County Hospital, Lincoln

Amit Shukla (principal investigator), Grainne O'Dwyer, Adam Wolverson, Ferdinand Adams and Laura Perrin-Brown.

Chesterfield Royal Hospital, Chesterfield

Tim White (principal investigator) and Sarah Beavis.

Queen's Medical Centre, Nottingham

Victoria Banks (principal investigator), John Abercrombie, Jonathon Mole, Avninder Chana, Ayan Banerjea, David Humes, Rajpal Dhingsa, John Wells and Stephanie Brown.

Kent, Surrey and Sussex

Queen Elizabeth The Queen Mother Hospital, Kent

Kenneth Adegoke (principal investigator), Samer Doughan, Barclay Tofte, Ana Alegria, Nat Natarajan and Mansoor Akhtar.

William Harvey Hospital, Kent

John Mackinnon (principal investigator), Biju Aravind, Esther Cook, Mark Snazelle and Matt Gardner.

Kent and Canterbury Hospital, Kent

John Mackinnon (principal investigator).

Tunbridge Wells, Kent

Lee Baldwin (principal investigator), Simon Bailey and Greg Lawton.

Medway Maritime Hospital, Kent

Nandita Divekar (principal investigator) and Neil Kukreja.

Darent Valley Hospital, Kent

Mansoor Sange (principal investigator), Mark Watson, Mallikarjunappa Satisha, Michael Protopapas, Zakaulla Belagodu, Shameem Sarfi, Pasupathy Raju and Brenda Stacey.

East Surrey Hospital, Surrey

Tim Campbell-Smith (principal investigator), Simon Parrington and Somi Desikan.

Manchester, Merseyside and Yorkshire

Bradford Royal Infirmary, Bradford

Andrew Brennan (principal investigator), John Griffith, Steve Fletcher, Catherine Farrow, Stewart Prestwich, Laura Graham, Martin Northey, Jay Gokhale and Frances Mosley.

Wythenshawe Hospital, Manchester

Peter Alexander (principal investigator), Abhiram Sharma, Will Brady, John Hopper and Oliver Hill.

Pinderfields Hospital, Yorkshire

Sandeep Varma (principal investigator), Christopher Macklin, Alastair Rose, Harjeet Narula, Sarah Buckley and Karen Simeson.

Whiston Hospital, Merseyside

Kevin Sim (principal investigator), Michael Chadwick, Preeti Kuduvalli, Susan Dowling and Amanda McCairn.

Countess of Chester, Chester

Lawrence Wilson (principal investigator), Dale Vimalchandran, Anita Jhamatt, Nicole Robin and David Monk.

Royal Liverpool University Hospital, Liverpool

Martin Leuwer (principal investigator), David Bottomley, Oliver Zuzan and Ingeborg Welters.

North East London

Royal London, London

Davina Ross-Anderson (principal investigator), Charles Knowles, Nick Bunker, Kirsty Everingham, Ying Hu, Marta Januszewska, Phoebe Bodger, Edyta Niebrzegowska, Carmen Correia, Richard Haslop and Tom Abbott.

Homerton University Hospital, London

Tabitha Tanqueray (principal investigator), Sanjay Wijeykoon, Susan Jain, Jens Full, Tamzin Cuming, Flora Bailey, Stelios Chatzimichail, Pedro Cunha, Almas Rehman, Manab Mohanty and Nicola Radford.

Newham University Hospital, London

Otto Mohr (principal investigator), Hitesh Patel and Dolores Mateo.

Whipps Cross University Hospital, London

Ashok Raj (principal investigator), Michael Machesney, Nazar Abdul, Kim Jemmet, Marta Campbell, David Inglis, Thomas Parker, Thomas Medici and Peter Chan.

Queen's Hospital, London

Nathan Borgeaud (principal investigator), Dipankar Mukherjee, Oluremi Odejinmi, Tomas Jovaisa and Elizabeth Harwood.

University College London Hospital, London

Ramani Moonesinghe (principal investigator), Jonathan Mccullough, Jigna Modha and Sanjiv Patel.

North-east of England**Darlington Memorial Hospital, Darlington**

James Limb (principal investigator), Sheshagiri Bengeri and Amir Rafi.

North Tyneside General Hospital, North Shields

Elizabeth Hall (principal investigator), James Brown, Bruce Gibson, Una McNelis, Mike Bradburn, Maria Lawson and Sara Pick.

Queen Elizabeth Hospital, Gateshead

Matthew Gaughan (principal investigator), David Browell, Vanessa Linnett, Jenny Ritzema and Paul O'Loughlin.

Sunderland Royal Hospital, Sunderland

Sean Cope (principal investigator), John Corson, Alistair Roy and Julie Furneal.

University Hospital Of North Durham, Durham

Anitha Holtham (principal investigator), Sophie Noblett and Chris Dawson.

Wansbeck General Hospital, Ashington

Elizabeth Hall (principal investigator), Mike Bradburn and Fiona McMenemie.

Royal Victoria Infirmary, Newcastle

David Saunders (principal investigator), Stefan Pulsa, Ian Clement, Verity Calder, Katherine Allen, Catherine Rimmer, Helen Reed and Christine Boyd.

James Cook Hospital, Middlesbrough

Diane Monkhouse (principal investigator), Peter Davies, Jost Mullenheim, Emanuel Cirstea, Martyn Cain and Kirsty Baillie.

North Lancashire and Cumbria**Royal Preston Hospital, Preston**

Tom Owen (principal investigator), Arnab Bhowmick, Keiarash Jovestani, Sean McMullan, Emma Durant, Alexandra Williams and Donna Doyle.

Blackpool Victoria Hospital, Blackpool

Jason Cupitt (principal investigator), Jonathon Barker, Nick Harper and Emma Brennan.

Royal Blackburn Hospital, Blackburn

Daren Subar (principal investigator), Robert Shawcross and Dominic Sebastian.

Furness General Hospital, Cumbria

Panna Patel (principal investigator), Gillian O'Connell, Jyrki Karvonen, Alison Hool, Karen Burns, Carol Mcarthur, Maitra Ishaan, Tezas Stergios, Singh Gursevak, Makvana Sonia, Heather Pratt and Kaighan Lynne.

Royal Albert Edward Infirmary, Wigan

Sean McAfee (principal investigator), Chris Lewis and Wael Khalaf.

Royal Lancaster Hospital, Lancaster

Chris Coldwell (principal investigator), Christine Bronder, Mark Wilkinson and Emma Davis.

North West London**St Mary's Hospital, London**

Glenn Arnold (principal investigator), Paul Ziprin, Rachel Bartlett and Martin Stotz.

Royal Free Hospital, London

Rovan D'souza (principal investigator), Phillippa Pemberton, Banwari Agarwal, Anita Sugavanam, Melanie Tan, Massimo Varcada and Craig Lyness.

Hillingdon Hospital, London

Andrew Thorniley (principal investigator), Ash Prabhudesai, Ruth Griffin, Shubha Vashisht, James Harris, Julie Wakeford, Sergei Vaganov, Yasser Mohsen, Alistair Myers, Qamar Iqbal, Simon Harris and Sami Ijaz.

Charing Cross Hospital, London

James Burrow (principal investigator), Paul Ziprin, Francesca Rubulotta, Nabil El-Masry and Nicola Stranix.

Northwick Park Hospital/St Mark's Hospital, London

Tamsin Rope (principal investigator), Lampros Liasis, Tariq Husain, Josef Watfah, Megan Griffiths, Janindra Warusavitarne, Charles Cartwright, Linden Baxter and Rakhee Visavadia.

Scotland**Western Infirmary, Glasgow**

Malcolm Sim (principal investigator), Chris Wilson and Paul Harrison.

Dumfries and Galloway Royal Infirmary, Dumfries

Dewi Williams (principal investigator), Maria Bews-Hair, Wayne Wrathall, Catherine Jardine, Paul McLaren, Fanus Dreyer and Paddy Collins.

Royal Alexandra Hospital, Paisley

Jennifer Edwards (principal investigator), Susan Moug, Kevin Rooney, Erin Mcilveen, Steven Henderson, Linda Graham, Gail Stark, Lynn Taylor, Mark Munro, Lynn Stewart, Natalie Dickinson, Laura Rooney, Lindsay Bailey and Diane Murray.

University Hospital Crosshouse, Kilmarnock

Tim Geary (principal investigator), Simon Gibson, Colin Pow, Kerwei Tan and Richard Stevenson.

Royal Infirmary of Edinburgh, Edinburgh

Ewen Harrison (principal investigator), Peter Lamb, Michael Gillies, Kate Carey, Laura Fitton, Fabian Cook and Magen Schwarz.

Wishaw General Hospital, Wishaw

Alan Morrison (principal investigator), Gavin Bryce, Khaled Razouk and Kathryn Cain.

South London**King's College London, London**

Gudrun Kunst (principal investigator), Savvas Papagrigoriadis and Phil Hopkins.

Kingston Hospital, London

Adrian Fawcett (principal investigator), Britta O'Carroll-Kuehn and Amira Girgis.

St Helier Hospital, London

Stas Janokowski (principal investigator) and Sami Farhat.

Croydon University Hospital, London

Stella Vig (principal investigator), Nada Hadi and Anthony Parsons.

St George's Hospital, London

Maurizio Cecconi (principal investigator), David Melville and Richard Hartopp.

Frimley Park Hospital, London

Justin Woods (principal investigator), Isabella Karat and David Gerrard.

Wales**Nevill Hall Hospital, Abergavenny**

Edward Curtis (principal investigator), Krishnamurthy Somasekar and Tom Morgan-Jones.

Glangwili General Hospital, Camarthen

Michael Martin (principal investigator), Mark Henwood, Gordon Milne and Ajit Sivasankaranand.

Prince Charles Hospital, Merthyr Tydfil

Alexandra Scott (principal investigator), Xavier Escofet, Piroska Toth-Tarsoly, Majed Al Shama, Valerie Hilton and Huw Davis.

Royal Glamorgan Hospital, Llantrisant

Gail Williams (principal investigator), Tim Harvard, Peter Fitzgerald, Dom Hurford and Tamas Szakmany.

Royal Gwent Hospital, Newport

Babu Muthuswamy (principal investigator), Gethin Williams, Jack Parry Jones and Nick Mason.

Glan Clwyd Hospital, Bodelwyddan

Richard Pugh (principal investigator), Ramesh Rajagopal, Shrisha Shenoy, Magdy Khater and Richard Morgan.

Thames Valley**Milton Keynes General Hospital, Milton Keynes**

Nikolaos Makris (principal investigator), Anil Hermandes and Andrew White.

Northampton General Hospital, Northampton

Guy Finch (principal investigator), Matt Outram, Jonny Wilkinson and Jennifer Spimpolo.

Luton and Dunstable University Hospital, Luton

Debbie Shaw (principal investigator), Marion Obichere, Giovanni Brescia, Flavia Menezes and Helena Stafford.

Great Western Hospital, Swindon

Malcolm Watters (principal investigator), Chris Thorn, Julian Stone, Sam Andrews, Nicola Lythell and Helen Langton.

Wexham Park Hospital, Slough

Clare Stapleton (principal investigator), Stephen Baxter and Roy Fernandes.

Stoke Mandeville Hospital, Aylesbury

Rame Sunthareswaran (principal investigator), Alastair Ankers and Kumar Panikkar.

Wessex**Salisbury District Hospital, Salisbury**

Simon Sleight (principal investigator), Belinda Cornforth and Louise Bell.

Royal Hampshire County Hospital, Winchester

Phil Dodd (principal investigator), Fenella Welsh and Geoff Watson.

Basingstoke & North Hampshire Hospital, Hampshire

Phil Dodd (principal investigator), Fenella Welsh and Geoff Watson.

Poole Hospital, Poole

Frankie Dorman (principal investigator), Guy Nash, James Bromilow and Fran Haigh.

Queen Alexandra Hospital, Portsmouth

David Pogson (principal investigator), Stuart Mercer and Vanessa Tucker.

Southampton University Hospital, Southampton

Carolyn Way (principal investigator), James Kirby-Bott, Jenny McLachan and Rob Chambers.

West Country**Bristol Royal Infirmary, Bristol**

Rachael Craven (principal investigator), Jane Blazeby, Dan Freshwater-Turner, Lorna Burrows and Helen Howes.

Derriford Hospital, Plymouth

Iain Christie (principal investigator), Mark Coleman, Gary Minto, Sam Waddy, Grant Sanders, Abigail Patrick, Catherine Pitman, Susan Tyson and Hannah Smith.

North Devon District Hospital, Barnstaple

Guy Rousseau (principal investigator), Mark Cartmell, Jan Hanousek, Nigel Hollister and Lynsey Kightly.

Dorset County Hospital, Dorchester

Mark Pulletz (principal investigator), Anjay Talwar, Susie Baker and Ruth Thomas.

Musgrove Park Hospital, Taunton

Richard Gibbs (principal investigator), Hamish Noble, Joseph Silsby, Helen Black, Thomas Evans, Robert DeBrunner, Nicola Cook, Stacy Hodges, Amanda Stevens and Rowena Felipe.

Royal Cornwall Hospital, Cornwall

Jonathan Paddle (principal investigator), Denzil May and Alison Pickford.

West Midlands**Queen Elizabeth Hospital, Birmingham**

Sid Riddington (principal investigator), Olga Tucker, Simon Smart, Jeremy Marwick, Nigel Suggett, Ewen Griffiths and David Riddington.

Sandwell Hospital, West Bromwich

Kathryn Gill (principal investigator), Neil Cruickshank, Jay Susarla, Emma Leno and Julie Colley.

Worcestershire Royal Hospital, Worcester

Andrew Burtenshaw (principal investigator), Stephen Lake, Jamie Greenwood, Sian Bhardwaj, Jessica Thrush and Julie Wollaston.

Russells Hall Hospital, Dudley

Julian Sonksen (principal investigator), Rajan Patel, Adrian Jennings, David Stanley, Jenny Wright, Chris Horner, Faisal Baig, Katie Cooke and Jagdeep Singh.

New Cross Hospital, Wolverhampton

Andrew Claxton (principal investigator), Nazzia Mirza, Simon Hester and Georgia Knight.

University Hospital Coventry, Coventry

Peeyush Kumar (principal investigator), Taj Saran (principal investigator), Gabriele Marangoni, Roger Townsend, Andy Thacker, Anne Scase, Meghna Sharma and Beth Hale.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

*This report presents independent research funded by the National Institute for Health Research (NIHR).
The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the
Department of Health and Social Care*

Published by the NIHR Journals Library