Effectiveness of a national quality improvement programme to improve survival after emergency abdominal surgery: A stepped-wedge cluster randomised trial

Carol J. Peden, Tim Stephens, Graham Martin, Brennan C. Kahan,

Ann Thomson, Kate Rivett, Duncan Wells, Gerry Richardson, Sally Kerry,

Julian Bion, and Rupert M. Pearse, on behalf of the Enhanced Peri-Operative

Care for High-risk patients (EPOCH) trial group*

*members of group listed in supplementary file

Prof Carol J. Peden MD(Res), Keck School of Medicine, University of Southern California, CA 90033, USA
Mr Tim Stephens MSc, William Harvey Research Institute, Queen Mary University of London, UK EC1M 6BQ
Prof Graham Martin PhD, Health Sciences, University of Leicester, UK LE1 7RH
Mr Brennan C. Kahan MSc, Pragmatic Clinical Trials Unit, Queen Mary University of London, UK E1 2AB
Ms Ann Thomson MSc, Pragmatic Clinical Trials Unit, Queen Mary University of London, UK E1 2AB
Mrs Kate Rivett (patient representative), London, UK
Mr Duncan Wells (patient representative), Chalfont St. Peter, Buckinghamshire, UK
Prof Gerry Richardson PhD, Centre for Health Economics, University of York, UK YO10 5DD
Mrs Sally Kerry MSc, Pragmatic Clinical Trials Unit, Queen Mary University of London, UK E1 2AB
Prof Julian Bion MD(Res), Institute of Clinical Sciences, University of Birmingham, UK B15 2TH
Prof Rupert M. Pearse MD(Res), William Harvey Research Institute, Queen Mary University of London, UK EC1M 6BQ

Correspondence to:
Rupert M. Pearse
Adult Critical Care Unit
Royal London Hospital
London
E1 1BB
United Kingdom
e-mail: r.pearse@qmul.ac.uk
Tel: +44 20 3594 0351

Main text: 4569 words **Summary:** 266 words

Study registration: ISRCTN80682973 and Lancet protocol 13PRT/7655

Keywords: Randomised trial; Quality improvement; Postoperative care/methods; Postoperative care/statistics & numerical data; Surgical Procedures, Operative/mortality;

Summary

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.

Methods
Stepped-wedge cluster randomised trial of patients aged ≥40 years undergoing emergency open major abdominal surgery. 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. The trial included an ethnographic study in six hospitals. The primary outcome measure was mortality within 90 days of surgery. 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.

Findings
Of 15,873 eligible patients from 93 NHS hospitals, primary outcome data were analysed for 8482 patients in the usual care group and 7374 in the QI group. The primary outcome occurred in 1393 patients in the usual care group (16%) compared with 1210 patients in the QI group (16%) (HR QI vs usual care: 1.11 [0.96-1.28]). There were only modest overall improvements in processes of patient care 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.

Interpretation
There was no survival benefit from a QI programme to implement a care pathway for patients undergoing emergency abdominal surgery. The success of the QI intervention may have been limited by the time and resources needed to improve patient care.

Funding
National Institute for Health Research, Health Services and Delivery Research.

Research in context

Evidence before this study

Emergency abdominal surgery is associated with poor post-operative outcomes. Around 30,000 patients undergo this type of surgery each year in the UK National Health Service (NHS), with 30-day mortality rates in excess of 10% and wide variation in standards of care between hospitals. We searched for peer reviewed publications describing the effects of quality improvement programmes on survival for adult patients using the terms 'emergency abdominal surgery' and 'emergency laparotomy'. Several groups have studied the effect of quality improvement 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 weak study designs associated with a high risk of bias. The feasibility and benefit of a national quality improvement programme to implement a more extensive acute care pathway for this patient group remain uncertain.

Added value of this study

We conducted a large national quality improvement programme to implement a care pathway for patients undergoing emergency abdominal surgery. In a stepped-wedge cluster randomised trial of 15,873 patients aged ≥40 years, in 93 NHS hospitals organised into fifteen geographical clusters, we did not identify any survival benefit at either 90 or 180 days after surgery. There was good engagement with the quality improvement programme but staff had limited time and resources to implement change. Consequently, there were only modest overall changes in the processes of patient care from before to after quality improvement implementation. There were wide variations in intervention fidelity between hospitals, with differences in the processes teams tried to change, the rate of change and eventual success.

Implications of all the available evidence

Despite the success of some smaller projects, there was no survival benefit from a national quality improvement programme to implement a care pathway for patients undergoing emergency abdominal surgery. To succeed, large national quality improvement programmes need to allow for differences between hospitals and ensure teams have both the time and resources needed to improve patient care.

Introduction

More than 1.53 million adults undergo in-patient surgery in the UK National Health Service (NHS) each year with a 30-day mortality of 1.5%.[1] However, patients undergoing emergency 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 post-operative admission to critical care. These variations have been associated with differences in mortality rates.[2,3]

In small studies, quality improvement 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, the Royal College of Surgeons of England 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 where discrete quality improvement interventions have been associated with improved patient outcomes,[9,10] but others yielded disappointing results.[11,12] This is especially true for complex interventions requiring co-ordinated change across a healthcare system.[13,14] The benefits of quality improvement initiatives are self-evident to some,[15] but others question the value of these projects, citing high costs, failure to engage clinicians and a lack of scientific rigour.[16,17] Despite this, the direction in healthcare policy is towards ever more widespread use of quality improvement to drive large scale change.[18]

The launch of the National Emergency Laparotomy Audit in December 2013,[2] provided a unique opportunity to study a quality improvement programme to implement a complex care pathway at a national level. We conducted a stepped-wedge cluster randomised trial, with an embedded ethnographic evaluation, to evaluate the hypothesis that implementing this pathway would improve survival following emergency abdominal surgery in NHS hospitals.

Methods

Study design and participants

EPOCH was a multi-centre, stepped-wedge cluster randomised trial of a quality improvement (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 (www.epochtrial.org/protocol). The trial was prospectively registered at isrctn.com on 27th February 2014 but a registration number was not issued until 7th March 2014 (ISRCTN80682973).

NHS hospitals delivering an emergency general surgical service were eligible for inclusion provided they undertook a significant volume of emergency abdominal surgery cases and contributed data to the National Emergency Laparotomy Audit (NELA). Hospitals were required to nominate specialty leads from surgery, anaesthesia and critical care, and to secure support from their NHS Trust Board or equivalent. Hospitals which were already implementing a care pathway to improve treatment for this patient group were excluded. Patients were eligible for inclusion in the data analysis if they were 40 years or older, and undergoing emergency open abdominal surgery in a participating hospital during the 85-week trial period from 3rd March 2014 to 19th October 2015. Patients were excluded from the analysis if they were undergoing a simple 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 (www.nela.org.uk), and then linked using unique patient identifiers to Hospital Episode Statistics and Office for National Statistics in England and Wales, and the Information Services Division of NHS Scotland, to provide data describing mortality and hospital re-admissions. The trial was approved by the East Midlands (Nottingham 1) Research Ethics Committee (Ref: 13/EM/0415). Data were analysed without individual patient consent in accordance with section 251 of the National Health Services Act 2006.

Randomisation and masking

We planned to include fifteen geographical clusters of five to seven hospitals. The QI intervention lasted 80 weeks with one geographical cluster commencing the intervention each five-week step from the 2nd to the 16th time period. Clusters were randomly assigned to one of 15 start dates for the QI intervention by an independent statistician using a computer-generated random allocation sequence. Because each geographical area started in the usual care group, and ended in the 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 quality improvement programme at their hospital. Because they were engaged in delivery of the intervention, it was not possible to mask hospital staff. Patients were masked to study group allocation. The organisation of hospitals into geographical clusters minimised any contamination between sites due to natural workforce movements between hospitals.

Trial intervention

The EPOCH trial care pathway was developed through an evidence based Delphi consensus process to update existing guidelines published by the Royal College of Surgeons of England.[8] A list of the 37 component interventions is provided in the supplementary file and Supplementary figure 1, and a full summary of evidence grading is available on the trial website (www.epochtrial.org). Because of the stepped-wedge trial design, the duration of the QI intervention varied between clusters from 5 to 80 weeks. We developed an evidence-based 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 a hospital wide improvement programme to implement the care pathway with the support and guidance of the national EPOCH QI team. The key features of the quality improvement methodology were 1) Reframing the high mortality rates for these patients as a 'social problem' requiring re-organisation of existing care processes rather than technical innovation; 2) Supporting QI leads to engage their frontline staff and executive leaders in the change process; 3) Training local QI leads in basic improvement skills based around the Model for Improvement;[19] and 4) Supporting teams to analyse and feed back key process measure data to their colleagues to drive change. The EPOCH QI team provided a one-day activation and education 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 leaders required to achieve change. Nominated QI Leads were informed 12 weeks before the date of activation to the intervention. Five weeks before activation, QI leads were sent a 'pre-activation' checklist which included planning a local stakeholder meeting, recruiting colleagues to their change teams and developing a presentation entitled 'Where we are now' including baseline data, local challenges and ideas for improvement to share at the cluster activation meeting. The EPOCH QI team provided further advice and support by phone and email. All QI resources, including data analysis tools, training materials and promotional documents were available online through a Virtual Learning Environment. Clusters were offered a half-day follow-up meeting 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 only eligible to attend these if their hospital had been activated to the trial intervention.

Outcome measures

The primary outcome measure was all-cause mortality within 90 days following surgery. Secondary outcomes were all-cause mortality within 180 days following surgery, duration of hospital stay after surgery and hospital re-admission within 180 days of surgery. We selected ten predefined process measures (key components of the care pathway) for inclusion in the main report: 1) consultant led decision to operate, 2) consultant review of patient before surgery, 3) pre-operative documentation of risk, 4) time from decision to operate to entry into operating theatre, 5) patient entered operating theatre within time-frame specified by their urgency (<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 & Hughes approach,[20] for an analysis with fixed time effects and random cluster effects, modified to exclude data collected during the five-week period in which the intervention commenced in individual clusters. Using Hospital Episodes Statistics data (www.epochtrial.org/protocol), we estimated that 27,540 eligible patients would be registered across 90 NHS hospitals over 85 weeks, with a 90-day mortality rate of 25% in the usual care group, and a between hospital coefficient of variation of 0.15. Assuming a constant case-load (18 patients per 5 weeks per hospital), independent hospital effects and a 5% significance level, the trial would have 92% power to detect a reduction in 90-day mortality from 25% to 22%. If the assumption of independent hospital effects was not met, and the 15 geographical clusters functioned effectively as 15 large hospitals, power would be reduced to 83%.

All analyses were conducted according to intention-to-treat principles. All eligible patients with available outcome data were included in the analysis, and analysed according to the randomisation schedule.[21] Patients who presented during the 5-week time period immediately after quality improvement activation were excluded from the analysis. Hospitals that initially agreed to participate but subsequently withdrew prior to the trial start date were excluded, however hospitals which withdrew after the trial start date, or did not implement the intervention, were included in the analysis. Hospitals which merged with other hospitals during the trial period were included in the analysis up to the point of the merger.

We were unable to procure data describing survival status after hospital discharge for patients in Wales. We therefore changed our primary analysis from binary to a time to event approach allowing inclusion of mortality events censored at hospital discharge. All analyses included time period as a fixed effect using indicator variables, and adjusted for age, gender, and indication for surgery using fixed factors.[22] Age was included as a continuous covariate, assuming a linear association with outcome.[23] Missing baseline data for indication for surgery were handled using a missing indicator approach.[24] 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 within hospital). This allowed additional correlation between patients in the same hospital and the same period, compared to patients in other periods, as is recommended.[25] All-cause mortality within 180 days was analysed using the same approach. Duration of hospital stay was analysed using competing risk time-to-event models, with mortality before the outcome event acting as the competing risk, and robust standard errors to account for clustering by geographical area. The hazard ratio from this analysis measures the relative probability of hospital discharge between treatment arms, with HR<1 indicating a lower probability of discharge in the QI group (and therefore longer hospital stay). Hospital readmission within 180 days was analysed using the same approach (with a HR<1 indicating a lower probability of re-admission).

Ethnographic study and process evaluation

As part of the wider EPOCH project, a prospective ethnographic evaluation was undertaken in six trial sites by researchers outside the main trial team. Ethnography draws on anthropological methods, including observation and interview, to provide a rich description of events that occur within a specific context. A maximum variation sample of sites was chosen with criteria focussed on size, surgical volume and discipline of the primary QI lead. A process evaluation was conducted to describe the delivery of the QI intervention. Data were collected describing the activity of QI teams and an exit questionnaire was completed by local QI leads to report their experience of the quality improvement process. All data were collected and analysed prior to the main trial analysis. Detailed methods are presented in the full reports.[26,27] In this report, we summarise key themes to provide the perspective needed to interpret our main findings.

Role of the funding source

The funder (National Institute for Health Research, Health Services & Delivery Research programme) had no role in study design, data collection, data analysis, data interpretation or writing of this report. The trial was sponsored by Queen Mary University of London. All authors had full access to the final dataset and approved the final submitted version of this report.

Results

Fifteen geographic areas underwent randomisation including 97 NHS hospitals. Four hospitals withdrew before the start of the trial, leaving 93 participating. Between 3rd March 2014 and 19th October 2015, 15,873 eligible patients underwent emergency abdominal surgery in participating hospitals with data recorded in the NELA database (8490 in the usual care group and 7383 in the QI group) (Figure 1). Baseline characteristics were similar between groups (Table 1).

Process measures

91/93 (98%) hospitals were represented at the initial QI meeting for the relevant geographical cluster and 53/93 (57%) were represented at the follow-up QI meeting. This representation included a named hospital QI lead for 89/93 (96%) hospitals at the first meeting and 47/93 (51%) hospitals at the second. Most meetings (n=13/15) occurred within two weeks of the activation date. Patient-level process measures are described in Table 2. In accordance with our analysis plan, we did not test these for statistical significance.

Clinical outcomes

Complete primary outcome data were available for more than 99% of patients (Figure 1, Supplementary tables 1 and 2). The primary outcome of 90-day mortality occurred in 1393 patients in the usual care group (16%) compared with 1210 patients in the QI group (16%) (Hazard ratio, QI vs usual care: 1.11 [0.96 to 1.28]) (Figure 2 and Table 3). Results were similar for mortality within 180 days (HR 1.12 [0.98 to 1.28]) (Supplementary figure 2). Patients in the QI group had a lower probability of hospital discharge (Hazard ratio for hospital discharge 0.90 [0.83 to 0.97]), leading to a marginally longer hospital stay (days in hospital, usual care: 8 [13 to 23] days vs. QI: 8 [13 to 24] days), although this difference was not clinically meaningful (Figure 3). There was no difference between groups in hospital re-admission within 180 days (usual care 1618 (20%) vs. QI 1242 (18%); Hazard ratio for re-admission 0.87 [0.73 to 1.04]) (Supplementary figure 3). In a secondary analysis, we found no evidence that the QI strategy became more effective the longer it had been adopted (Supplementary table 3). To assess the impact of missing mortality data following hospital discharge from patients in Wales, we assessed the number of mortality events which occurred after hospital discharge

but before 90 days in English and Scottish hospitals. Only 5% (631/13,034) of patients died between hospital discharge and 90 days, suggesting few outcome events in Wales were missed.

Quality improvement and ethnographic findings

Our prospective ethnographic study and process evaluation are reported in full elsewhere.[26,27] The findings showed that teams reflected positively on the QI programme, in particular the practical nature of the activation and education meetings, and the opportunity to share ideas and learn from others as well as the utility of the online resources. However, staff in each of the six sites studied encountered multiple challenges as they attempted to improve patient care during the intervention period and often had little or no additional time in their job plans to accommodate this. In particular, the task of collecting and entering data into the NELA database was more time consuming than expected. In addition, we observed differences in the fidelity with which teams used our recommended QI methods, differences in the clinical processes teams chose to attempt to change, the rate of this change and the eventual degree of success. Even amongst those sites that adhered to the QI intervention more closely, local adaptations to the care pathway were required to make this fit with the prevailing conditions of the hospital. The ethnographic evaluation confirmed the primarily social nature of the trial intervention. To a large extent, more successful QI teams drew on existing relationships within their hospital to influence colleagues and make change happen. Successful change seemed to be linked to the strength and number of these relationships; where these were lacking, additional effort was required to garner support for change. These findings suggest that whilst the QI programme may have provided QI leads and their teams with additional capabilities to lead change, the capacity to make change happen, especially in terms of protected time, was lacking. The extent to which the QI programme was delivered as intended, as well as enablers and barriers to change, are described in full in the report of the EPOCH trial process evaluation.[27]

Discussion

The principal finding of this trial was that there was no survival benefit associated with a national quality improvement programme to implement an evidence-based care pathway for patients undergoing emergency abdominal surgery. Furthermore, there was no beneficial effect on 180-day mortality, hospital stay or hospital readmission. At a national level, there were only modest improvements amongst the ten measures selected to reflect key processes of care within the pathway. In some cases, the baseline rate of adherence to process measures was higher than anticipated. Experience from individual hospitals suggested wide variations in which of the 37 pathway elements local QI teams chose to tackle, the rate of change they achieved, and their eventual success. The baseline contexts of participating hospitals also differed. Implementation of change was slower where existing relationships within and beyond the perioperative team were weaker, and so QI leads had to spend time developing relationships with stakeholders. At the time of trial design, the EPOCH care pathway was widely agreed to represent an achievable standard of care that informed clinicians would wish to deliver for their patients, but commonly failed to provide because of poor awareness amongst the perioperative team. Our findings reveal that implementation of such an extensive care pathway was a more complex challenge than expected by our clinical community. It is important to interpret the results of this trial alongside those of the ethnographic study and process evaluation,[26,27] which together suggest that quality improvement programmes designed to implement complex care pathways require more resources, with dedicated time for clinical teams to focus on making change happen.

There are several published reports of the impact of small scale quality improvement projects to improve outcomes for patients undergoing emergency abdominal surgery. In the UK, the ELPQuiC group examined the implementation of a care bundle of five interventions in four NHS hospitals in an uncontrolled before and after study.[4] They reported a reduction in mortality (risk ratio 0.61) amongst 726 patients. This study design is more prone to bias than a stepped-wedge cluster randomised trial.[28] The difference in findings may additionally relate to the simpler intervention, and stronger pre-existing relationships between staff leading implementation in these early adopter hospitals. The simpler objective was more readily achieved than that of the national EPOCH trial which set more ambitious targets in hospitals

where there may have been a less favourable context for change. Researchers from Denmark reported differing results from three separate studies of perioperative quality improvement interventions for patients undergoing emergency abdominal surgery. The PULP trial group used an uncontrolled before and after design with historical controls to study the effect of a 'multidisciplinary perioperative care protocol' in seven hospitals and reported a considerable reduction in 30-day mortality in comparison.[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) amongst 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 uncontrolled before and after study with historical controls, finding a more modest reduction in 30-day mortality from 22% amongst 600 control patients to 16% amongst 600 intervention patients.[7] It is possible that a background trend to improved mortality may explain the findings of these previous studies, especially given the growing international focus on poor patient outcomes following emergency abdominal surgery. Whilst our analysis accounts for temporal trends during the EPOCH trial, it is possible that a decreasing mortality beforehand may explain why the mortality rate was lower than that predicted from NHS registry data. Meanwhile, recent studies of quality improvement in other clinical areas have delivered mixed results.[29-32] These findings suggest that more focussed, discrete clinical interventions may be more successfully implemented than interventions that include larger numbers of care processes. The evidence is less clear in defining the optimal improvement methods. There are several theoretical models of implementation including the Consolidated Framework for Implementation Research and the COM-B model.[33,34] These provide frameworks for designing and evaluating effective implementation, clinical process and behaviour change. However, none of these models gives emphasis to institutional support or protected leadership time. Our findings suggest these more practical considerations are essential for clinicians to successfully lead quality improvement projects. In the EPOCH trial, teams were encouraged to begin with easier interventions, before building toward full pathway implementation. However, our process evaluation reveals that many

teams did not have the time or capacity to progress beyond simpler interventions (e.g. documentation of patient risk) to implementation of more important but challenging interventions such as admission to critical care. It is also important to note that the National Emergency Laparotomy Audit was launched only three months before the EPOCH trial commenced. Our 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 five-week period for the transition between usual care and the launch of the quality improvement programme in each cluster. Longer transition and intervention periods with dedicated time for QI leads to plan, negotiate and implement change may have led to more successful implementation. However, we also note that there was no evidence of survival benefit amongst hospitals exposed to the quality improvement programme for longer than 10 weeks, which included hospitals exposed for up to 80 weeks.

The strengths of this trial include wide generalisability (large number of consecutive patients enrolled by many hospitals), robust trial design and the devolved leadership to local clinical QI teams. The EPOCH care pathway was developed through a Delphi consensus process to update national professional guidelines.[8] As with many evidence-based treatment guidelines, some recommendations were graded as strong although the available evidence was weak. The choice of component interventions such as intensive care admission and consultant led care was primarily based on expert opinion; it is unclear how this evidence base could be improved. Partnership with the National Emergency Laparotomy Audit allowed an efficient trial design with no additional data collection for participating staff. However, our final dataset 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 datasets. On several occasions, organisations changed their position on information governance regulations, requiring revision of previous agreements between each of the parties involved. In hindsight, we would have encountered fewer problems had we confined the trial to the jurisdictions of fewer organisations with information governance oversight. Despite the large sample, fewer patients than expected underwent emergency abdominal surgery, and the 90-day mortality rate was lower than anticipated. The sample size calculation was based on Hospital Episodes Statistics data which do not provide a specific diagnostic code for emergency abdominal surgery. Instead we identified a series of codes for

relevant procedures. We chose to power the trial to detect a very modest treatment effect partly to accommodate the possibility that these data were poorly representative of the EPOCH trial population. However, the 95% confidence interval for our primary effect estimate was narrow, with a lower limit which indicates a maximum potential mortality reduction of 4%. Our findings are unlikely to change with a larger sample size. Due to difficulty in obtaining post-discharge survival data in Wales, we changed our primary analysis from a binary to a time to event approach allowing inclusion of mortality events censored at hospital discharge. However, post-discharge data from England and Scotland suggest few events were missed through this approach. The additional application required to obtain post-discharge mortality data for Wales would have further delayed the trial results by many months.

Conclusions

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

Contributions

CP, TS, GM, BK, AT, KR, DW, GR, SK, JB and RP all contributed to protocol development and design of the EPOCH trial. CP and TS designed and delivered the EPOCH Quality Improvement Programme. TS, RP and CP led the process evaluation with input from GM. GM led the ethnographic study with input from TS, CP and RP. RP, KE, AT, TS, CP, BK and SK were responsible for conduct of the trial. All authors read and approved the final manuscript. We wish to thank all members of the EPOCH trial group who are listed in the supplementary file.

Conflict of interest statement

RP holds research grants, has given lectures and/or performed consultancy work for BBraun, GlaxoSmithkline, Medtronic, Intersurgical and Edwards Lifesciences. CP has performed consultancy work for Merck and for the Institute for Healthcare Improvement. All other authors declare they have no conflicts of interest.

Acknowledgements

Data sharing

Due to information governance restrictions imposed by organisations governing data access, we are unable to share the trial data unless applicants secure the relevant permissions. All trial materials are freely available on the trial website (www.epochtrial.org).

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**(2): 249-57.

2.   NELA 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, Network UKEL. Variations in mortality after emergency laparotomy: the first report of the UK Emergency Laparotomy Network. *Br J Anaesth* 2012; **109**(3): 368-75.

4.   Huddart S, Peden CJ, Swart M, et al. Use of a pathway quality improvement care bundle to reduce mortality after emergency laparotomy. *Br J Surg* 2015; **102**(1): 57-66.

5.   Vester-Andersen M, Lundstrom LH, Moller MH, et al. Mortality and postoperative care pathways after emergency gastrointestinal surgery in 2904 patients: a population-based cohort study. *Br J Anaesth* 2014; **112**(5): 860-70.

6.   Moller MH, Adamsen S, Thomsen RW, Moller AM, Peptic Ulcer Perforation trial g. Multicentre trial of a perioperative protocol to reduce mortality in patients with peptic ulcer perforation. *Br J Surg* 2011; **98**(6): 802-10.

7.   Tengberg LT, Bay-Nielsen M, Bisgaard T, et al. Multidisciplinary perioperative protocol in patients undergoing acute high-risk abdominal surgery. *Br J Surg* 2017; **104**(4): 463-71.

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 of England, 2011.

9.   Haynes AB, Weiser TG, Berry WR, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009; **360**(5): 491-9.

10.   Kuper M, Gold SJ, Callow C, et al. Intraoperative fluid management guided by oesophageal Doppler monitoring. *BMJ* 2011; **342**: d3016.

11.   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**(2): 150-65.

12.   Schouten LM, Hulscher ME, van Everdingen JJ, Huijsman R, Grol RP. Evidence for the impact of quality improvement collaboratives: systematic review. *BMJ* 2008; **336**(7659): 1491-4.

13.   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**(9): 1045-51.

14.   Benning A, Dixon-Woods M, Nwulu U, et al. Multiple component patient safety intervention in English hospitals: controlled evaluation of second phase. *BMJ* 2011; **342**: d199.

15.   Batalden PB, Davidoff F. What is "quality improvement" and how can it transform healthcare? *Qual Saf Health Care* 2007; **16**(1): 2-3.

16.   Dixon-Woods M, Martin GP. Does quality improvement improve quality? *Future Hospital Journal* 2016; **3**(3): 191-4.

17.   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**(6): 608-13.

18.   Leatherman S, Sutherland K. The quest for quality in the NHS: refining the NHS reforms. Policy analysis and chartbook. . London: Nuffield Trust, 2008.

19.   Langley G, Moen R, Nolan K. The improvement guide: a practical approach to enhancing organizational performance. 2nd edition ed. San Francisco: Jossey-Bass; 2009.

20.   Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* 2007; **28**(2): 182-91.

21.   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.

22.   Kahan BC, Jairath V, Dore CJ, Morris TP. The risks and rewards of covariate adjustment in randomized trials: an assessment of 12 outcomes from 8 studies. *Trials* 2014; **15**: 139.

23.   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.

24.   White IR, Thompson SG. Adjusting for partially missing baseline measurements in randomized trials. *Stat Med* 2005; **24**(7): 993-1007.

25.   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**(2): 318-33.

26.   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 Illn* 2017; **39**(8): 1314-29.

27.   Stephens T, Peden C, Pearse RM, et al. Improving care at scale: Process evaluation of a complex quality improvement intervention to reduce mortality after emergency abdominal surgery (EPOCH Trial). *under review* 2018.

28.   Sedgwick P. Before and after study designs. *BMJ* 2014; **349**: g5074.

29.   Bion J, Richardson A, Hibbert P, et al. 'Matching Michigan': a 2-year stepped interventional programme to minimise central venous catheter-blood stream infections in intensive care units in England. *BMJ Qual Saf* 2013; **22**(2): 110-23.

30.   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**(7): e014333.

31.   Presseau J, Mackintosh J, Hawthorne G, et al. Cluster randomised controlled trial of a theory-based multiple behaviour change intervention aimed at healthcare professionals to improve their management of type 2 diabetes in primary care. *Implement Sci* 2018; **13**(1): 65.

32.   Williams L, Daggett V, Slaven JE, et al. A cluster-randomised quality improvement study to improve two inpatient stroke quality indicators. *BMJ Qual Saf* 2016; **25**(4): 257-64.

33.   Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci* 2009; **4**: 50.

34.   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.

Table 1. Baseline patient characteristics. Data presented as n (%) unless otherwise indicated.
ASA: American Society of Anesthesiologists physical status score; P-POSSUM: Portsmouth Physiological and Operative Severity Score for the enumeration of Mortality and morbidity score.

	Number of patients with missing data (n %)		Summary measure	
	Usual care (n=8490)	Quality improvement (n=7383)	Usual care	Quality improvement
Baseline characteristics				
Female	0 (0)	0 (0)	4550 (54)	3938 (53)
Age – mean (SD)	0 (0)	0 (0)	68 (13)	68 (13)
Indication for surgery	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)
Pre-operative characteristics				
Estimated risk of death	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	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 (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 (median [IQR])	4103 (48)	2870 (39)	1.6 (1.1-2.8)	1.5 (1.0-2.6)

Table 2. Patient level process measures. Data presented as n (%). [a] 29 patients in the usual care group and 27 patients in the QI group died during surgery.

Process measure	Number of patients with missing data (n %)		Summary measure	
	Usual care (n=8490)	Quality improvement (n=7383)	Usual care	Quality improvement
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)
Pre-operative 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 from decision to operate to entry into operating theatre (hours)	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)

Table 3. Patient outcomes. Data presented as median (IQR), n (%) or hazard ratio with 95% confidence intervals.

	Number of patients included in analysis		Summary outcome measure		
	Usual care (n=8490)	Quality improvement (n=7383)	Usual care	Quality improvement	Hazard ratio (QI vs. usual care)
All-cause mortality within 90 days of surgery	8482 (>99)	7374 (>99)	1393 (16)	1210 (16)	1.11 (0.96-1.28)
All-cause mortality within 180 days of surgery	8482 (>99)	7374 (>99)	1698 (20)	1440 (20)	1.12 (0.98-1.28)
Duration of hospital stay (days)	8320 (98)	7353 (>99)	8 (13-23)	8 (13-24)	0.90 (0.83-0.97)
Hospital re-admission within 180 days of surgery	7969 (94)	6723 (91)	1618 (20)	1242 (18)	0.87 (0.73-1.04)

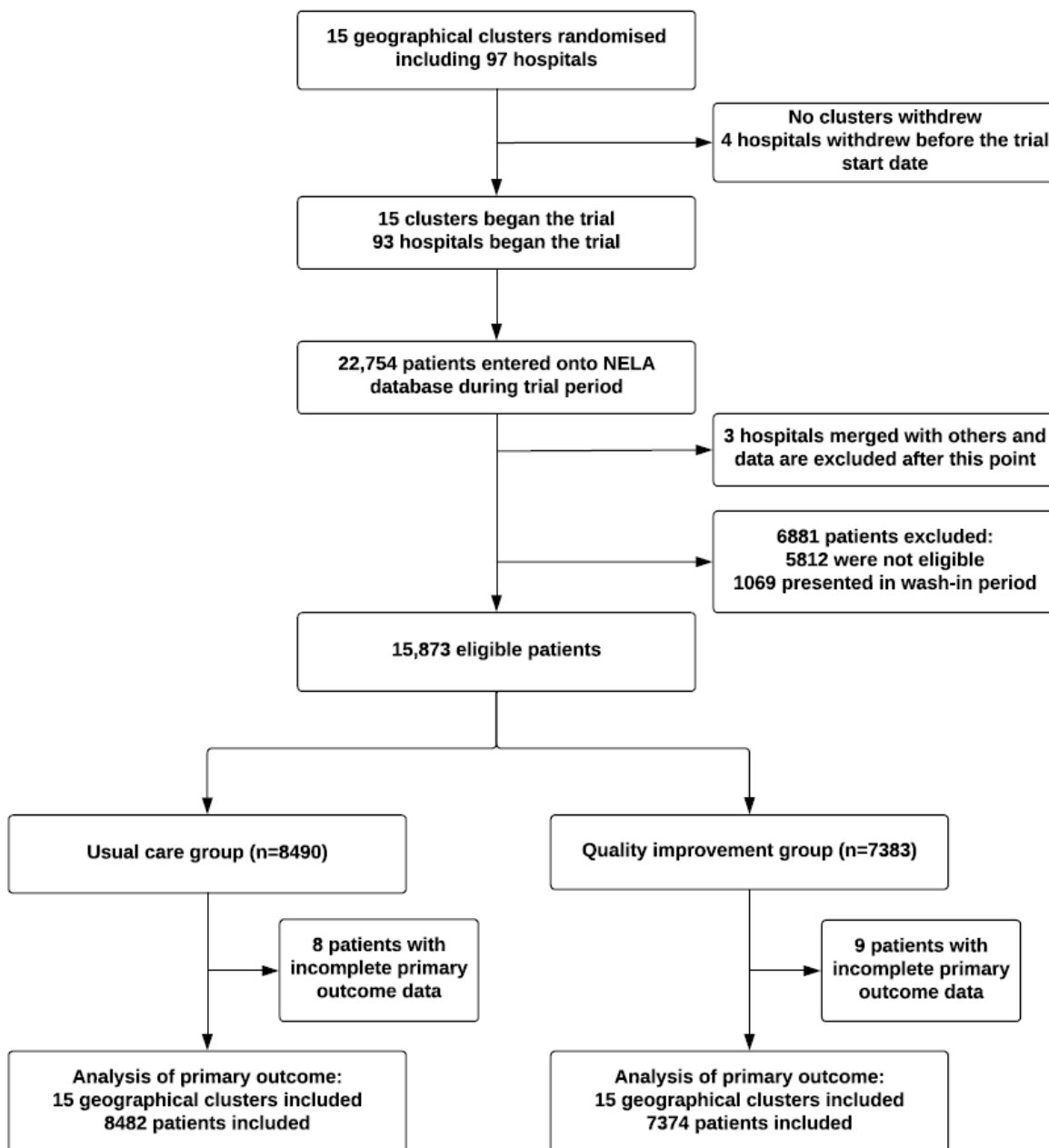**Figure 1. Inclusion of hospitals and patients in the trial.** NELA: National Emergency Laparotomy Audit.

Figure 2. Mortality within 90 days of emergency abdominal surgery. QI: quality improvement group.

Number at risk
Usual care 8482   7216   6968   6811
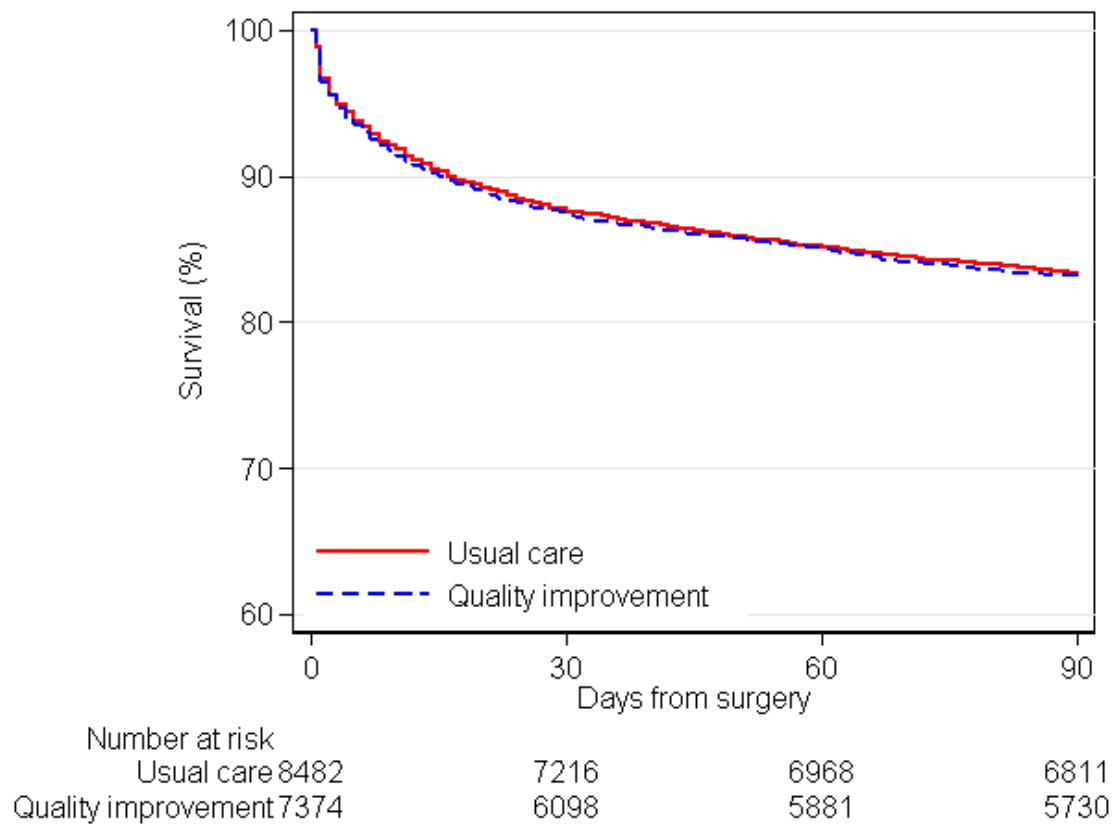Quality improvement 7374   6098   5881   5730

Figure 3. Duration of hospital stay after emergency abdominal surgery. QI: quality improvement group.

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