Safer delivery of surgical services: a programme of controlled before-and-after intervention studies with pre-planned pooled data analysis

Peter McCulloch,[1]* Lauren Morgan,[1] Lorna Flynn,[1] Oliver Rivero-Arias,[2] Graham Martin,[3] Gary Collins[4] and Steve New[5]

[1]Nuffield Department of Surgical Sciences, University of Oxford, John Radcliffe Hospital, Oxford, UK
[2]Nuffield Department of Population Health, University of Oxford, Oxford, UK
[3]Department of Health Sciences, University of Leicester, Leicester, UK
[4]Centre for Statistics in Medicine, University of Oxford, Botnar Research Centre, Oxford, UK
[5]Saïd Business School, University of Oxford, Oxford, UK

*Corresponding author

Declared competing interests of authors: Peter McCulloch is a member of the Health Technology Assessment panel on Interventional Procedures, and codirector of the Patient Safety Academy, sponsored by Health Education Thames Valley and the Oxford Academic Health Sciences Network.

Published December 2016
DOI: 10.3310/pgfar04190

Scientific summary

Safer delivery of surgical services

Programme Grants for Applied Research 2016; Vol. 4: No. 19
DOI: 10.3310/pgfar04190

NIHR Journals Library www.journalslibrary.nihr.ac.uk

Scientific summary

Background

The frequency and severity of inadvertent harm to patients in modern hospitals began to give rise to concern in the 1990s. A series of observational studies indicated that between 3% and 14% of patients were harmed by their care, with as many as 1 in 200 being killed by it. Analytical studies of typical incidents in the early 2000s revealed that serious harm was usually multifactorial, requiring the concurrence of a set of unrelated adverse factors, such as poor staff relations, faulty equipment, or failure to institute or follow safe procedures. The work of ergonomists and psychologists studying error and harm in other industries was reviewed to find a theoretical framework for understanding these findings, and the work of Reason and Helmreich in particular appeared relevant. Analogies with civil aviation became popular and schemes to train clinical personnel to communicate and co-operate effectively as part of a team were developed using the aviation crew resource management (CRM) model, although usually without proper evaluation. At the same time, interventions focused on rationalising systems of work to design error out were promoted by the Institute of Healthcare Improvement and other organisations. On reviewing the literature we noted that the direct evidence of benefit for most interventions to improve safety was weak and there was a disconnect between those which focused on systems improvement and those, such as teamwork training (TT), which focused on staff culture. We formed the hypothesis that interventions that addressed both aspects of the problem would be more successful than those which addressed only one. In addition, we sought to use qualitative methods to seek insights into the behavioural characteristics of improvement activity and to understand the obstacles that such initiatives face. This programme of work was developed to address the integrated intervention hypothesis together with this broader agenda of understanding mechanism in addition to outcome.

Objectives

- In operating theatre teams, does TT, to improve co-operation and communication, lead to improved technical and non-technical (teamwork) performance?
- In such teams, does quality improvement (QI), based on the Toyota lean system, lead to improved technical and non-technical (teamwork) performance?
- In such teams, does the co-operative design of a standard operating procedure (SOP) system lead to improved technical and non-technical (teamwork) performance?
- In such teams, do combinations of systems improvement (using either lean or SOP approaches) together with culture change via TT result in a better overall clinical outcome than TT, lean systems or SOP alone?
- What factors (apart from deliberate interventions) affect teamwork behaviour in the operating theatre?
- What factors (apart from deliberate interventions) affect team technical performance in the operating theatre?
- What are the challenges and obstacles for implementing safety interventions in a NHS hospital setting?
- What are the underlying mechanisms which influence the context for safety interventions?
- What are the costs of developing and implementing specific safety interventions and can they be assessed alongside identifiable benefits?

Methods

Study design

The programme was designed as a suite of controlled interrupted time series experiments, using identical methods for outcome evaluation and delivering the same three interventions in different combinations to improve safety and reliability in operating theatres. This was followed by a final 'all-systems'

before-and-after study combining learning from the previous studies to intervene in a surgical service not only in the operating theatre but throughout the patient pathway. Each individual study had pre- and post-intervention observation periods of 3–4 months and an intervening intervention period of 4 months.

Settings

We studied operating theatres at five sites in three trusts; the majority were performing routine orthopaedic procedures, but we also studied vascular, trauma and plastic surgery teams. The final study was performed on a regional neurosurgery unit.

Data collected

We evaluated theatre team technical performance using the 'glitch count' method, which we derived and validated from previous work by our group. We evaluated non-technical performance using the Oxford Non-Technical Skills (NOTECHS) II scale score, a modification of our previous validated method for doing this. Both of these methods require real-time observation of the entire procedure by a pair of observers: one clinically trained and the other trained in human factors. The same pairs observed the time-out (T/O) and sign-out (S/O) procedures of the World Health Organization (WHO)'s surgical safety checklist and recorded compliance with it. Data were subsequently collected from hospital records on patient outcomes including length of stay, return to hospital and to theatre, death within 30 days, recorded complications and recorded patient safety incidents. The European Quality of Life-5 Dimensions questionnaire was administered at 6 months after surgery to a subset of patients to assess quality of life. Theatre process data and costings were collected to allow economic analysis. The final study was conducted mainly on surgical wards and generated a variety of problem-specific process data. A qualitative study of the process of intervention across studies was performed using semistructured interviews, with researchers and participating clinical staff, and analysed using an approach based on the constant comparison method.

Interventions

The three interventions used were (1) TT using the CRM model from aviation; (2) SOP development (staff were encouraged to develop a formal SOP for their work with support from ergonomists); and (3) lean-style systems QI. Two additional experiments studied combinations of interventions (4) SOP and TT and (5) lean systems and TT. Each of these interventions began with a 1- to 2-day training course for all staff involved in a theatre team and were followed up by 6–8 weeks of support and coaching. For the final project, the same format of training was used, but elements of all three intervention approaches were integrated.

Statistical analysis

We used conventional probability testing, regression and correlation methods to study associations between potential confounders (e.g. hospital site) and outcomes (e.g. glitch rate). We used two-way analysis of variance with time as one factor to compare the changes in outcomes in the active groups before and after the intervention with those in the control group. We used individual patient pooled analysis to synthesise the results of the five similar studies.

Results

The development work on the scales used demonstrated appropriate reliability and validity for both NOTECHS II and the glitch count. NOTECHS II scale score correlated with WHO checklist performance as expected but not with glitch count, giving us orthogonal measures for technical and non-technical performance. Both NOTECHS II scale score and glitch count showed considerable baseline variation, and for NOTECHS II scale score we identified that surgical specialty appeared to affect average NOTECHS II scale score but hospital site did not. For glitch count the opposite was noted: hospital site predicted glitch rate while specialty did not. We noted that 40% of glitches happen in the first quarter of an operation and only 10% in the final quarter.

We observed poor compliance with the WHO T/O and S/O procedures. In a sample of 294 operations, T/O was attempted in 86.7%, but S/O in only 8.9%. Full compliance with T/O procedures (all staff present, all items covered, active participation of team) occurred in 38% of the operations observed. Hospital site predicted the level of compliance, but surgical specialty did not.

Teamwork training alone produced a significant improvement in the NOTECHS II scale scores, but glitch count scores declined. WHO T/O performance rose in both active and control groups, suggesting some bias from contamination, but S/O improved only in the active group.

Standard operating procedure development alone did little to change team performance: WHO checklist compliance and NOTECHS II scale scores were unchanged and glitch counts deteriorated in both active and control groups.

Lean QI alone did not affect the NOTECHS II scale score, WHO checklist compliance or glitch count.

Standard operating procedure plus TT improved the NOTECHS II scale scores substantially and improved S/O performance compared with controls. There was no apparent effect on T/O performance or glitch rate.

Lean plus TT improved both NOTECHS II scale score and glitch count substantially and led to an improvement in WHO S/O compliance, although T/O compliance was unaffected.

Five out of 15 projects in the final summary study combining all intervention approaches yielded measurable improvements in process or outcome. A project to reduce patient falls succeeded in reducing the rate by 50%. Projects to improve ward rounds and communication, to rationalise and prioritise urgent tertiary referrals, to improve the management of subarachnoid haemorrhage and to improve learning and feedback among surgical staff also yielded evidence of success in varying degrees.

Pooled analysis confirmed the additive value of TT together with systems improvement methods, particularly lean systems. Qualitative analysis identified major barriers to safety improvement in attitudes and incentives, hospital management structures and professional culture and highlighted challenges relating to the ambiguous role of an external team in facilitating QI. Complexity, uncertainty and rapid change also contributed to the challenges, as did financial constraints and performance targets. An implementation strategy designed to address these problems was developed and used in the neurosurgery study, where it seemed moderately effective.

Conclusions

Measures of technical and non-technical team performance reveal wide variations among teams. Some of these appear related to the hospital site and others to the surgical specialty involved. Although the needs of specific types of surgery may produce unavoidable differences in team performance, those related to hospital site may be related to local culture or working conditions and need to be addressed.

Brief staff training interventions with follow-up support at the micro-system level can be effective in improving team performance in both technical and non-technical skills related to patient safety. The combination of TT with an intervention focused on systems improvement is more effective than either TT alone or systems improvement alone. Only TT plus lean QI produced substantial improvement in all three outcomes.

The selection of appropriate outcome measures in complex intervention programmes with an element of staff autonomy is problematic, because interventions need to demonstrate impact on overall objectives, but measures directed at these may be inappropriate for demonstrating the effectiveness of process improvements selected by staff. A combination of prespecified programme outcome measures and individual project measures is required.

There is considerable room for improvement in compliance with the WHO checklist and the S/O procedure as currently performed seems poorly fitted for purpose and may need to be revised.

Introducing systems change in NHS surgical units is challenging for cultural, structural, financial and workload-related reasons. An explicit implementation strategy is required to optimise the chances of success.

Future research

Further research is required to:

1. Determine the impact of integrated systems and culture improvement training for surgical staff on clinical outcomes for patients. This will require a large cluster randomised trial.
2. Describe and understand the mechanisms underlying barriers to beneficial change in NHS surgical units, and how they can be overcome.
3. Test methods for disseminating the integrated training programme at scale: a 'train the trainers' approach seems most appropriate, but will require significant ergonomic support for trainers.
4. Define the optimum implementation strategy for theatre safety interventions.

Funding

Funding for this study was provided by the Programme Grants for Applied Research programme of the National Institute for Health Research.

Programme Grants for Applied Research

ISSN 2050-4322 (Print)

ISSN 2050-4330 (Online)

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

Editorial contact: nihredit@southampton.ac.uk

The full PGfAR archive is freely available to view online at www.journalslibrary.nihr.ac.uk/pgfar. 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 *Programme Grants for Applied Research* journal

Reports are published in *Programme Grants for Applied Research* (PGfAR) if (1) they have resulted from work for the PGfAR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Programme Grants for Applied Research programme

The Programme Grants for Applied Research (PGfAR) programme, part of the National Institute for Health Research (NIHR), was set up in 2006 to produce independent research findings that will have practical application for the benefit of patients and the NHS in the relatively near future. The Programme is managed by the NIHR Central Commissioning Facility (CCF) with strategic input from the Programme Director.

The programme is a national response mode funding scheme that aims to provide evidence to improve health outcomes in England through promotion of health, prevention of ill health, and optimal disease management (including safety and quality), with particular emphasis on conditions causing significant disease burden.

For more information about the PGfAR programme please visit the website: http://www.nihr.ac.uk/funding/programme-grants-for-applied-research.htm

This report

The research reported in this issue of the journal was funded by PGfAR as project number RP-PG-0108-10020. The contractual start date was in November 2009. The final report began editorial review in March 2015 and was accepted for publication in December 2015. As the funder, the PGfAR programme agreed the research questions and study designs in advance with the investigators. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The PGfAR 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, CCF, NETSCC, PGfAR or the Department of Health. 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 PGfAR programme or the Department of Health.

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