

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

Peter McCulloch, Lauren Morgan, Lorna Flynn, Oliver Rivero-Arias, Graham Martin, Gary Collins and Steve New



***National Institute for
Health Research***

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 Collins⁴
and Steve New⁵

¹Nuffield Department of Surgical Sciences, University of Oxford, John Radcliffe Hospital, Oxford, UK

²Nuffield Department of Population Health, University of Oxford, Oxford, UK

³Department of Health Sciences, University of Leicester, Leicester, UK

⁴Centre for Statistics in Medicine, University of Oxford, Botnar Research Centre, Oxford, UK

⁵Saïd 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

This report should be referenced as follows:

McCulloch P, Morgan L, Flynn L, Rivero-Arias O, Martin G, Collins G, *et al*. Safer delivery of surgical services: a programme of controlled before-and-after intervention studies with pre-planned pooled data analysis. *Programme Grants Appl Res* 2016;**4**(19).

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.

© Queen's Printer and Controller of HMSO 2016. This work was produced by McCulloch *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. 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).

Programme Grants for Applied Research Editor-in-Chief

Professor Paul Little Professor of Primary Care Research, University of Southampton, UK

NIHR Journals Library Editor-in-Chief

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the EME Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andree Le May Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Professor Aileen Clarke Professor of Public Health and Health Services Research, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

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

Professor Geoffrey Meads Professor of Health Sciences Research, Health and Wellbeing Research Group, University of Winchester, UK

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

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), 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 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 Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

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

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: nihredit@southampton.ac.uk

Abstract

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 Collins⁴ and Steve New⁵

¹Nuffield Department of Surgical Sciences, University of Oxford, John Radcliffe Hospital, Oxford, UK

²Nuffield Department of Population Health, University of Oxford, Oxford, UK

³Department of Health Sciences, University of Leicester, Leicester, UK

⁴Centre for Statistics in Medicine, University of Oxford, Botnar Research Centre, Oxford, UK

⁵Saïd Business School, University of Oxford, Oxford, UK

*Corresponding author peter.mcculloch@nds.ox.ac.uk

Background: High rates of iatrogenic harm have been confirmed in observational studies of surgery. Most interventions designed to reduce this have been targeted at either workplace culture or operational systems. We hypothesised that an integrated intervention addressing both culture and system might be more effective than either approach alone.

Objective: To evaluate interventions designed to improve surgical team performance by impacting culture or systems in isolation or combination.

Design: Five controlled intervention experiments, addressing system, culture or both, were performed in operating theatres. A final whole-system intervention study integrated approaches that showed benefit in these experiments. The five linked studies were subjected to a pre-planned pooled analysis to identify the effects of interventions, combinations and confounders. A qualitative interview study provided explanatory data on the mechanisms of intervention success and failure. An economic analysis was conducted.

Setting: Operating theatres in five hospitals, performing orthopaedic, trauma, vascular and plastic surgery were used for the linked studies. The final study occurred in a tertiary referral neurosurgery unit.

Participants: The main study subjects were clinical staff. Patient outcomes, both clinical and patient reported, were collected as secondary outcome measures.

Interventions: The interventions tested were (1) teamwork training (TT) based on the aviation crew resource management model, (2) the development of a set of standard operating procedures (SOPs), (3) a safety improvement programme based on lean principles, (4) TT plus SOPs and (5) TT plus lean. The final intervention used elements of all three strategies.

Main outcome measures: Primary outcomes were team non-technical skills [as measured by the Oxford Non-Technical Skills (NOTECHS) II scale score] and team technical performance (via the 'glitch count'). Secondary outcomes were compliance with the World Health Organization (WHO)'s checklist procedures, patient length of stay, readmissions, 30-day mortality, complications and patient-reported outcome measures [as measured by the European Quality of Life-5 Dimensions (EQ-5D)]. A qualitative interview study provided explanatory data on the mechanisms of intervention success and failure. An economic analysis was conducted.

Data sources: Direct observation of whole operations, clinical records, hospital information systems and EQ-5D questionnaires. The qualitative study used semistructured interviews.

Statistical methods: Individual studies were analysed using two-way analysis of variance, and an overall individual patient pooled analysis was performed. Methods validation studies and other analyses used chi-squared test, correlation and regression methods as appropriate.

Results: We studied 453 operations. The results of single interventions were inconsistent. TT alone improved non-technical skills and WHO compliance ($p < 0.001$) but not technical performance, whereas the systems interventions (lean and SOP) improved non-technical skills and technical performance ($p < 0.001$), but were less effective in improving WHO compliance. The integrated intervention approaches improved all aspects of team performance except time-out attempt rate, whereas the single approaches were significantly poorer at improving checklist compliance ($p < 0.001$) and failed to improve glitch rate. Combining all three strategies did not increase the percentage of successful projects. The qualitative analysis confirmed that integrated interventions better addressed the breadth of challenges that face surgical safety but also indicated that differences in implementation between integrated- and single-intervention studies amplified their differential effect.

Conclusions: A combination of TT plus systems improvement training appears more effective in improving team performance than either approach alone. An implementation strategy based on an understanding of the barriers to change in hospitals is important for success.

Future work: More work is required to understand and measure barriers to safety improvement. Implementation strategies need to be tested empirically. Methods for delivering integrated interventions on a larger scale need development. A cluster randomised trial of the integrated-systems/culture-improvement approach is warranted.

Funding: The National Institute for Health Research Programme Grants for Applied Research programme.

Contents

List of tables	xiii
List of figures	xvii
List of boxes	xix
List of abbreviations	xxi
Plain English summary	xxiii
Scientific summary	xxv
Chapter 1 Introduction	1
Recognition and quantification of patient harm: history	1
Analysis of causes: human factors	2
Focus on surgery: specific problems	3
Attempts to correct problems	4
Models of harm in health care and their implications: critique	5
Chapter 2 Study design	7
Study design background	7
Study design and rationale	8
Study logistics	9
Study settings	10
Initial liaison and manner of working with trusts	11
Study team and expertise	11
Final programme	12
<i>Primary hypotheses</i>	12
<i>Secondary hypotheses</i>	13
<i>Safer delivery of surgical services workstreams</i>	13
Chapter 3 Methods	15
Observational methods	15
<i>Methodological approach</i>	15
<i>Observer background and training</i>	15
Non-technical skills	15
<i>Development of Oxford Non-Technical Skills II scale</i>	16
<i>Validation</i>	17
<i>Method for Oxford Non-Technical Skills II scale</i>	18
Glitch count	19
<i>Description of the glitch count method</i>	19
<i>Development of the glitch count</i>	20
<i>Reliability and validity of the glitch count</i>	20
The World Health Organization's checklist evaluation method	21
<i>Development of the World Health Organization's checklist evaluation</i>	21
<i>Description of the World Health Organization's checklist evaluation method</i>	22
Clinical outcome measures	22
Patient-reported outcome measures	23

Health economic evaluation methods	23
Qualitative methods	23
<i>Design</i>	23
<i>Development of the interview guide</i>	23
<i>Participants</i>	23
<i>Interviews</i>	23
<i>Analysis</i>	23
Statistical methods	24
General approach to interventions	24
Lean	25
<i>Background theory and content delivered</i>	25
<i>Training and follow-up support approach</i>	26
Standard operating procedures	26
<i>Background theory and content delivered</i>	26
<i>Training and follow-up support approach</i>	26
Teamwork training	27
<i>Background theory and content delivered</i>	27
<i>Training and follow-up support approach</i>	28
Chapter 4 Standard operating procedures: study A	29
Introduction	29
Methods	29
<i>Design</i>	29
<i>Setting</i>	29
<i>Interventions</i>	29
<i>Primary intervention</i>	29
<i>Secondary intervention</i>	29
<i>Evaluation</i>	30
<i>Data analysis</i>	30
Results	30
<i>Primary intervention</i>	30
<i>Secondary intervention</i>	32
Discussion	32
Chapter 5 Teamwork training: study B	35
Introduction	35
Methods	35
<i>Design</i>	35
<i>Setting</i>	35
<i>Intervention</i>	36
<i>Evaluation</i>	36
<i>Data analysis</i>	37
Results	37
<i>Overview</i>	37
<i>World Health Organization checklist compliance</i>	37
<i>Non-Technical Skills II scale</i>	38
<i>Glitch rate</i>	38
<i>Clinical outcomes</i>	38
Discussion	39

Chapter 6 Lean process improvement: study C	41
Introduction	41
Methods	41
<i>Design</i>	41
<i>Setting</i>	41
<i>Interventions</i>	42
<i>Evaluation</i>	42
<i>Data analysis</i>	43
Results	43
<i>Primary intervention</i>	43
<i>Secondary intervention</i>	44
Discussion	46
Chapter 7 Standard operating procedures and teamwork training: study A and study B	49
Introduction	49
Methods	49
<i>Design</i>	49
<i>Setting</i>	49
<i>Interventions</i>	49
<i>Primary intervention</i>	49
<i>Secondary intervention</i>	50
Evaluation	50
<i>Primary intervention</i>	50
<i>Secondary intervention</i>	50
Data analysis	50
Results	50
<i>Primary intervention</i>	50
<i>Secondary intervention</i>	51
Discussion	52
Chapter 8 Teamwork training and lean process improvement: study B and study C	55
Introduction	55
Methods	55
<i>Design</i>	55
<i>Setting</i>	55
<i>Interventions</i>	55
<i>Primary intervention</i>	55
<i>Secondary intervention</i>	56
Evaluation	56
<i>Primary intervention</i>	56
<i>Secondary intervention</i>	56
Data analysis	56
Results	56
<i>Primary intervention</i>	56
<i>Secondary intervention</i>	58
Discussion	58

Chapter 9 Integrated interventions: study A, study B and study C	61
Introduction	61
Methods	61
<i>Design</i>	61
<i>Setting</i>	61
<i>Interventions</i>	62
Evaluation	62
<i>Primary intervention</i>	62
<i>Secondary intervention</i>	63
<i>Data analysis</i>	63
Results	63
<i>Primary intervention</i>	63
<i>Secondary interventions</i>	64
Discussion	78
<i>Intervention and implementation strategy</i>	81
<i>Limitations on success: investigator strategy</i>	81
<i>Limitations on success: culture and incentives for key staff</i>	82
<i>Other external factors</i>	83
Conclusions	83
Chapter 10 Qualitative analysis	85
Introduction	85
Methods	85
<i>Design</i>	85
<i>Participants</i>	85
<i>Analysis</i>	85
Results	85
<i>The importance of senior-level clinician buy-in</i>	86
<i>Ability of staff to undertake quality improvement work</i>	86
<i>Sustaining change</i>	87
<i>Incentives</i>	88
<i>Effort required versus rewards received</i>	88
<i>'Learned helplessness'</i>	89
<i>Change fatigue</i>	89
<i>Staff issues with lean</i>	90
<i>Complacency</i>	90
<i>Time</i>	91
<i>Opportunity</i>	92
<i>The external team</i>	93
<i>Team rapport</i>	93
<i>Team support</i>	94
Conclusion	95
Chapter 11 Health economic analysis	97
Introduction	97
Objectives	97
Methods	97
<i>Deviations from original proposal and final data sources</i>	97
<i>Aggregate-level teamwork training data (point 1)</i>	98
<i>Patient-level hospital-related resource data (point 2)</i>	98
<i>Cost analysis of training programmes</i>	99
<i>Hospital-related data statistical analysis</i>	100

Results	100
<i>Training programme costs</i>	100
<i>Hospital-related data</i>	101
Discussion	108
Chapter 12 Final synthesis and analysis	115
Measurement	115
Interventions	116
Implementation strategy	117
Critique	118
Chapter 13 Learning and development	121
Introduction	121
What have we learned?	121
<i>Engagement</i>	121
<i>Sustainability</i>	122
<i>Staff-led improvement</i>	122
<i>Single compared with integrated interventions</i>	123
How have we developed?	123
Lessons: for researchers studying safety and quality intervention strategies	124
<i>External team status</i>	124
<i>Maintenance of objectivity in evaluation</i>	124
<i>Practicalities of data collection and evaluation</i>	124
<i>Logistics and communication</i>	125
Lessons: for independent quality improvement groups	125
<i>Organisational context</i>	125
<i>Relationships with senior management</i>	125
<i>Using evidence and data</i>	125
<i>Building rapport</i>	126
<i>Technical support from intervention team</i>	126
<i>Incentives</i>	126
<i>Making benefits visible</i>	127
<i>Multiple projects</i>	127
<i>Project focus</i>	127
Lessons: trust management	128
<i>Staff-led improvement compared with external consultants</i>	128
<i>Management role</i>	128
<i>Opportunity</i>	129
<i>Resourcing</i>	129
<i>Policy</i>	129
Lessons: national training bodies	130
<i>Frontline staff training</i>	130
Lessons: national focus	130
<i>We can do better</i>	130
<i>The problem with targets</i>	131
<i>Measurement</i>	131
Chapter 14 Overall conclusions	133
Our hypothesis	133
Why change is hard	133
What can be done now?	134
What are the next questions?	134

Chapter 15 Knowledge translation	137
Introduction	137
Publications	137
Conferences	137
Work in our trust	137
<i>Pharmacy</i>	137
<i>System for Electronic Notification and Documentation</i>	137
<i>Oxford Acute Referral System</i>	140
<i>Staff-implemented feedback survey</i>	140
Local and national dissemination	140
Teaching	140
Future developments	140
<i>The Patient Safety Academy</i>	140
Summary	142
Chapter 16 Recommendations for future research	143
Models for upscaling the integrated quality intervention programme	143
Addition of a technology arm	144
Analysis of organisational barriers to system change	144
Action research on implementation strategies	145
Experimental compared with observational studies	145
Acknowledgements	147
References	149
Appendix 1 Intentional rounding area sheet	155
Appendix 2 Intentional rounding individual patient sheet	157
Appendix 3 Ward round data collection pro forma	159
Appendix 4 Subarachnoid haemorrhage care project case note template	163
Appendix 5 Subarachnoid haemorrhage care project process map	165
Appendix 6 Subarachnoid haemorrhage pro forma	167
Appendix 7 Senior house officer leavers survey (August 2014)	169

List of tables

TABLE 1 Schematic of intervention phases	9
TABLE 2 Description of trust site and surgical interactions	11
TABLE 3 The operating theatre team Oxford NOTECHS assessment tool	16
TABLE 4 The scoring scheme for NOTECHS II	17
TABLE 5 Inter-rater agreement	18
TABLE 6 Example data entry sheet for NOTECHS II	18
TABLE 7 Glitch categories with definition and examples	19
TABLE 8 Difference in observed glitches between observer specialties (<i>p</i> -values based on chi-squared test)	20
TABLE 9 Agreement between observers on observed glitches	21
TABLE 10 Adherence to the WHO's checklist	22
TABLE 11 World Health Organization checklist completion	31
TABLE 12 Primary intervention results	31
TABLE 13 World Health Organization checklist completion	38
TABLE 14 Intervention results	38
TABLE 15 World Health Organization checklist completion	44
TABLE 16 Primary intervention results	45
TABLE 17 Secondary intervention results	45
TABLE 18 World Health Organization checklist completion	51
TABLE 19 Primary intervention results	52
TABLE 20 World Health Organization surgical safety checklist compliance	57
TABLE 21 Summary of outcome measures	57
TABLE 22 Projects arising from the intervention	58
TABLE 23 European Quality of Life-5 Dimensions scores	63
TABLE 24 Plan–do–check–act cycles	64

TABLE 25 Total number of falls per intervention	67
TABLE 26 Ward round changes	68
TABLE 27 Audit results	75
TABLE 28 Morbidity and mortality post-intervention results	79
TABLE 29 Minor secondary interventions	80
TABLE 30 Discontinued secondary interventions	80
TABLE 31 Hospital frontline staff interviewed	85
TABLE 32 Resource-use data needed to calculate the costs of the interventions	99
TABLE 33 Standard UK orthopaedic operating theatre staff and invested training time	99
TABLE 34 Non-research and research staff preparation time of course set-up	100
TABLE 35 Resource-use and associated costs related to the TT programme	102
TABLE 36 Resource-use and associated costs related to the lean training	102
TABLE 37 Resource-use and associated costs related to the SOP training programme	103
TABLE 38 Frequency (percentage) of operations performed by consultants participating in the S3 project during the study period by intervention before and after change and in the active and control groups	105
TABLE 39 Mean (SD) age (years) across interventions before and after change and in the active and control groups, and parametric <i>t</i> -test comparison before and after change	105
TABLE 40 The frequency (percentage) of males across interventions before and after change and in the active and control groups, and chi-squared test comparison before and after change	106
TABLE 41 Frequency (percentage) of patients readmitted within 90 days across interventions before and after change and active and control groups. Before and after comparison using adjusted odds ratio (95% CI)	106
TABLE 42 Frequency (percentage) of patients with at least one complication across interventions before and after change and in the active and control groups. The before and after comparison uses the adjusted odds ratio (95% CI)	108
TABLE 43 The mean (SD, minimum, maximum) frequency of complications before and after change and in the active and control groups. The before and after change mean difference (95% CI) was adjusted for age and sex using linear regression	110

TABLE 44 The mean (SD, minimum, maximum) length of stay (in days) across interventions before and after change and in the active and control groups. The before and after change mean difference (95% CI) was adjusted for age and sex using linear regression	112
TABLE 45 The S3 study publications to date	138
TABLE 46 Safer Delivery of Surgical Services conferences	139
TABLE 47 Safer Delivery of Surgical Services teaching	141

List of figures

FIGURE 1 Schematic of studies	6
FIGURE 2 Oxford NOTECHS II scale score: HF observer and clinical observer	17
FIGURE 3 Nurse/care support workers visits to patients observed	65
FIGURE 4 Completion and documentation of intentional rounding	66
FIGURE 5 Weekly incidence of falls on the intervention ward	66
FIGURE 6 Weekly incidence of falls on the control wards ($n = 38$ wards)	67
FIGURE 7 Ward round documentation with/without pro forma	69
FIGURE 8 Documentation on ward round	70
FIGURE 9 Roles present during ward round	71
FIGURE 10 Post-operative consideration during ward round	72
FIGURE 11 Glitches observed during ward round	72
FIGURE 12 Driver diagram for MMMs	77
FIGURE 13 Flow diagram of hospital-related data available at the individual patient level available in the S3 study	104
FIGURE 14 The before and after change adjusted odds ratio and associated 95% CI of patients readmitted within 90 days across the active and control groups	107
FIGURE 15 The before and after change adjusted odds ratio and associated 95% CI of patients with at least one complication across the active and control groups	109
FIGURE 16 The before and after change adjusted mean difference and associated 95% CI of the frequency of complications across active and control groups	111
FIGURE 17 Before and after change adjusted mean difference and associated 95% CI of length of stay days across active and control groups	113

List of boxes

BOX 1 Choices to consider when constructing a SOP	27
BOX 2 Morbidity and mortality meeting outcome measures	78

List of abbreviations

3D	three-dimensional	NOTECHS	Non-Technical Skills
ANOVA	analysis of variance	PDCA	plan–do–check–act
CI	confidence interval	PROM	patient-reported outcome measure
CORESS	Confidential Reporting System for Surgery	PSA	Patient Safety Academy
CRM	crew resource management	QI	quality improvement
CV	curriculum vitae	S/O	sign-out
EPR	electronic patient record	S3	Safer Delivery of Surgical Services
EQ-5D	European Quality of Life-5 Dimensions	SAH	subarachnoid haemorrhage
HF	human factor	SEIPS	Systems Engineering Initiative for Patient Safety
IT	information technology	SEND	System for Electronic Notification and Documentation
ITU	intensive therapy unit	SHO	senior house officer
MDT	multidisciplinary team	SOP	standard operating procedure
MMM	morbidity and mortality meeting	T/O	time-out
NIHR	National Institute for Health Research	T&T	track and trigger
NOPEs	non-operative procedural errors system	TT	teamwork training
		WHO	World Health Organization

Plain English summary

Modern surgical treatment carries high risks of inadvertent patient harm, estimated at around 10%. Previous research has identified risky work systems and poor teamwork and communication as the main causes. Interventions have generally tried to correct either the teamwork problems or the system faults. We conducted five experiments, using three different interventions alone or in combination, to see whether or not tackling both system and teamwork faults would be more effective than dealing with either alone. We studied operating theatre teams performing surgery in five hospitals and measured their teamwork and technical performance by directly observing them. We used a training programme from aviation to improve teamwork and either a standard operating protocol (SOP) or the Toyota lean quality improvement process to correct system faults. We interviewed staff and researchers to find out more about what worked and why.

We found that teamwork training (TT) always improved teamwork performance, but on its own did not improve technical performance. Neither SOP nor lean systems alone improved team performance at all. The best results were obtained with integrated training. Only TT plus lean systems improvement caused major improvements in both technical and teamwork performance.

Our interview studies revealed a great deal about barriers to change in the NHS and how to overcome them, and about why integrated interventions work better.

Our studies were designed to measure the effects of the training on team performance. There are many other things that can affect patient outcomes and so we were not surprised that we were not able to prove that the interventions made things better for patients. To show this would require a much larger study. We recommend a large trial of lean systems integrated with TT to see if this can show a difference in patient outcomes.

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.

Chapter 1 Introduction

Recognition and quantification of patient harm: history

Health-care workers, since time immemorial, have had to recognise that their efforts are ultimately doomed to failure. The mortality rate among human beings is 100% and, although much can be done to extend life and relieve suffering, deterioration in function and death are, in the final analysis, unavoidable. Hippocrates taught that the most important function of a physician was not diagnosis or treatment but prognosis: allowing the patient to know whether they could expect to survive or whether they should make plans for their own death. Early medical and surgical treatments were of limited efficacy, and this was well recognised for millennia, leading human cultures for most of recorded history to put more faith in the power of religion or magic than in the powers of those who used chemical (herbal medicine) or physical (surgical) methods. In this context, the attitude of patients generally gave physicians a paradoxically privileged position in regard to the results of their efforts. As the expectation of cure was generally low, any success was likely to receive praise and reward, whereas most failures were attributed to the inevitable course of nature, fate or the will of divinities. This did not relieve physicians from all risk of blame by any means: the risk of deliberate or accidental poisoning by physicians is recognised in some stories and histories from ancient times, and the Code of Hammurabi from around 1780 bc recommends penal amputation of the hand for unsuccessful surgeons, an early example of 'shame and blame' audit cycle completion which must surely have led to some problems in workforce planning for the Babylonian health system. In the main, however, the recognised low probability of success meant that the reliability of most medical treatments was not seriously scrutinised for centuries, as it was considered impossible to distinguish failure through imperfect or incorrect treatment from failure because of the natural course of disease and injury and the divinely ordained nature of the world. This is reflected in the very similar attitudes taken to disease, death and the efforts of physicians and healers in writings from diverse cultures from China to ancient Rome to the early Muslim empire.^{1,2} The gradual development of science in Europe after the Reformation led, over the course of four centuries, to the beginnings of a different culture in which the effects of treatments could be explained according to theories which corresponded with verifiable facts, and the understanding of human physiology and pathology increased dramatically. The effects of treatments began to be more predictable and, in the case of surgery after the discovery of antiseptics and anaesthesia, much improved. As this change came about, it began to be more obvious that some clinicians had better results than others and the idea of medicine as a profession became prominent, with the medical degree marking out, from less trustworthy competitors, those whose skill and judgement could be trusted because of their recognised training and adherence to a common set of guiding principles. The improvement in efficacy brought about by more scientific treatment therefore had the effect of putting more pressure on the physician to behave according to a set of accepted norms decided by his (and at this stage it was always his) peers, but there was still a wide margin of leeway for variation in outcomes.

In the late nineteenth century and throughout the twentieth, the ever-increasing industrialisation of the process of medical treatment in hospitals, together with continuing advances in the effectiveness of medicine and surgery, led to wider understanding among the profession, and eventually among the public, of the wide variation in success rates that existed between centres for the treatment of the same condition. As treatments became more invasive, potentially toxic and complex, evidence of deaths and harm due to treatment began to emerge, but, more importantly, public attitudes to these occurrences changed. Having become accustomed to the expectation that medical treatment could achieve something in the majority of cases, patients began to question whether or not all adverse effects of treatment were inevitable. Professional concerns of this kind were raised in the nineteenth century by Semmelweis and in the early twentieth century by Codman,³ but the savage treatment of both of these pioneers reflects the complete rejection of critical self-examination which characterised the medical profession throughout much of its existence. Modern medicine was hailed in the second half of the twentieth century (with much

justification) as one of the great achievements of the scientific revolution. Its failings and costs were generally regarded within the profession as both acceptable and unavoidable but critical voices now challenged this perception, beginning with luminaries such as Ivan Illich⁴ and Archie Cochrane.⁵ The inexorable rise of lawsuits against doctors in the USA was a stimulus to the first systematic research on patient harm caused by treatment, and this showed from studies of case records that treatment caused specific identifiable harm in about 3.5% of patients.⁶ The US Institute of Medicine produced a landmark report⁶ in 1999, which extrapolated from such figures to calculate that treatment was killing 44,000–98,000 Americans every year and called for urgent reform.⁷ Since then a number of studies in a range of countries and health systems have attempted to quantify the problem. A reliable, objective and verifiable standard definition of treatment-related harm has not yet emerged, so that rates of harm in different studies cannot be directly compared. Most studies have been conducted in hospitals and reported rates of harm are considerably higher than the initial US studies, running between 8% and 14%.^{7,8} It has therefore been established that treatment-related harm and avoidable failure of effective treatment delivery are major causes of cost, morbidity and mortality in modern health care.

Analysis of causes: human factors

Beginning shortly after the first attempts to accurately describe the size and severity of the problem of patient harm, a literature has developed around analysis of the causes. This has been informed from its inception by earlier work done in other fields in which complex human work systems are at risk of catastrophic accidents. The seminal work on industrial accidents by James Reason⁹ formed the intellectual foundations for most theories of health-care safety. Reason emphasised the multifactorial nature of accidents and the complexity of their causes preventing prediction. He provided taxonomy for the types of errors humans make in work situations and developed the famous ‘Swiss cheese’ model to illustrate how barriers to harm need to be multiple and well maintained to reduce risk. Studies from the aviation industry were also very influential. The success of US aircraft carrier deck teams in performing difficult and dangerous tasks under extreme pressure was studied for clues as to how to develop ‘high-reliability organisations’.¹⁰ The use of mnemonics, prompts and checklists, as well as other formal methods of structuring communication among aircrew was studied, inspiring the well-known World Health Organization (WHO)’s surgical safety checklist¹¹ among other ideas for improvement. The work of Helmreich¹² on aircrew communication and relationships and their association with the risk of accidents was another pillar of the emerging theory of patient safety. His work showed the importance of clear communication in sharing mental models and ensuring high levels of situational awareness. His work on ‘power distance’ and the authority gradient in cockpits, and the adverse effect this has on the communication necessary for safe operation, was something that rang a chord with investigators studying the equally hierarchical world of medicine. The training courses adopted by airlines in response to Helmreich’s findings, generally known as crew resource management (CRM), were extensively adopted and tested in health-care settings, on the assumption that the underlying causes of error and harm were analogous. Another school of thought emphasised the role of inappropriate or poorly defined systems of work in ‘setting up’ clinical staff to fail^{9,13} and, therefore, advocated systems analysis and improvement using industrial techniques such as the Toyota production system.

Analytical studies of samples of errors and accidents leading to patient harm provided support for both schools of thought. Observational studies of operating theatre teams provided abundant evidence of communication and teamwork problems, many of which did seem very similar to those observed in aviation.^{12,14–16} Analyses of very serious incidents, such as those resulting in the inadvertent administration of intrathecal vincristine to cancer patients, showed how systems of work led ineluctably to a situation in which there was a high risk of a fatal error.¹⁴ The literature from both schools was, however, agreed on the need to de-emphasise the responsibility of the individual health-care professional. Rene Amalberti,¹⁵ one of the most creative thinkers on patient safety, described medical professionalism as one of the biggest barriers to safe health care. This apparent paradox is explained by the adverse results of the focus on the individual, which the medical model of professionalism promotes so strongly. The professional

ethos of medicine sets an unattainable standard of perfection, wherein the good clinician never forgets, never omits, never errs in judgement, is never too tired or emotionally upset to function and constantly updates their knowledge so that their expertise is always adequate to deal with the problems they face. When harm has occurred and an individual clinician has fallen below this standard in any way, the typical response is to attribute the harm entirely to the breach of professional standards. This results in high levels of guilt among individuals, but also encourages hypocrisy, dissimulation and attempts to shift responsibility onto others rather than objective analysis of what happened. These attempts are far from irrational because this approach to professional behaviour has led to a culture of blame in which it is generally accepted that all errors must be someone's responsibility, and that there is a moral necessity to punish that person for their lapse so that professional standards will be protected. Psychological and ergonomic studies have of course demonstrated beyond doubt the error-prone nature of even the most attentive and dedicated human professional and, therefore, punishment for error is unlikely to decrease the chance of its repetition, while the fear of it distorts communication and co-operation between workers anxious to avoid the possibility of blame. In actual practice, the use of the professional 'blame culture' is strictly related to hierarchy, so that the same error by an eminent and respected senior clinician and a new recruit from another institution are dealt with completely differently. The possibility of future error is not, however, decreased in either case.

Focus on surgery: specific problems

The operating theatre has attracted more attention than any other part of health care when it comes to analysing hazards and proposing solutions. It is difficult to quantify the amount of patient harm that can be attributed to errors in theatre, but it has been suggested that it is the highest-risk environment for hospital patients. These data reinforce the common-sense argument that the site where clinicians deliberately invade the body, risking damage to vital structures and the introduction of infection, is bound to be one in which serious inadvertent harm may occur. There is a significant literature around the observation of teamwork and communication in the operating theatre and its relationship with patient outcome. The negative effects of hierarchy and inter-professional 'tribal' barriers on communication and co-operation have been described and discussed, and evidence has been produced for a relationship between technical error rates and the quality of team interactions. Several groups have developed systems for evaluation of the quality of non-technical skills and teamwork behaviours in operating theatre personnel, as either individuals or teams. Some of these have been integrated into training systems designed to improve team outcomes by enhancing team interactions.

However, some surgeons argue that the operating theatre may in fact be safer than other parts of the surgical patient pathway, as the operation is generally performed by senior team members with a high degree of skill, assisted by a group of appropriately skilled colleagues, whereas post-operative care can sometimes be organised so that very junior staff members are faced with crisis situations outside normal hours, which they are not equipped to deal with. This view has been lent some support by recent work on 'failure to rescue'. Several studies have highlighted a recurring theme in reports of death and serious harm after surgery, which is the frequency with which deterioration is recognised but left untreated for long periods of time. A number of classic behaviour patterns appear to be responsible for this, including diffusion of responsibility and reluctance to send bad news up the chain of command in hierarchical organisations. Recent work has shown that high-volume units with excellent outcomes do not have substantially lower rates of serious complications than lower-performing lower-volume units, but they deal with these situations more effectively and, therefore, more frequently avoid death or other serious sequels. There are no reliable data on whether preoperative or intraoperative factors are more influential in generating adverse outcomes and, given the intimate inter-relationship between the two, attempts to assign relative risk to either are probably futile. Both the study of error and process breakdown and the efforts to minimise it necessarily require quite different approaches in the two environments. Operations require strongly co-ordinated efforts over a limited period of time from a disparate multidisciplinary team (MDT) using highly specialised and sophisticated equipment. Postoperative care takes place continuously

over a much longer period and involves a larger group of staff working in shifts, making transfer of information a key element in achieving success. The bulk of the routine work falls on nursing staff and junior doctors and, therefore, the willingness and ability of these staff groups to access advice and support from more senior medical staff or specialists outside the immediate team is another critical determinant of performance.

Attempts to correct problems

Serious efforts to identify the underlying causes of accidents and errors leading to patient harm in medicine lagged several decades behind similar efforts in some other areas of work. This delay can perhaps be attributed in varying degrees to the effects of the strong 'person-based' professional ethos in medicine which inhibited thinking about wider systems issues as causes of ineffectiveness, and partly to the fact that individual incidents in medicine normally affect a single patient. They are therefore less likely to create widespread concern, media attention and threats to the survival of corporate entities than, for example, the large-scale disasters represented by aeroplane crashes or industrial accidents in the power-generation industry. Whatever the reasons, the fact that medicine came to this field of work relatively late meant that there was already an established body of theory and practice available, some of which has been described earlier (see *Chapter 1, Analysis of causes: human factors*). Medicine undoubtedly benefited from the availability of a paradigm within which ideas about causes, effects and solutions for medical error could be developed, but the transmigration of ideas and principles from very different fields of work into the medical arena also carries serious risks of inappropriate extrapolation. What holds true in civil aviation, military operations or nuclear power generation may not be equally true in the very different context of medical care. Expert practitioners from other fields of work have repeatedly emphasised how disconcerting they found the hospital environment when asked to give help or advice within it. The features that characterise hospital medicine, as against practice in other fields in which safety work is more advanced, are the complexity of the process of patient care and the distributed nature of responsibility for action. The professional model in health care was developed from a simpler one in which treatment was directed by professional doctors with a high degree of specialist knowledge or training, either at university or within a guild or craft society. They were assisted by (almost invariably female) carers whose focus was on patient comfort and psychological well-being but who were charged with implementing the instructions of doctors for the performance of the key curative treatments. The separation of these roles has survived vast changes in the nature and capacity of treatments, the industrialisation of hospital medicine and ever-increasing specialisation but a host of other specialists have been superimposed on the basic model: physiotherapists, dietitians, pain specialists, pharmacists, infection specialists, wound care specialists and many others. As the process of care has become more complex and expensive, the role of professional business managers has become ever more prominent, leading to their eventual emergence as the leaders of the hospital organisation, responsible for employing and directing the doctors as well as all the other specialties. However, the persistence of professional roles and attitudes means that managers do not have the direct line management authority common in other industries over the activities of specialists within their area of responsibility. In order to achieve change in the work of their department, managers commonly need to negotiate with a hierarchy of doctors and nurses who have responsibility for policy within their specialty cadre for the same part of the hospital's work. The creation of true MDTs under unitary control is therefore impossible and progress can only be made by consensus. At a lower level, the traditional model of doctor responsibility for the care of the individual patient has been diffused by modern multidisciplinary care to the extent that it is often unclear who has the authority to make decisions about, for example, whether or not a patient should receive a particular antibiotic or nutritional supplementation technique. Change over time has also affected relationships between doctors caring for the same patient. A traditional model attributed responsibility for patient care to a single consultant, assisted by one or two junior doctors who followed his or her instructions. As the technological and pharmacological possibilities for more and more intensive care increased during the twentieth century, this model became unworkable, since the number and timing of the decisions required expanded to exhaust even the most dedicated professionals. Specialisation and shift work made the jobs of hospital doctors physically possible once again, but the cultural transformation required to fully implement a shift and specialty-based system have not occurred. Doctors continue to be reluctant to

hand over responsibility for patients they regard as 'theirs' and, conversely, do not always assume total responsibility for patients they accept as part of their case load if they regard them as 'belonging' to another specialist. This is a particular problem in surgery, as the fact of having operated on a patient creates an obligation in the mind of the specialist (and his or her colleague) to continue to take ultimate responsibility for the patient's care until the underlying problem can be regarded as resolved. In this complex environment, it is not surprising that models, tools and procedures devised to prevent error and enhance safety in very different, and usually simpler workplaces may not always be effective.

Models of harm in health care and their implications: critique

Although Reason's initial description of the multiple imperfect barriers to harm in complex systems, the so-called 'Swiss cheese' model,⁹ was a major conceptual breakthrough, it did not attempt to detail or classify either the potential sources of harm or the barriers which may prevent it. Given the complexity of the modern health-care environment, models which provided a structure for thinking about the opposing forces involved in increasing or protecting against risk to the patient were very necessary, and several attempts have been made to build these. Charles Vincent¹³ was among the earliest authors to attempt a comprehensive description of the hospital environment from the point of view of patient safety risk and possible sources of harm. The final iteration of this framework is commonly known as the 'London Protocol'.¹³ The protocol has been recommended by some authorities for use in 'root cause analysis' investigations of safety incidents in health care, but its greatest value is perhaps to ensure that the influences on harm and error which are more distant from any specific clinical episode are not forgotten in an overall assessment. Although it certainly achieves this, the difficulty in either defining or quantifying influences such as 'management culture' can be problematic. When it is unclear if an influence is important or not, and it can be neither defined nor measured, there is a strong tendency to ignore it and focus on easier parts of the problem. This is particularly likely if, as is usually the case, there is potential risk to any investigator in incorporating within their search the activities of individuals with power and influence in the organisation. For many clinical incidents the more proximate influences are indeed the ones which most urgently need fixing, but defining these very often leads remorselessly to questions about the governance structure that allowed these errors to occur.

Another popular framework for discussing risk and error in health care is the Systems Engineering Initiative for Patient Safety (SEIPS) model of Pascal Carayon *et al.*¹⁷ The five dimensions of SEIPS are tools/technology, person, tasks, organisation and environment. This format does seem able to encompass most or all of the common influences on risk. However, the categories of the five-dimension classification are not mutually exclusive: it is relatively common for a problem to be classifiable in more than one category and the dimension categories are not of a comparable nature. It would be helpful for those considering potential risks or analysing real incidents to be able to use a simple system that categorised sources of risk more conclusively, using a system which made it easy to decide how any particular factor should be described. We therefore developed the three-dimensional (3D) model specifically for the analysis of risk at the microsystem level, which is between health-care worker and patient.¹⁸ This simplified model essentially ignores the higher organisational influences and concentrates instead on factors that can be easily identified by visiting the workplace. The model postulates that all influences at this level can be described in terms of systems of work, workplace team culture and the technology used to complete the work. These three dimensions interact in unpredictable and bidirectional ways, potentially modifying the effects of external interventions directed at just one of the dimensions. By analogy with thermodynamics, we would predict that an imbalance represented by an increase or decrease in risk in one of the dimensions will tend to be countered by any interactions with another dimension that allow diffusion or diminution of that risk. So, for example, if teams receive better training in teamwork to improve their culture, but are forced to continue using an inefficient and risky work system, much of the benefit will 'bleed away' via this interaction between culture and system. The model is illustrated in *Figure 1*. Initial validation work on this model showed that it seemed informative in analysis and explanation of typical safety incidents in operating theatres. We therefore adopted it as a basis for thinking about safety problems in surgery at the microsystem level. This led us to develop a hypothesis that formed the basis for the work described in this report.

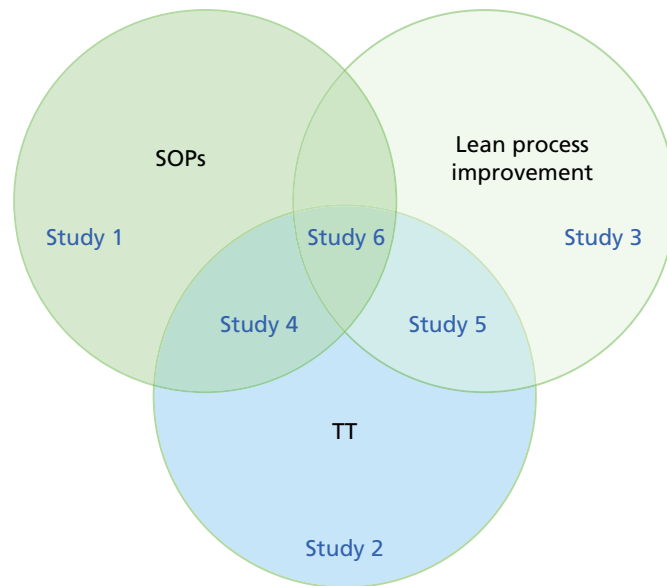


FIGURE 1 Schematic of studies. SOP, standard operating procedure; TT, teamwork training.

Chapter 2 Study design

Study design background

Having created the 3D model for safety in surgery,¹⁸ we spent some time thinking about whether or not it was literally true in all circumstances. We imagined three dimensions that were completely mutually exclusive, so that every influence on the risk of harm to patients could be described as an example of culture, system or technology, or a defined mixture of two or more of these. We soon realised that many influences were in fact mixtures, which made it all the more interesting to consider whether or not they could be modified by an intervention that affected one part of the mixture. We imagined the risk of patient harm as similar to a gas, which had a certain pressure in each of the globes we drew to represent a dimension. The higher the pressure, the more likely would be patient harm. The interactions between the dimensions, we had said, could be of any type. They might allow two-way flow of risk, or only one way or they might only affect the speed or direction of other interactions, for example, making the flow of risk intermittent instead of continuous or forcing an interaction that normally allowed flow for a limited time only to remain open all the time. What would be the real-life equivalents of these theoretical constructs? An example of an interaction between system and culture might be the reaction of staff to a new handover system. Owing to poor culture, the staff decide to corrupt the system so that it no longer fulfils its purpose but is easier for them. The high-risk concentration in the culture dimension has now passed into the system dimension via the interaction around the handover system, which increases the overall risk. We realised that this was a key point. If risk is merely exchanged between dimensions during interactions, overall risk remains constant. This is not a good description of reality. In order to model reality, therefore, the interactions between the dimensions have to be capable of creating or destroying more risk. This modification of the model allows us to make predictions about the effects of modifying two or more of the dimensions in the model system. Each interaction point between dimensions can now allow net risk to flow into or out of a dimension, but it could also create or destroy overall risk. If we attempt to lower the total risk in the system by concentrating only on the culture, for example, the interactions between culture and the other two dimensions may affect our likelihood of success. If the net effect of all the interactions is to allow net flow of risk into the culture domain, or to create more total risk, our efforts are likely to be frustrated. It is likely that this is in fact the case in any real-life system in which we are trying to reduce risk, as if the net effect of interactions were already in the direction of reducing risk, the need for our activity would not have arisen. If we now consider an intervention that reduces overall risk in two or more dimensions, the potential for synergy becomes obvious. Not only is the overall risk being reduced over a larger percentage of the overall organisation but the malign interactions between the dimensions are being dismantled from two directions, destroying the possibility of a homeostatic correction of the risk pressure in either of the dimensions affected via interaction with the other.

It is possible that culture-focused interventions are generally more or less effective than systems-focused interventions, but it is also possible that there are important elements of context which would make it very difficult to extrapolate from a comparison of the two in any particular setting. We felt that collecting data about the effectiveness of different interventions in settings that were as similar as possible might help us draw some tentative conclusions about this, but that a systematic literature review would probably be a more appropriate way of trying to reach a general conclusion. We therefore rejected the idea of focusing on whether systems or culture interventions were better as our primary research question.

Our hypotheses therefore are that interventions addressing two or more dimensions of the three dimensions of patient risk are more likely to be successful than those which address only one. Our analysis of the current literature shows that most interventions have focused heavily on one dimension. Training sessions held with staff at which they consider their relationships with other staff, for example, are clearly about culture, although a small-system aspect may creep in via, for example, discussion of formal systems

and routines for communications such as situation, background, assessment, recommendation. Methods for improving the reliability of systems by data-driven analysis and measurement, on the other hand, have little direct impact on culture but address systems problems. Integrated approaches may have greater potential than either type of intervention alone. This was the underlying rationale for our programme. For purely practical reasons we decided to focus on system and culture and not to work on technology. An overhaul of technology in one small area of a hospital is not a practical proposition for organisational reasons, and the decisions needed to bring about a radical reform of a hospital's use of technology would require very large resources and a long period of time to implement. On the other hand, we could conduct small-scale studies of culture and system interventions without major disruption to overall hospital function or budgets. This omission leaves the theoretical construct partly tested at best and further studies to evaluate the impact of technology rationalisation with or without attention to the other dimensions in the future would be highly desirable. We decided to focus our efforts on the two dimensions over which we felt that we could exert some control within a study setting with present resources. We were influenced in doing so by our recognition that the current literature on safety interventions has also focused largely on these two dimensions, but surprisingly has largely neglected the potential for interaction between approaches dealing with them.

The primary focus of the study is patient safety and quality of outcome in surgery. The main questions relate to the effectiveness of methods to improve safety and quality. The selected primary outcome measures are discussed in *Chapter 3*. The three intervention approaches tested were culture enhancement [teamwork training (TT)] and two contrasting approaches to systems improvement [lean systems and a systems approach based on development of standard operating procedures (SOPs) through application of ergonomic methods].

Study design and rationale

We reviewed the literature looking particularly at the detail of interventions used to reduce risk of harm to patients or to increase the reliability with which surgical processes were carried out. We identified TT methods based on CRM as a common type of intervention that seemed to be focused largely on improving workplace culture. We identified a range of methods for improving workplace systems and considered the possibility that they might interact differently with a culture change intervention. We therefore chose two different systems' improvement approaches to study because they were clearly quite distinct in several important respects.

We were acutely aware of the practical constraints on research interventions to improve safety identified in previous research: intervention studies need to be carried out at a certain scale in a real-life environment. Working with an individual clinician, for example, would not be appropriate. The smallest unit feasible without moving to an artificial controlled environment such as a simulator is the 'microsystem' of a single ward or operating theatre. With a realistic set of resources, we had a choice of conducting a small number of studies at a scale larger than this, or a larger number at this scale. It seemed clear that testing our hypothesis would require several parallel studies examining different possibilities, so we opted for a number of small studies rather than one or two larger ones.

This approach also allowed us to compare the effects of different interventions indirectly, using a comprehensive pooled analysis of all of the results in five identical studies. The interventions used in these studies were TT alone, SOPs alone, lean alone, TT and SOPs and TT and lean.

Another important consideration was the need for controlled studies. From previous work we were aware that the live hospital environment is a highly unstable one in which influences over which the experimenter has no control arise regularly and which can grossly distort outcomes of interest. Recent publication of two large-scale studies had shown the importance of controlling for such influences even in large-scale experiments.^{19,20} In both studies the control group revealed the existence of a secular trend to improved

outcomes, which would have been wrongly attributed to the intervention if the studies had not been controlled. We therefore considered that a control group was essential for each study. The complications involved in setting up studies with two interventions and a control group are significant in any clinical study, but in studies for which the interventions being compared are training programmes, which require long-term staff co-operation and which need to be set up so as to avoid contamination of control groups, they become daunting. We therefore took a different approach. Rather than attempt to compare interventions directly we decided to carry out a series of controlled studies of single interventions and of the combinations of each systems intervention integrated with the culture intervention. The controlled studies permit analysis of the difference in change after the intervention between the active study group, which received it, and the control group, which did not.

We therefore concluded we could assess the interactions with six individually controlled studies, illustrated in *Figure 1*.

We also expected that we would find out which aspects of each of the quality improvement (QI) approaches were useful and that we would learn a good deal about what was effective during the process of conducting our five studies; therefore, we were interested in creating a final intervention study in which we integrated TT with aspects of both lean and the SOP approach. We chose to extend the range of the clinical context for this final study to encompass not only the operating theatre but also the entire patient journey from admission to discharge. This was the sixth and final intervention project in the programme. The aim was to pilot test an integrated intervention strategy based on the learning from the previous five studies, with a view to conducting future controlled studies at a larger scale.

We recognised the need to develop both a pre-planned statistical approach to analysis of these studies and a health economic analysis of the costs and effects of the interventions. It was clear that there would be a considerable amount of practical learning on an experiential basis during the study, particularly on what proved successful and what did not, and about the challenges and barriers to delivery. We therefore ensured that we had a qualitative study plan which would allow us to analyse the underlying causes of success and failure, barriers and delays, and to point to potential ways of avoiding these problems in future interventions. We also recognised the importance of knowledge translation: being able to explain and disseminate the essential learning from the studies in an effective way that would result in adoption of the valuable parts of our work by others, with actual benefit for NHS patients.

Study logistics

The research team aimed to study 100+ operations in each individual study, maintaining a 1 : 1 active-to-control ratio when possible. This would result in 25 active and 25 control pre-intervention operations, and 25 active and 25 control post-intervention operations. The control cases used were to be parallel teams, when available, and were not subjected to any intervention.

The interventional studies all follow a controlled interrupted time series design, incorporating a 3-month pre-intervention phase, an intervention phase (usually also 3 months) and a post-intervention phase (at least 3 months) (*Table 1*).

TABLE 1 Schematic of intervention phases

Intervention	Phase		
	Pre intervention (3 months)	Interventional (3 months)	Post intervention (\geq 3 months)
Active	25 cases		25 cases
Control	25 cases		25 cases

Process data collection was to be completed through observation of the 25 cases within each section. Clinical outcome data were to be collected from the patients undergoing treatment, with the surgical teams involved in the study arm within the 3-month phases. For example, during the 3-month pre-intervention phase, all patients of surgeons participating in the intervention would have their clinical outcome data collected (identified only by time frame and operating surgeon).

The duration of the studies ranged from 6 months to 20 months. Approximately 600 staff subjects were observed over five sites in three NHS trusts.

Study settings

To examine the hypothesis that combined interventions functioned better than single-dimension interventions, we needed to conduct studies of both types of interventions while keeping the factors that might introduce bias and error to a minimum. We therefore decided to identify a hospital environment that was as standardised as possible. This meant finding a surgical specialty that carried out the same types of operation in more or less the same way at high volume in many different hospitals. We rejected studying minor procedures such as endoscopy, as both the level of complexity and the degree of risk in such procedures are much lower than those in the environment in which we were actually interested, which is the surgical operating theatre. Emergency general surgery, on the other hand, carries a high risk of harm to the patient and the variability of the procedures performed is potentially large. These factors meant that identifying and studying a stable team in a stable situation was unlikely.

We therefore chose to target a small selection of elective procedures from elective adult orthopaedic surgery, which in the UK is very much focused on a small repertoire of operations performed at high volume: hip and knee replacement, arthroscopic surgery and cruciate ligament repairs. In the latter stages of the study, it became necessary to either expand beyond elective orthopaedic operations or extend beyond the original hospital trusts involved. We chose to extend the operation types, as the negotiations associated with changing trusts were logistically challenging. The additional types of surgery included were elective plastic reconstructive surgery, orthopaedic trauma surgery and vascular surgery.

These operations comprise, at least in adult practice, a small repertoire of relatively complex procedures that are repeated at high volume in many centres across the UK. This allows for a choice in setting of study (from university teaching hospital to district general hospital; *Table 2*). The large number of these procedures (mainly hip and knee replacements) also makes them highly relevant to NHS priorities and any improvement in outcomes would be of major importance to the NHS. We sought the opportunity of studying our methods in three quite distinct contexts and, therefore, approached three trusts that carry out this kind of work within travelling distance of our Oxford base. The Nuffield Orthopaedic Centre in Oxford is a free-standing specialist hospital that focuses on elective orthopaedics and plastics. The University Hospitals Coventry and Warwickshire NHS Trust is a large university hospital with a high orthopaedic workload. Part of this is dealt with at the main site, a very large operating theatre suite with more than 30 theatres. The remainder of surgeries are performed at Hospital of St Cross (Rugby), several miles from the main site; this is a small elective unit specialising in high-volume surgical procedures. Kettering General Hospital is a small semirural district hospital that has developed a busy general orthopaedic unit but has very little academic or training input from outside.

For the final study, utilising the learning from the suite of five intervention studies, we chose to work with the neurosurgery department at the Oxford University Hospitals NHS Foundation Trust. By this time we felt that further interventional work with the orthopaedic surgeons in the trusts we had worked with was impractical, as the staff were thoroughly familiar with the team, and most had been exposed to one or another study in the past. Neurosurgery represented a relatively high-risk discrete specialty, with a unique patient pathway that we could study from beginning to end.

TABLE 2 Description of Trust site and surgical interactions

Hospital Trust	Hospital site	Description	Surgical specialties partaking in study
Oxford University Hospitals NHS Trust	Nuffield Orthopaedic Centre	Specialist centre	Orthopaedics and plastic
	John Radcliffe Hospital	University teaching hospital	Neurosurgery and vascular
University Hospitals Coventry and Warwickshire NHS Trust	University Hospital Coventry	University teaching hospital	Elective and trauma orthopaedics
	Hospital of St Cross (Rugby)	District general hospital	Elective orthopaedics
Kettering General Hospital NHS Foundation Trust	Kettering General Hospital	District general hospital	Elective orthopaedics and vascular

Initial liaison and manner of working with trusts

We discussed the programme with senior management at all three trusts at an early stage. A senior clinician or manager acted as liaison on each site. With their help we identified the consultants and associated theatre teams and who would be best suited to involvement in the study, and arranged to hold information and engagement meetings with the surgical teams from the surgeries involved. For logistic reasons some of these meetings involved only the surgical consultants and others only theatre staff, but all groups were engaged and had the programme explained and discussed with them. We appointed four research fellows of whom two worked largely in Oxford and the other two were managed from Warwick and worked largely there and at Kettering. Dr Catchpole and, latterly, Dr Morgan were responsible for ensuring day-to-day communication and co-ordination with the clinical staff and management at the hospitals involved, together with the programme manager.

Study team and expertise

The initial team of investigators was assembled with the needs of the studies we had planned very much in mind. The co-investigators each brought with them expertise that contributed an important element to the team.

Mr Peter McCulloch (PM; Principal Investigator) set up the Quality Safety Reliability and Teamwork Unit in Oxford in 2005 specifically to conduct scientific studies of interventions to improve surgical practice and patient safety. His expertise in study design and methodology together with his experience as a consultant surgeon helped him to design studies appropriately.

Dr Ken Catchpole (KC) brought experience and knowledge in the application of ergonomics [human factors (HFs)] to health care, and to surgery in particular. Dr Catchpole joined the team after a successful 3-year collaboration with the paediatric cardiac surgery team at Great Ormond Street Hospital, during which he devised a number of methods for analysing their work in theatre and demonstrated important principles, such as a demonstrable link between accumulation of small errors and the likelihood of a major error, and between the non-technical and technical performance of a theatre team.

Dr Steve New (Lecturer in Operations Management at the Saïd Business School) is an expert on process improvement methods, particularly lean methodology, and has considerable experience of real-life improvement projects using these methods.

Professor Doug Altman (Head of the Centre for Statistics in Medicine at Oxford) provided advice on the development of the statistical plan and model.

Dr Gary Collins (Senior Statistician at the Centre for Statistics in Medicine) was in operational charge of the analysis plan and delivery of the statistical analyses.

Professor Alastair Gray (Professor of Health Economics at Oxford) devised a plan for collecting both costs and outcomes to allow the analysis of the cost-effectiveness of the interventions we instituted. This work was then implemented and subsequently led during the evolution of the study by Dr Oliver Rivero-Arias.

Professor Crispin Jenkinson was brought into the programme to assist with the study of patient-reported outcome measures (PROMs), on which he is a recognised expert, as part of the overall health economic analysis.

Professors Renee Lyons and Alison Kitson contributed to the development of a knowledge translation plan for qualitative analysis, explanation and dissemination of learning from the programme.

Professor Damian Griffin (Professor of Orthopaedic Surgery at Warwick University) acted as liaison and project manager for the practical studies conducted at that site, and supervised the work of the research staff involved.

Dr Tony Berendt, Dr Karen Barker and Ms Elaine Strachan-Hall acted as liaison and facilitators for the study at the Nuffield Orthopaedic Centre site.

Dr David acted as liaison and facilitator for the study at the Kettering General Hospital site.

Professor Graham Martin (GM) was brought into the team towards the end of the project to assist with revision of the plans for qualitative analysis. His expertise in working with Professor Mary Dixon-Woods on analysis of health-care improvement interventions proved invaluable in ensuring delivery of this important part of the programme.

Dr Lauren Morgan (LM) joined the team in 2011 as a replacement for Dr Catchpole when the latter accepted a post in California, USA. As well as HF's expertise, she brought to the team organisational and negotiating skills that proved highly valuable in implementing the study programme.

The investigators who carried out the bulk of the data collection and intervention delivery were Dr Lauren Morgan (above), Dr Mohammed Hadi (MH) and Dr Eleanor Robertson, Ms Sharon Pickering (SP) and, latterly, Ms Lorna Flynn (LF), Ms Laura Bleakley (LB) and Ms Julia Matthews (JM).

Final programme

The final programme of studies contained a number of modifications of the original plan, some for logistic reasons and others because discussions between the members of the research team led to an evolution in ideas on how best to answer the main questions we set out to address. Our final hypotheses were:

Primary hypotheses

Single interventions

Teamwork training: CRM-based training given to a theatre team improves teamwork performance and process error rates.

Standard operating procedures: SOPs developed and applied by a theatre team improve teamwork performance and process error rates.

Lean process improvement: lean process improvement developed and applied by a theatre team improves teamwork performance and process errors rates.

Integrated interventions

Teamwork training and SOPs: TT and SOPs together result in greater improvements to teamwork performance and process error rates than either intervention alone.

Teamwork training and lean process improvement: TT and lean process improvement together result in greater improvements to teamwork performance and process error rates than either intervention alone.

Secondary hypotheses

The measures being collected include a set of clinical outcomes or surrogates (e.g. length of hospital stay, returns to theatre, 30-day mortality), a set of observational process measures (e.g. operative durations, WHO checklist completion) and a set of PROMs [e.g. European Quality of Life-5 Dimensions (EQ-5D)]. With each of the primary hypotheses, there are three associated sets of secondary hypotheses:

Clinical outcomes

Intervention (lean, TT or SOPs, or a combination of these) results in improvements in clinical outcomes compared with current practice or a single intervention for pairs.

Observational outcomes

Intervention (lean, TT or SOPs, or a combination of two of these) results in improved theatre process efficiency compared with current practice or a single intervention for pairs.

Patient-reported outcome measures

Intervention (lean, TT or SOPs, or a combination of two of these) results in more improved patient-reported outcomes than current practice or a single intervention for pairs.

Safer delivery of surgical services workstreams

Our initial proposal to the National Institute for Health Research (NIHR) called for four workstreams: (1) a preparatory stream involving literature review, development of statistical, knowledge translation, health economic models and project planning; (2) a stream focused on the development of more advanced and appropriate measurement techniques for team performance; (3) a stream focused on the experimental implementation and evaluation of improved intervention methods; and (4) an analytical stream comprising statistical, health economic and knowledge translation analyses and dissemination of learning. The detailed planning of the third workstream became the strongest focus of the programme, as detailed and explained above (see *Chapter 2, Study design and rationale*). The second workstream was modified partly in response to the needs of the third and partly for practical reasons.

A number of pilot attempts, led by Dr Catchpole, to develop a simplified universal measure for team non-technical skills which could be used in any part of the health-care systems, ended in failure and re-evaluation of our goals. We concluded that this part of the task we had set ourselves was infeasible and focused instead on improving our existing Oxford Non-Technical Skills (NOTECHS) scale for theatre team non-technical skills. At the same time it became evident that we needed to revise our methodology for evaluating the technical performance of theatre teams. An important part of workstream 2 therefore became the development of validated measures for evaluating technical performance (the glitch count) and for measuring compliance with the mandatory procedures of the WHO's surgical checklist. Our final study design strengthened considerably the potential for interpretation of our five parallel studies. Our final qualitative assessment programme was, we feel, also an improvement on the initial plan, allowing the incorporation of learning from an extremely successful specialist team with expertise in this area. Our health economic analysis, on the other hand, was significantly affected by the problems which arose over administration of PROMs and extraction of clinical details from individual patient records. Data protection regulations proved more challenging than we had appreciated and it proved impossible to collect all the data we wanted. One final difference between the proposal and the plan as delivered was the element of delay that developed during the study. This was partly because of the difficulties of negotiating training time for clinical staff with hospital management, which proved much more

challenging than we had expected, and partly because of protracted initial negotiations over the contract, which resulted in an unrealistic start date for the programme. This means that we are unable, at this stage, to provide a complete picture of the final intervention on neurosurgery, as some studies are still in the process of analysis. Despite these changes, the final body of work is substantial, coherent and successful in integrating mixed methods in the analysis of the key question underlying the programme: how can we most effectively combine methods to improve safety and quality in surgery? The results carry important messages which should influence thinking on the design of such interventions in the future.

Chapter 3 Methods

Parts of the text in this chapter have been reproduced from Robertson *et al.*²¹ © 2014 Robertson *et al.* This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited; and from Morgan *et al.*²² This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>. Furthermore, text has also been reproduced with permission from Pickering SP, Robertson ER, Griffin D, Hadi M, Morgan LJ, Catchpole KC, *et al.* Compliance and use of the World Health Organization checklist in UK operating theatres. *British Journal of Surgery*, © 2013 British Journal of Surgery Society Ltd. Published by John Wiley & Sons Ltd;²³ and Flynn L, McCulloch P, Morgan L, Robertson E, New S, Stedman F, *et al.* The Safer Delivery of Surgical Services Programme (S3): explaining its differential effectiveness and exploring implications for improving quality in complex systems. *Annals of Surgery* vol. 264, iss. 6, pp. 997–1003.²⁴

Observational methods

Methodological approach

The observational methods are used to collect real-time prospective data in the clinical setting. The benefits of this approach, along with the requirements to achieve robustness, are described in detail by Carthey,²⁵ who includes the requirement for training of, and demonstration of reliability between, observers. Structured observation is chosen over unstructured ethnographical approaches because of the requirement to demonstrate reliability and the requirement to translate observational data into quantitative data to facilitate the comparison required for the evaluation.

Observer background and training

The clinical observers were either surgical trainees (ER and MH) or operating department practitioners (JM) with greater than 1 year of theatre experience. The HF specialists had at least an undergraduate qualification in HFs (LM, SP and LB). The clinical observers gained experience of HF principles from in-house lectures and literature reviews and the HF observers gained experience of the theatre environment from theatre observational practice and mentoring by clinical observers. All observers trained in the use of the observational methods over a 2-month period with self-study and group practice sessions using video-recordings of operating teams in simulated settings.

The clinical observers developed a process map of the main operation types to be observed, which took the form of a descriptive list of the operative process, including relevant procedures and steps. These process maps formed the basis for the training and subsequent structured observation.

Non-technical skills

Non-technical skills assessment tools adapted from aviation have been used in the operating theatre to understand the influence of behaviour on outcome. A number of methods have been developed for assessing teamwork skills in operating theatres, based on direct observation or video analysis.^{26–28} The development of the first NOTECHS assessment method was conducted via a European collaboration,²⁹ which has since been followed by further developments. The Observational Teamwork Assessment for Surgery,³⁰ Oxford NOTECHS scale,²⁶ Objective Structured Assessment of Technical Skills³¹ and Effective Practice and Organisation of Care (EPOC)³² provide whole-team assessments; Anaesthetists' Non-Technical

Skills,³³ Non-Technical Skills for Surgeons³⁴ and Scrub Practitioners' List of Intraoperative Non-Technical Skills³⁵ focus on subteam performance.

As the complexity of surgical teamworking becomes clearer, so the demands for more sophisticated methods of measurement have followed. The validity and reliability of the Oxford NOTECHS system was demonstrated in live theatre environments,²⁶ but a number of imperfections were noted in discussion and through experience. The system tended to group results close to the median and, therefore, had a suboptimal capacity to discriminate between teams whose performance was not extreme. We therefore sought to develop a modified version of the non-technical skills assessment method, which we termed Oxford NOTECHS II, and which is of greater utility than the original.

Development of Oxford Non-Technical Skills II scale

The Oxford NOTECHS II scale is built on extensive work in developing, evaluating and validating the original Oxford NOTECHS.²⁶ The final version was developed through group discussion between members of the research team. The four original Oxford NOTECHS domains of leadership and management, teamwork and co-operation, problem-solving and decision-making and situation awareness remain unchanged. The behavioural markers for each of the Oxford NOTECHS II parameters are largely unchanged from before (*Table 3*). There is no alteration to the consideration of the theatre team as three subteams: surgical (operating and assisting surgeons); anaesthetic (anaesthetists and anaesthetic nurses/practitioners) and nursing (scrub and non-anaesthetic circulating nurses and practitioners).

Oxford NOTECHS II differs from the original Oxford NOTECHS scale in that it uses an 8-point rather than a 4-point scale for dimensions, and it assigns all teams a baseline score of 6, a behavioural marker of 'consistently maintaining an effective level of patient safety and teamwork' (*Table 4*), with subsequent observations of behavioural markers potentially resulting in deviation upwards or downwards.²⁶

TABLE 3 The operating theatre team Oxford NOTECHS assessment tool

Leadership and management	
Leadership	Involves/reflects on suggestions/visible/accessible/inspires/motivates/coaches
Maintenance of standards	Subscribes to standards/monitors compliance to standards/intervenes if deviation/deviates with team approval/demonstrates desire to achieve high standards
Planning and preparation	Team participation in planning/plan is shared/understanding confirmed/projects/changes in consultation
Workload management	Distributes tasks/monitors/reviews/tasks are prioritised/allots adequate time/responds to stress
Authority and assertiveness	Advocates position/values team input/takes control/persistent/appropriate assertiveness
Teamwork and co-operation	
Team building/maintaining	Relaxed/supportive/open/inclusive/polite/friendly/use of humour/does not compete
Support of others	Helps others/offers assistance/gives feedback
Understanding team needs	Listens to others/recognises ability of team/condition of others considered/gives personal feedback
Conflict-solving	Keeps calm in conflicts/suggests conflict solutions/concentrates on what is right
Situation awareness	
Notice	Considers all team elements/asks for or shares information/aware of available of resources/encourages vigilance/checks and reports changes in team/requests reports/updates
Understand	Knows capabilities/cross checks above/shares mental models/speaks up when unsure/updates other team members/discusses team constraints
Think ahead	Identifies future problems/discusses contingencies/anticipates requirements

Source: reproduced from Robertson *et al.*²¹ © 2014 Robertson *et al.* This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

TABLE 4 The scoring scheme for NOTECHS II

Behaviour	Frequency	Oxford NOTECHS II scale score
Compromises patient safety and effective teamwork	Consistently	1
	Inconsistently	2
Could directly compromise patient safety and effective teamwork	Consistently	3
	Inconsistently	4
Maintains an effective level of patient safety and teamwork	Inconsistently	5
	Consistently	6
Enhances patient safety and effective teamwork	Inconsistently	7
	Consistently	8

Source: reproduced from Robertson *et al.*²¹ © 2014 Robertson *et al.* This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Validation

An initial live test of inter-rater reliability in 20 elective orthopaedic operations across multiple sites showed good inter-rater agreement. Subsequent analysis on the whole data set confirmed the maintenance of good inter-rater agreement for total Oxford NOTECHS II scale score between the HF and clinical observers (Figure 2 and Table 5).

The Oxford NOTECHS II scale scores correlated with the quality of completion of the WHO's surgical safety checklist. We found a weak correlation between NOTECHS II scale score and glitch rate.

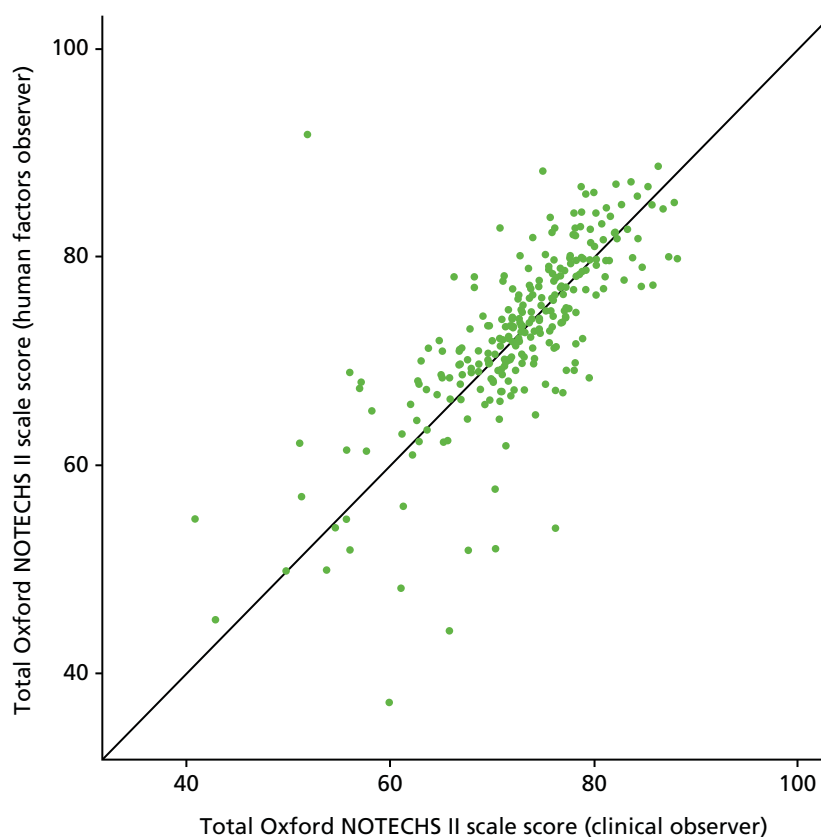
**FIGURE 2** Oxford NOTECHS II scale score: HF observer and clinical observer.

TABLE 5 Inter-rater agreement

Domain	Team (% agreement, intraclass correlation coefficient)		
	Surgical	Nursing	Anaesthetic
Leadership and management	59, 0.881	59, 0.777	64, 0.739
Teamwork and co-operation	55, 0.757	45, 0.343	64, 0.676
Problem-solving and decision-making	55, 0.725	63, 0.397	78, 0.385
Situational awareness	53, 0.770	48, 0.683	56, 0.675

Source: reproduced from Robertson *et al.*²¹ © 2014 Robertson *et al.* This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Method for Oxford Non-Technical Skills II scale

The three theatre subteams are each scored on four parameters (leadership and management, teamwork and co-operation, problem-solving and decision-making, and situation awareness) resulting in a total of 12 scores. The final scores were calculated at the end of the operation (during suturing/applying dressings) and entered into procedure-specific data collection books in an independent manner.³⁶ Once the operation was complete, each observer's scores were individually entered on a secure database.

The record sheet for scoring the Oxford NOTECHS II scale is shown in *Table 6*.

The total NOTECHS II scale score for each operation was used for the evaluation of the operating team's non-technical skills performance.

TABLE 6 Example data entry sheet for NOTECHS II²¹

Domain	Behaviours	Surgeon team	Anaesthetic team	Nursing team			
Leadership and management	Leadership						
	Maintenance of standards						
	Planning and preparation						
	Workload management						
	Authority and assertiveness						
Teamwork and co-operation	Team building/maintaining						
	Support of others						
	Understanding team needs						
	Conflict-solving						
Problem-solving and decision-making	Definition and diagnosis						
	Option generation						
	Risk assessment						
	Outcome review						
Situation awareness	Notice						
	Understand						
	Think ahead						
1	2	3	4	5	6	7	8
Consistent	Inconsistent	Consistent	Inconsistent	Inconsistent	Consistent	Inconsistent	Consistent
Behaviour compromises patient safety and effective teamwork		Behaviour in other conditions could directly compromise patient safety and effective teamwork		Behaviour maintains an effective level of patient safety and teamwork		Behaviour enhances patient safety and teamwork, a model for all other teams	

Glitch count

Events within the theatre process have been given several descriptive terms in the literature, including 'minor problems', 'operating problems'³⁷ and 'surgical flow disruptions'.³⁸ In each case, categories to aid description were provided including (but not exclusively) technical and environmental factors, technology and instruments, issues relating to training and procedures, teamwork and patient factors. The theoretical basis for deriving the proposed set of categories was to extend the theories proposed by Reason⁹ and Helmreich,¹² in acknowledging the variety of system-sourced factors that can contribute to the visible imperfections in system performance. The categories are broadly based on the SEIPS model.¹⁷

Description of the glitch count method

Glitches are defined as 'deviations from the recognised process with the potential to reduce quality or speed, including interruptions, omissions and changes, whether or not these actually affected the outcome of the procedure'.⁸ To capture these, direct observations are made of entire operations from the time the patient entered the operating theatre to the time they left, by pairs of observers comprising one clinical and one HF's researcher. The glitches were collected independently by each observer, individually noting the time and detail of the glitch within data collection booklets. This results in a set of glitches captured by each observer. These are de-duplicated and summed to provide a total glitch count for an operation. We recorded the detail of the glitch (e.g. 'diathermy not plugged in when surgeon tried to use it') along with the associated time point. All glitches were categorised post hoc (*Table 7*) and entered into a secure database. The observers spent a period of 1 month in training and orientation to the data collection

TABLE 7 Glitch categories with definition and examples

Glitch category	Definition	Example
Absence	Absence of theatre staff member, when required	Circulating nurse not available to get equipment
Communication	Difficulties in communication among team members	Repeat requests; incorrect terminology; misinterpretations
Distractions	Anything causing distraction from the task	Telephone calls/bleeps; loud music requiring to be turned down
Environment	Aspects of the working environment causing difficulty	Low lighting during operation causing difficulties
Equipment design	Issues arising from equipment design, which would not otherwise be corrected with training or maintenance	Compatibility problems with different implant systems; equipment blockage
Maintenance	Faulty or poorly maintained equipment	Battery depleted during use; blunt equipment
Health and safety	Any observed physical risk to personnel	Mask violations; food/drink in theatre
Planning and preparation	Instances that may otherwise have been avoided with appropriate prior planning and preparation	Insufficient equipment resources; staffing levels; training
Patient related	Issues relating to the physiological status of the patient	Difficulty in extracting previous implants
Process deviation	Incomplete or re-ordered completion of standard tasks	Unnecessary equipment opened
Slips	Psychomotor errors	Dropped instruments
Training	Repetition or delay of operative steps as a result of training	Consultant corrects assistants operating technique
Workspace	Equipment or theatre layout issues	De-sterilising of equipment/scrubbed staff on environment

Source: reproduced from Morgan *et al.*²² This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>.

methods before any real-time data were collected. Whole operating lists were observed whenever practical. If a patient left theatre mid-operation (e.g. to go to radiology), the observations were paused until the patient returned to theatre.

Development of the glitch count

A sample set of 94 glitches were collected during the initial training phase, grouped into common themes, and then assigned titles and definitions (Table 8). The reliability of the categorisation process was assessed using Cohen's kappa. Agreement was good between the four observers [$\kappa = 0.70$, 95% confidence interval (CI) 0.66 to 0.75]. Two observers (one clinical, one HF) were present at each operation. In contrast with previous methodologies,³⁷ no immediate evaluation of the glitch significance was made, as the impact of a particular glitch on process or outcome is context dependent. Prior to the final analysis, all four observers reviewed the glitch data jointly. Glitches noted by both observers were categorised by consensus when there was a difference and an overall glitch score was assigned comprising the sum of all unique glitches seen (i.e. those unique to observer A plus those unique to observer B plus those in common). Some glitches were deleted (if the team considered this event was not a glitch), split (if the contextual data contained more than one glitch occurrence) or recategorised during this consensus process.

Reliability and validity of the glitch count

A total of 429 operations were observed between November 2010 and July 2012 and 5742 glitches were observed. The total number of glitches observed in a single operation ranged from 0 to 83 (mean 14 glitches).

We investigated possible differences in the profile of glitches that each observer collected in theatre (see Table 7). Of the 5742 glitches, 64% were observed by the HF observers and 76% were observed by the clinical observers ($p \leq 0.001$). The clinical observers consistently noted more glitches per operation than

TABLE 8 Difference in observed glitches between observer specialties (p -values based on chi-squared test)

Glitch category	Observer, n (% of category)			Total observed, n (% of total)	Difference, % (95% CI)	p -value
	HFs and clinical	HFs	Clinical			
Absence	123 (42.1)	202 (69.2)	213 (72.9)	292 (5.1)	3.8 (-3.9 to 11.5)	0.362
Communication	128 (38.3)	218 (65.3)	244 (73.1)	334 (5.8)	7.8 (0.5 to 15.1)	0.036
Distractions	585 (43.6)	887 (66.1)	1039 (77.4)	1342 (23.4)	11.3 (7.9 to 14.8)	< 0.001
Environment	5 (33.3)	8 (53.3)	12 (80.0)	15 (0.3)	26.7 (-12.4 to 65.7)	0.245
Equipment design	224 (37.6)	379 (63.7)	440 (73.9)	595 (10.4)	10.3 (4.9 to 15.7)	< 0.001
Equipment maintenance	146 (52.5)	206 (74.1)	218 (78.4)	278 (4.8)	4.3 (-3.1 to 11.7)	0.273
Health and safety	171 (40.4)	243 (57.4)	350 (82.7)	423 (7.4)	25.3 (19.1 to 31.5)	< 0.001
Patient related	36 (30.0)	49 (40.8)	107 (89.2)	120 (2.1)	48.3 (37.1 to 59.6)	< 0.001
Planning and preparation	304 (38.5)	495 (62.7)	596 (75.5)	789 (13.7)	12.8 (8.2 to 17.4)	< 0.001
Process deviation	227 (37.0)	375 (61.1)	465 (75.7)	614 (10.7)	14.7 (9.4 to 20.09)	< 0.001
Slips	256 (50.4)	386 (76.0)	377 (74.2)	508 (8.8)	-1.8 (-7.3 to 3.7)	0.562
Training	36 (23.4)	70 (45.5)	120 (77.9)	154 (2.7)	32.5 (21.6 to 43.4)	< 0.001
Workspace	67 (24.1)	165 (59.4)	180 (64.7)	278 (4.8)	5.4 (-3.0 to 13.8)	0.221
Overall	2308 (40.2)	3683 (64.1)	4361 (75.9)	5742	11.8 (10.1 to 13.5)	< 0.001

Reproduced from Morgan *et al.*²² This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>.

the HFs observers, but the difference varied markedly between glitch categories. Clinical observers noted a much larger proportion of environmental, training, health and safety, and patient-related glitches, while there was minimal difference between the observers for absence, slips and equipment maintenance.

Agreement between observers was assessed using Fleiss's kappa for multiple observers (a chance-corrected proportional agreement). A value of zero indicates no agreement better than chance and a value of one indicates perfect agreement. Values can be interpreted³⁹ as < 0.20, poor; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, good; and 0.81–1.00, very good. The agreement between the four observers is shown in *Table 9*.

The World Health Organization's checklist evaluation method

The WHO launched the Safe Surgery Saves Lives campaign⁴⁰ in January 2007 to improve consistency in surgical care and adherence to safety practices. In June 2008, the WHO's surgical safety checklist¹¹ was designed to help operating room staff improve teamwork and ensure the consistent use of safety processes. The WHO's checklist has become one of the most significant and widely used innovations in surgical safety of the last 20 years.^{41–44}

Large benefits have been reported following implementation of the checklist, including reductions in adverse events⁴¹ and cost savings.⁴³ Based on these reports, the WHO's surgical safety checklist has been regarded as highly successful. Claims have been made for its capacity to induce indirect changes, such as improved situational awareness, in line with the evidence that structured briefings and checklists improve factors such as team communication and information sharing.³⁹

The WHO's surgical safety checklist has since been taken up by national health-care governance organisations in the USA and the UK to ensure target compliance. In the NHS, hospitals are required to audit and report their adherence rates to meet set targets, and compliance is encouraged by a financial incentive. However, clinical audit studies of the WHO's surgical safety checklist compliance have questioned the quality of compliance with the time-out (T/O) and sign-out (S/O) sections of the checklist.⁴⁴

Audits of the WHO's surgical safety checklist compliance to fulfil regulatory requirements intend to record whether or not all sections of the checklist are performed. However, these audits commonly record only the fact that an attempt was made, and not whether or not the attempt was adequate to fulfil its intended purposes. Reports of problems with implementation and application⁴⁵ of the WHO's surgical safety checklist in the operating room have suggested that achieving compliance to a level in which benefit can reasonably be expected is more complex than expected.⁴⁶

Development of the World Health Organization's checklist evaluation

A sample of the WHO's form for each site is provided within the data collection booklet to assist the observers in noting whether or not the WHO's form is completed or which parts are consistently missed.

The categories for scoring the WHO's checklist completion are taken from Lingard *et al.*³⁹

TABLE 9 Agreement between observers on observed glitches

Summary	Cases agreed
Agreement between the four observers	
Agreement between all four observers, <i>n</i> (%)	28 (56)
Agreement between three or more observers, <i>n</i> (%)	42 (84)
Kappa (95% CI)	0.70 (0.66 to 0.75)

The time (hour:minute) of commencing the WHO checks and the duration of checks (minute:second) are collected along with free-text contextual data regarding the content.

Description of the World Health Organization's checklist evaluation method

Observers recorded whether or not a T/O or S/O was attempted and, if it was, the attempted process was critiqued on three quality parameters: whether or not all information was communicated, whether or not all the team was present, and whether or not active participation was noted. These parameters each received a yes or no result individually. 'All team present' required all the team taking part in the operation at its commencement to be present; 'all information communicated' required all points on the checklist to be verbally addressed; and 'active participation' required whole-team interaction and engagement with checklist completion. Observers recorded these parameters independently for the T/O and S/O, including the onset time and the duration of the process. Who led the T/O and S/O was also recorded, being defined as the team member who read out loud the checklist questions to the rest of the team. Following the conclusion of each operation, observers compared findings and resolved any disagreement by discussion.

The observation of the WHO T/O and S/O was performed within the context of a wider study of work, the Safer Delivery of Surgical Services (S3), which aims to quantify the effect of QI interventions on theatre processes and safety. The observation of the WHO T/O and S/O was selected as an outcome measure for the S3 programme. Within this paper we are presenting all pre-intervention assessment data.

Sign-in data were not available as they are completed in the anaesthetic room. Each of the points below is considered with a binary yes or no answer (*Table 10*).

Clinical outcome measures

The clinical outcome measures used were taken from hospital administrative data sets. Data on readmissions within 90 days, whether or not a complication has occurred and length of stay were extracted from hospital records in the three sites participating in the study [Nuffield Orthopaedic Centre, University Hospital Coventry and Warwick (including Rugby), and Kettering]. Baseline demographic information on age and sex was also extracted. For each consultant participating in the study in the active or control group, data were obtained for all his or her patients 6 months before and 6 months after the intervention was delivered. In order to further ensure anonymity and to avoid linking consultants to a particular case, consultants were combined into groups. These groups were then used to identify cases for a particular intervention for the before and after intervention periods, and the active and control groups in the extracted data set.

TABLE 10 Adherence to the WHO's checklist

Status	Yes/no
Content (relevant information communicated?)	
Occasion (patient awake?)	
Audience (all team present?)	
Participation (active?)	
Nature of completion	Yes/no
Concurrent task completion	
Concurrent conversations	
Concurrent absence	

Patient-reported outcome measures

Decisions made early in the process of developing the research ethics application for this study effectively prevented us from collecting the intended large representative sample of quality-of-life measures. Our intention to analyse EQ-5D responses to determine the effect of our interventions on patient quality of life was therefore frustrated.

Health economic evaluation methods

These methods are described in greater detail in the chapter on health economic aspects of the programme (see *Chapter 11*). Costs of training were estimated based on a standard programme of preliminary meetings, main training events and follow-up visits. Resource volumes were then multiplied by appropriate unit costs to calculate the cost of the intervention.

Qualitative methods

Design

A number of qualitative options were considered for the evaluation of the S3 programme. Semistructured interviews were deemed appropriate because they are less restrictive than other options such as structured interviews, surveys or questionnaires. The use of open-ended questions allows the interviewer to explore participants' experiences and attitudes.⁴⁷ They allow for a broader research question, which can be adapted, and altered, throughout the analysis process.⁴⁸ The purpose of the qualitative evaluation in this study was to retrospectively explore participants' experiences and perceptions of the S3 study in an attempt to better understand its successes and challenges.

Development of the interview guide

The initial interview guide used was designed by the main interviewer (LF) and an external collaborator (GM), who had not been involved in the rest of the S3 study. Discussions of the S3 project in general were held individually with some members of the research team (PM, LM, ER and SN) to give the main interviewer an overview of the entire S3 study. This was used in the development of the interview guide.

Participants

To enable as broad an understanding as possible of the S3 study, we decided to include the researchers involved in S3 in the pool of participants for interviews alongside clinicians and managers in the participating hospitals. A mixture of staff (management, doctors, nursing and clinical support workers) from each intervention arm were selected randomly from a consented staff database for recruitment, with some additional staff who had led the improvement projects added to this list. E-mails and/or telephone calls were made to each of these individuals. In total, 36 staff and 12 researchers were targeted for recruitment; 34 individuals (23 frontline staff and 11 researchers) were interviewed.

Interviews

Two researchers who had a lesser involvement in the wider S3 (LF and FC) carried out the interviews. Each interview was conducted in person, in a quiet room away from other individuals. The interviews were digitally recorded, and transcribed by the main interviewer (LF).

Analysis

Analysis of interviews was based on the constant comparison method.^{48,49} The analysis was conducted partly concurrently with data collection, in order to allow for development of the interview guide. The analysis process consisted of breaking down data into component conceptual units and generating codes and categories based on these, informed by both ideas developed prior to interviews (based on

previous literature and discussions with members of the S3 research team) and through more inductive analysis of the data. These codes were then developed, broken down and merged into wider themes that were used in our presentation in *Chapter 10*. This coding and analysis was conducted by the main interviewer (LF) using NVivo software version 10 (QSR International, Warrington, UK) to assist. The expert collaborator (GM) independently explored the data set and verified and further developed themes constructed by LF.

Statistical methods

Our primary outcome measures were the total NOTECHS II scale score, WHO checklist measures and glitch rate. For each of the five linked studies, differences between the control and the active arms was assessed using two-way analysis of variance (ANOVA) (group × time), with treatment (control vs. active) and time (pre intervention vs. post intervention) as factors. Differences between groups were assessed by the group–time interaction, effectively comparing the pre–post change in the active and control groups. Pre- and post-intervention differences are reported as 95% CIs.

In the pooled analysis of data from all five studies (see *Chapter 9*), the same approach was adopted using the relevant pooled data sets. For the clinical outcome measures, binary variables in the before and after periods were compared using odds ratios and 95% CIs from a logistic regression, and mean length of stay using linear regression, both adjusted for age and sex.

Agreement between observers (clinical and HF) for each dimension of the Oxford NOTECHS instrument was assessed using agreement and the intraclass correlation coefficient. Mean Oxford NOTECHS II scale scores were compared for operations with high and low WHO T/O scores²⁶ using the *t*-test. Differences between mean Oxford NOTECHS II scale scores for different specialties and sites were examined using ANOVA. When differences between specialties or between sites were observed, the *t*-test was used to explore where the difference lay; no adjustment was made for multiple testing. The correlation between Oxford NOTECHS II scale scores and glitch rate was explored graphically and quantified by Pearson's correlation coefficient. Differences in mean glitch rates per operation between the sites and specialties were examined by one-way ANOVA and *t*-tests. We considered *p*-values of < 0.05 to be statistically significant (with no adjustment for multiple testing). All statistical analyses were carried out in R version 3.0.1 (The R Foundation for Statistical Computing, Vienna, Austria), except for the analysis of clinical outcome data that was conducted in Stata version 12 (StataCorp LP, College Station, TX, USA).

To analyse the influence of culture and system interventions, we pooled the data from the relevant arms of the five studies to conduct comparisons addressing the following predefined questions based on our hypothesis:

- Is intervention generally more effective than no intervention? (Pooled results of all active groups, studies 1–5, vs. all control groups.)
- Are integrated system/culture strategies more effective than single-dimension interventions? (Pooled results of active study 4 and study 5 vs. pooled results of active studies 1, 2 and 3.)
- Are systems interventions more effective than a culture intervention? (Study 3 vs. studies 1 and 2.)

General approach to interventions

Each of the interventions followed a specific implementation format that included the following:

- encouragement of multidisciplinary involvement
- training content taught away from the clinical setting
- coaching/support delivered within the clinical setting.

Involvement of all staff grades and types was essential and required liaison with management to ensure that nurses, doctors, management, cleaning, clerical and ancillary staff were able to take part.

The interventions were all delivered by individuals with significant experience in the area. There were three main interventions (lean, SOPs and team training); the approach taken to the single and combined interventions is described in the following text.

Lean

Background theory and content delivered

Lean in health care is often taken to refer to a holistic philosophy of organisation which covers a broad range of issues including, for example, patient flow, the management of waiting times and the levelling of demand. Here we have adopted a view often in line with the interpretation of Spear and Bowen⁵⁰ which emphasises the workplace and process organisation, plan–do–check–act (PDCA) cycles and process visibility.

- (a) Workplace and process organisation focuses on the redesign of practices to minimise the possibility of human error. The initial step was to map out the processes that take place in the workplace in considerable detail, insisting on describing what actually occurs rather than what is mandated or supposed to happen. The points in the process at which problems tend to occur are then identified largely by the staff team themselves, with assistance from pre-intervention observations by the research team. The staff team were then tasked with proposing a redesign of the work process to make the errors or problems they have experienced impossible. This process relies greatly on detailed staff experience of what works in their specific environment.
- (b) PDCA (the Deming cycle)⁵¹ describes an approach to participative exploration and experimentation in which new ideas for improvement are subjected to systematic testing prior to roll-out. The use of PDCA is essential to ensure that complex systemic problems are appropriately assessed and that participant ideas for solutions were subjected to testing and further adapted, rather than being implemented at too immature a stage in their development. The testing process relied on repeated small, simple audits that allowed very rapid assessment of the likely success of innovations.
- (c) Process visibility ensured that participants can easily see the state of a system, ensuring that potential problems are identified early. As part of each improvement cycle, a means of ensuring this is built into the process redesign. It is important that the methods used are themselves very simple and, as far as possible, require no effort from staff.

An initial stage in the classic lean process was a one-off exercise in tidying and rationalisation of the work environment to eliminate waste and ensure that equipment was always available and functional when required. Following this, specific projects were selected for action. Theatre staff decided among themselves on the areas to be addressed and prioritised solutions themselves, with support from the research team. During the 3-month period we expected the team to tackle up to four specific problems using the lean process, but the number may be more or less than this depending on staff enthusiasm and capacity. Regular brief ad hoc QI meetings were held to review progress and decide on changes within each project. Each meeting was led by the member of the theatre staff team who had volunteered to take lead responsibility for that particular project, initially with support from members of the research team, and involved as many other members as practicality and necessity allowed. These meetings were conducted in a confidential and non-punitive environment. With agreement from management (director of nursing and director of surgery), the staff were given licence to make changes to working practices and policies within their own theatre. For every project, a method of audit, developed by the staff, was put in place to examine the efficacy of the intervention and inform the PDCA cycle. This aided the research team in understanding the progress of the team towards increased safety and quality of care. The nature of lean as a method for addressing the unique problems of each theatre meant it was not possible to identify exactly what improvement projects emerged. Early demonstration projects were led by the

research team, with the aim of handing over implementation of lean to the ward staff by the end of the intervention phase.

Training and follow-up support approach

Training was led by Dr New, through a series of short seminars beginning in the first week of the intervention, and continuing throughout. The total contact time required from the staff team was about 8 hours and was broken up, as required, into sessions of ≥ 2 hours.

The classroom-based session was run as an off-site 'away day' and concluded with the team identifying an area for process improvement and developing an action plan for taking the work forward. When staff were unable to attend on the main day, and in lieu of this, an extra half-day session was run a few days earlier. Preparation for the away day included a pre-course visit by the trainer and consultation with the observers who had spent time within the site collecting observational data.

Standard operating procedures

Background theory and content delivered

The concept of SOPs was explained, together with the tools used by HFs professionals to evaluate existing processes, such as the various types of task analysis, process mapping and failure mode effects analysis. The concept of PDCA cycles to introduce change was explained. Rather than SOPs imposed by authority with mandatory compliance audited, we preferred to use staff-led projects to develop standardised approaches, as this approach encourages team discussion and may produce more effective interventions through more intimate knowledge of potential system gaps and common deviations from correct procedures. The intervention was designed as a facilitated introductory course on the principles of SOPs, with the surgical team given choice on where to apply standardisation and the format of any associated SOP processes and documentation.

1. Detailed written instructions to achieve uniformity of the performance of a specific function.
2. An organised list of directives that establish a standardised course of action.
3. A method of functioning that has been established over time in order to execute a specific task or react to a specific set of circumstances.
4. Chronological steps to follow and decisions to make in carrying out a task or function.

These encapsulate some emphasis on the 'document' (1 and 2) and some on the 'doing' (3 and 4). Clearly both of these aspects are essential.

Standard operating procedures can be a confusing term in a hospital (because of the inclusion of the term 'operating') and, therefore, care was taken to ensure that it was clear that the focus was on all aspects of the process in and around the operating room, rather than the operative procedure itself.

The team were given ideas about the choices to make when considering a SOP (*Box 1*) and asked to make their own selection dependent on the perceived needs of the situation.

The team were then supported in the development of the SOP(s) specific to the area of work they had chosen to focus on.

Training and follow-up support approach

As with the lean interventions, training was led by Dr Steve New through a series of short seminars beginning in the first week of the intervention and continuing throughout. The total contact time required from the staff team is about 8 hours, which can be broken up, as required, into sessions of ≥ 2 hours.

BOX 1 Choices to consider when constructing a SOP

Something for a crisis.

Something that exists to achieve better outcomes.

Designed to be followed exactly.

A way to do something.

Linear (do a, then b, then . . .).

Something you do once/rarely.

Something that reminds you of what you know.

Something for the inexperienced (learning).

Designed to enhance collectivity.

Something for routine use.

Something that exists for protection.

Designed to be followed approximately.

The way to do something.

Contingent (if x, then y).

Something you do repeatedly.

Something that tells you something you don't know.

Something for the experienced (doing).

Designed for individual.

The classroom-based session was run as an off-site 'away day' and concluded with the team identifying an area for process improvement and developing an action plan for taking the work forward. When staff were unable to attend on the main day, and in lieu of this, an extra half-day session was run a few days earlier. Preparation for the away day included a pre-course visit by the trainer and consultation with the observers who had spent time within the site collecting observational data.

Teamwork training

Background theory and content delivered

The TT was delivered by external consultants, in one morning and two evening sessions. The training consisted of educational content on the aetiology of human error from a psychological perspective,

together with a discussion of teamwork based on the aviation CRM model, including the importance of situational awareness and explaining the value of checklists.

Included in the training was the introduction to when technical excellence was not enough, concerning the issues that make a working day easier or more challenging. The impact of adverse events in the NHS was discussed, with a discussion of the common themes (cognition errors, decision-making faults, loss of situation awareness, communication issues, inappropriate team dynamics and lack of assertion). Some relevant findings from psychological research on decision-making, working memory, social inhibition and control and perception and attention problems were included and their relevance to theatre work discussed. Personality types and their impact on team interactions were also included together with learning styles. Lessons from industries with a strong safety record were discussed, along with useful models of error applicable to health care.^{9,13,18} The role of non-technical skills was introduced, along with the role of checklists and briefings.

Training and follow-up support approach

The training was delivered in one 8-hour or two 4-hour sessions off site. Following this teaching, the active theatres received 5 days' in-theatre coaching focusing on aspects of non-technical skills and the WHO's surgical safety checklist completion. The training and the coaching were delivered by Atrainability Ltd (Cranleigh, UK), a commercial provider of this type of training in the NHS.

Chapter 4 Standard operating procedures: study A

Introduction

There have been widespread efforts to address the issue of iatrogenic injury in surgery.^{41,52} In an attempt to mitigate risks, many efforts have focused largely on either work culture or standardisation.^{53,54} Such approaches to reducing risk and error have been seen across many high-reliability organisations. SOPs have been used across many such industries including the military, manufacturing and aviation. They involve devising a comprehensive, written framework to make functions and tasks more explicit and uniform. The use of SOPs has been demonstrated to be beneficial in many health-care domains.⁵⁵ The WHO's surgical checklist is one of the most recent and best known attempts to implement a SOP to improve surgical safety.⁴⁶ There has been some research to demonstrate its benefits; however, there have also been studies suggesting that effective implementation of checklists to standardise can be challenging.

Methods

Design

The study design was a controlled interrupted time series. It consisted of three phases: pre intervention (baseline data collection), intervention (active only) and post intervention (follow-up data collection). Each of the phases lasted 3 months, with a total study duration of 18 months.

Setting

This intervention was set in a tertiary referral centre that specialised in orthopaedic and reconstructive surgery. The centre had 106 beds and six operating theatres. Both the control and the active teams specialised in lower limb orthopaedic surgery, with typical procedures including operations such as knee arthroscopic procedures, knee primary and revision arthroplasty, and hip primary and revision arthroplasty. Care was taken to ensure that contamination between the active and the control teams did not occur.

Interventions

This study was split into two types of interventions: the primary intervention and the secondary intervention. The primary intervention consisted of the S3 training in SOPs and any support provided by the S3 research team. The secondary intervention was the staff-led improvement project. This included any improvement activity or projects conducted by the staff who were trained as part of the primary intervention.

Primary intervention

The primary intervention consisted of training in the principles of SOPs and follow-up support from the research team for the duration of the intervention period. Staff were invited to attend two 2-hour training sessions. This was delivered on-site by an expert in the field. In addition to covering the principles of SOPs, training sessions covered the concepts of PDCA cycles and implementing change. Those who attended the training included surgeons, nurses, anaesthetists and administrators. Following training, staff were asked to identify areas which they believed would benefit from improvement. An action plan was then designed to take this forward. Follow-up support was then provided by the team, which included facilitation and encouragement, along with additional support for some data collection and analysis.

Secondary intervention

The improvement area identified was how to reliably record and communicate the tasks required for the operating lists. A project team was developed, consisting of a consultant surgeon, a registrar and two theatre nurses. The decision was made to trial a briefing tool, which would take the form of a whiteboard, on which all information that would usually be shared verbally would be written up to share with the whole

team and, importantly, information transfer would be standardised across team members. In addition, questions could be posted on the board and answered later by others. The physical nature of the board evolved from a temporary single-use whiteboard attached to the theatre wall through to a permanently fixed board. An accompanying SOP was developed to guide and assist in the standardising of recording of pertinent list-related information including surgical and anaesthetic plan, equipment required and any high-risk pieces of information.

Evaluation

Primary intervention

The impact of the primary intervention was evaluated using both process measures and clinical outcomes. The process measures used were team non-technical skills (Oxford NOTECHS II scale), intra-operative process reliability (glitch count) and WHO checklist compliance. These measures were chosen as they reliably quantify process and safety measures which may have been affected by the initial training programme. (Please see *Chapter 3* for further information on these measures). The clinical impact was evaluated by measuring mortality, readmissions, complications and length of stay.

Observational data in theatre were collected 3 months pre and post intervention for evaluation purposes. Observations were conducted by two members of the team: one with a surgical background and one with a HF/psychology background. Each of these observers were oriented to the other field and were both provided training in data collection methodologies prior to the study.

Secondary intervention

Frontline staff identified information flow and communication of tasks for operating as an issue. The whiteboard briefing tool was developed by the project group and evolved based on feedback from the staff. Whether or not the briefing tool was used effectively was recorded for each case observed post intervention.

Data analysis

A two-way ANOVA was used to examine differences across time (pre- vs. post-intervention groups) and condition (active vs. control). Differences between groups were assessed by the group–time interaction. This statistical analysis was carried out using R version 3.01. Mean age was compared before and after the intervention using *t*-tests. Chi-squared tests were used to compare sex distribution. Three of the clinical outcome variables were binary in nature (mortality within 30 days, readmissions within 90 days and having at least one complication). These variables were compared before and after the intervention using 95% CIs and odds ratios from a logistic regression, and were adjusted for age and sex. Linear regression was used to compare mean length of stay across time, controlling for sex and age. The significance level was set at 1% based on the number of comparisons made. This statistical analysis was conducted in Stata version 12.

Results

Primary intervention

Overview

Twenty-five operations were observed in the control group before and after the intervention (total of 50). Twenty-six cases were observed pre intervention in the active group and 29 were observed post intervention (total of 55). The average operating times were similar across pre-intervention (control, 1 hour 52 minutes; active, 1 hour 46 minutes) and post-intervention (control, 1 hour 44 minutes; active, 2 hours 7 minutes) groups.

World Health Organization compliance

There was minimal change in T/O pre and post intervention across groups (*Table 11*). The difference observed in the active and control groups was not significant. All three requirements for T/O were

TABLE 11 World Health Organization checklist completion

Section of WHO checklist	Group, n/N (%)			
	Control		Active	
	Pre intervention	Post intervention	Pre intervention	Post intervention
T/O performed	21/25 (84)	23/26 (88)	4/25 (96)	25/29 (86)
Three components (communication, all team present, active participation)	14/25 (56)	5/26 (19)	11/24 (44)	8/29 (28)
S/O performed	0/25 (0)	1/26 (4)	0/25 (0)	2/29 (7)

No significant differences in change before and after the intervention between the active and control groups (two-way ANOVA).

satisfactory pre intervention in both the active and the control groups. A decrease was observed in both the active and the control groups post intervention. There was no difference in the before-and-after change between groups. S/O was the same across pre-intervention groups. A small increase was observed in groups post intervention, but the difference was not significant.

Non-Technical Skills II scale

Minimal change was observed in the mean NOTECHS II scale scores post intervention in both the active group (before, 74.84; after, 73.79) and the control group (before, 72.52; after, 72.88) (*Table 12*). There was no significant difference in the change in groups. Analysis of staff subgroup level revealed no differences between groups for surgeons, nurses and anaesthetists.

Glitch rate

Mean glitch rate per hour per operation was similar across both groups (active, 4.75; control, 4.92) before the intervention (see *Table 12*). A rise in mean glitch rate was observed after the intervention in both groups (active, 7.80; control, 9.79). There was no significant difference in change between the groups.

TABLE 12 Primary intervention results

Outcome measure	Group			
	Control		Active	
	Pre intervention	Post intervention	Pre intervention	Post intervention
NOTECHS II, mean score (SD)	72.52 (9.09)	72.88 (8.65)	74.84 (7.43)	73.79 (8.36)
WHO T/O attempted, n/N (%)	21/25 (84)	23/26 (88)	24/25 (96)	25/29 (86)
WHO T/O satisfactory, n/N (%)	14/25 (56)	5/26 (19)	11/24 (44)	8/29 (28)
WHO S/O attempted, n/N (%)	0/25 (0)	1/26 (4)	0/25 (0)	2/29 (7)
Glitch rate per hour (SD)	4.92 (3.54)	9.79 (4.12)	4.75 (2.68)	7.80 (4.79)
90-day readmissions, n (%)	9 (1)	4 (1)	4 (1)	8 (2)
Complications, n (%)	61 (10) ^a	79 (20)	17 (4) ^a	35 (7)
Length of stay in days (SD)	5.7 (8.4) ^a	5.4 (7.6)	1.7 (4.3) ^a	(3.7)

SD, standard deviation.

^a $p < 0.05$ (t-test) for baseline values: no difference in change before and after the intervention between the active and control groups (two-way ANOVA).

Clinical outcomes

Complications were significantly higher in the control group than in the active group. Length of stay was also significantly higher in the control group, but remained relatively stable in both groups across time (from pre to post intervention). Readmission was low across groups, both pre and post intervention (see *Table 12*).

Secondary intervention

The board supported a written SOP to standardise recording of list-related information including surgical and anaesthetic plan, equipment required and any high-risk pieces of information. The briefing tool summarising this information was used in 29 out of 29 of the observed post-intervention operations.

Discussion

This study hypothesised that, by training and supporting staff in the use of SOPs for improvement, an increase in quality and safety in the operating theatre would be likely observed. The staff engaged in improvement activity and identified an area for improvement as intended: standardisation of the methodology and content used for pre-operation briefing. Although the secondary intervention had successes in that it was maintained and it also led to similar changes in practice by other teams, the primary intervention did not result in any statistically significant improvements in the primary outcome measures.

There are a number of features of the study design which add substantially to the validity of the findings. Primary outcome measures used were semiobjective, validated process measures or clinical outcome measures. The fact that data collection was conducted by two observers increases the reliability of the findings. Looking at the findings of the study, it is notable that the use of a control was highly important. By conducting a controlled study, we avoided any misinterpretation of results, which often happens when controls are not used. The use of a control allowed for the identification of secular trends in outcomes that otherwise may have been attributed to the intervention. This was the case for both glitch rate and complications. Two significant events occurred during the intervention which might offer an explanation for these trends: the death of a key staff member and the introduction of a new information technology (IT) system. Although glitch rate did increase, the increase was smaller in the active team, suggesting that the intervention may have actually lessened the trend in this case.

There are many plausible reasons as to why the primary intervention did not appear to have any direct effect on the outcomes. Although standardisation is frequently used as a tool for improvement, it has rarely been evaluated objectively in the past. In this study there did appear to be benefits in terms of the secondary intervention; the fact that the staff-led change was sustained and subsequently adopted by other teams in the unit supports this. However, this improvement was not reflected in the primary outcomes, suggesting that the metrics may not have been adequate or suitable in terms of capturing the benefits of this particular change. This issue with objective measurement is not a new one in the QI literature.

It is also possible, because the study was relatively small (based on a methodological choice to conduct observations of whole operations), that the study was not powerful enough to detect an effect in the primary outcome measures. Another plausible reason is that, although all care was taken to avoid contamination of active or control teams, some contamination still occurred. However, it was likely that there was no contamination of secondary intervention, as the briefing tool (the whiteboard) was only present in the theatre of the active group. There is also the risk that there was observer bias, as it was not possible to blind observers to the study group.

Another potential reason for limited impact on primary outcome measures could have been the choice of secondary intervention, the standardised briefing, tool. It could be argued that the staff targeted an area of the system which was perhaps not the most suitable and tackled a very small part of the system, making it difficult to link to quality and safety. Although the staff demonstrated engagement and

implemented a solution which reduced handover repetition and wasted less time at the start of the theatre day, this change in pre-list work would not have been captured in any of the metrics used in theatre.

The staff-led improvement approach required that staff choose their area for improvement and develop their own solution. The reason, as previously discussed in the report (see *Chapter 3, Lean*), was that staff were more likely to engage if it was staff led and changes were then more likely to be sustained (evidenced by observers noting the use of the whiteboard among more than just the active team a year on). This study suggests that a balance is possibly required between the experts (those delivering training and providing support) and staff-led improvement activity. Furthermore, as the improvement approach was for all activity to be staff led, this also made the selection of relevant metrics in the pre-intervention phase extremely difficult. During this phase staff had not yet been involved in training or identified their improvement project therefore making it near impossible to identify very specific metrics.

There were a number of issues during intervention implementation. The first issue was that it was difficult to find the time and opportunity to get all of the frontline staff together. It appeared that the benefits of any potential improvements rarely outweighed the cost of lost staff activity. In future, to make wide-scale improvement, it appears likely that organisational change to support safety interventions will be necessary. A second issue was that it was difficult to generate enthusiasm and support for standardisation, with many negative views towards it from frontline staff. This was believed to be related to a fear of loss of professional autonomy. Many similar issues and challenges have been echoed by others previously.^{56,57}

Training and support in SOPs and how to make change resulted in staff-led improvements that were maintained and generalised to other teams a year later. This improvement activity and subsequent benefits were not, however, reflected in the primary outcome measures. Staff-led improvement, while potentially leading to greater engagement and sustainability, is not without its challenges. Change from an organisational level may be required to scale up and promote further improvement efforts.

Chapter 5 Teamwork training: study B

Introduction

As described in *Chapter 1*, enhancement of the working culture has been one of the main interventional approaches used in attempts to improve the safety and reliability of health care, and the majority of the available evidence has been in the context of surgical operations. Most approaches to improving culture have been based on the aviation CRM model and there is now a significant literature describing intervention studies using this approach.^{13,15,58,59} Evaluation of this approach was a very important part of our programme, as CRM was the only consistent culture-modifying technique we identified in the literature, and we therefore chose it as our culture intervention to be tested alongside systems interventions to shed light on our main hypothesis. As described in *Chapter 2*, the nature of improvement interventions that address safety places significant constraints on the choice of study design, and for this reason we elected to conduct a series of identical controlled, interrupted time series experiments, using either single-intervention approaches or combinations of a culture intervention (TT) with a systems intervention (either SOP or lean; see *Chapter 2*). Thus, although evaluation of the effects of a TT intervention has been performed by numerous groups in the past in different settings, it was important for us to evaluate it using the same tools to implement and evaluate it as we used for the other studies in our programme. Controlled studies of TT effectiveness are in fact quite rare, so we were hopeful that our results would be of interest in their own right, but the main rationale for conducting this study was to allow comparisons of the effectiveness of TT alone versus TT in conjunction with systems interventions.

Methods

Design

This was one of five controlled interrupted time series studies using an identical format. Three-month baseline data collection in theatres was followed by a 3-month intervention period and a final 3-month observation period. Operations were observed on the same days of the week whenever possible; the observed group therefore represents a large convenience sample of the whole population operated on. Patient outcome data were extracted from hospital information systems for the 6 months before and after the intervention. Because of the small, tight-knit nature of the orthopaedic theatre set-up, it was necessary to use a non-orthopaedic control group. We therefore used the vascular surgery theatre as our control group. Interchange of staff between this group and the orthopaedic teams was minimal, reducing the risks of 'contamination' of the control theatre. The main operations performed were hip and knee replacements in the intervention group and varicose vein surgery, femoropopliteal artery bypass and inguinal hernia repair in the control group.

Setting

This study was conducted at a medium-sized district general hospital in the Midlands with a busy orthopaedic service dealing mainly with elective joint replacement. The hospital has 532 beds, of which 31 are for elective orthopaedic surgery. Two operating theatres are dedicated to elective orthopaedics on a semipermanent basis, in a main theatre suite of six theatres. The hospital has a separate day-case surgery unit that performs a range of more minor orthopaedic, vascular and general surgery procedures. There is a relatively small degree of nursing staff rotation between orthopaedics and other theatres, so the orthopaedic theatre nursing teams are generally quite familiar with each other and have developed an implicit modus operandi based on the preferences of a small number of orthopaedic surgery consultants who work with them regularly. There is no trainee programme and support for the consultants is provided by staff grade (non-training) junior staff. Staffing levels are therefore not as generous as in the other trusts

involved in the S3 study. Patients are cared for postoperatively on two wards, one of which also deals with orthopaedic trauma. Anaesthetic cover is likewise provided by a small pool of consultants with assistance from a small team of operating department practitioners. Radiology and anaesthetic equipment are adequate, but not as 'state of the art' as in the other trusts involved in the study. The hospital serves a diverse community, which is partly urban and partly semirural.

Intervention

The intervention was a course of teamwork and communications training based closely on aviation CRM. The course consisted of two 4-hour sessions of interactive classroom teaching, delivered by active and retired civil aviation pilots from Atrainability Ltd, who had an extensive background in CRM training for aircrew, and several years' experience of adapting this to training theatre staff. The concepts forming the core syllabus were:

- inevitability of human error
- cognitive limitations of humans
- types of human error
- latent errors and risks in systems
- models of risk and failure in work systems
- importance of standardised communication protocols and checklists
- work culture and safety – barriers to communication
- importance of sharing mental models
- situational awareness – red flags.

Specific attention was given to the relevance of the training to the performance of the WHO's surgical checklist. After completing classroom training, the trainers returned regularly to provide on-the-job coaching to each theatre over the next 6 weeks. We attempted to give training to all members of the surgical, nursing and anaesthetic staff who regularly worked in the intervention group theatres. We provided several opportunities to attend and negotiated free time and staff back-fill with management, as well as publicising the training in a number of different ways. In preliminary discussions we attempted to gain the engagement of the consultant surgeons and anaesthetists, theatre team leaders and theatre and surgical managers. We held meetings with theatre nursing staff to explain the ideas behind the training, to reassure them and to answer questions.

Evaluation

Intervention

Non-technical skills

The Oxford NOTECHS II (described in detail in *Chapter 3*) was used to evaluate non-technical team performance. As was our standard practice, paired surgical and HF observers conducted all the observations. The NOTECHS II scale produced a maximum possible score of 96, based on the addition of separate scores for the surgery, nursing and anaesthetic subteams. The four dimensions scored in each team are teamwork and co-operation, leadership and management, situational awareness, and problem-solving and decision-making.

Technical performance

The glitch count method was used to evaluate the frequency and nature of technical slips and errors by the team. The method is described in detail in *Chapter 3*. Variations from the intended activity are observed by the same observers as for NOTECHS II and classified into 1 of 13 categories (see *Chapter 3* for details). The glitch rate is calculated from the duration of the procedure and then the number of glitches agreed by the two observers.

World Health Organization checklist compliance

We evaluated compliance with the T/O and S/O parts of the WHO checklist in two ways, again using dual-direct observation. For both T/O and S/O we recorded whether or not an attempt was made to conduct the relevant part of the checklist and report the percentage of cases in which this occurred for both parts. We also scored the teams as yes or no on whether or not all team members were present, whether or not all parts of the T/O were performed and whether or not there was active participation of the team. We report the percentage of operations with a three out of three score for quality using this system, which we have described in detail elsewhere.²³

Clinical outcome data

Hospital Episode Statistics data were extracted for all patients undergoing operations in the relevant operating theatres under the involved consultants during the 6-month periods immediately before and after the intervention. This therefore represents a large group of patients, of which those whose operations were observed represented a large convenience sample. Data were independently extracted by trust staff and supplied to the research team in anonymised form. The information extracted for each patient detailed their age, sex, diagnosis, consultant, operation, operating time, length of hospital stay, complications (any) and nature, readmission within 90 days of operation and reoperation. The parameters used in comparisons between active and control groups were length of stay and number (%) of patients with any complication and readmissions within 90 days.

Data analysis

Differences between the control and active arms were assessed using two-way ANOVA (group–time interaction), with treatment (control vs. active) and time (pre intervention vs. post intervention) as factors. Differences between groups were assessed by the group–time interaction. Pre- and post-intervention differences are reported as 95% CIs. All statistical analyses were carried out in R version 3.0.1. For clinical outcome data, *t*-tests for mean age and chi-squared test for sex distribution were used to compare the before and after periods. Binary clinical outcome variables before and after intervention were compared using odds ratios and 95% CIs from a logistic regression, and mean length of stay using linear regression, controlling for age and sex in both regression models. Given the number of before and after comparisons performed, a 1% significance level was selected. This analysis was conducted in Stata version 12.

Results

Overview

Twenty-six operations were studied before the intervention in the active theatres and 11 in the control theatres, compared with 25 and 10 operations, respectively, after the intervention. The types of surgery performed remained stable throughout in both groups. The average operating time reduced from 1 hour 38 minutes to 1 hour 11 minutes in the control group, but remained static at about 1 hour 55 minutes in the active group. The total number of procedures performed in the theatres involved during the pre- and post-intervention periods were 650 and 690 for the intervention theatres and 598 and 413 for the control theatres.

World Health Organization checklist compliance

Time-out was attempted in 51 of the 72 observed operations. The T/O attempt rate improved significantly in the active group ($p < 0.001$), but also in the control group ($p < 0.001$) (Table 13). There was no significant difference between the degree of improvement between the active and the control groups. The completion of all three T/O components (communication, all team present and active participation) increased in the post-intervention phase, but also increased in the control arm. The increase in compliance was significantly higher in the active group ($p = 0.003$).

There was a small difference in the attempt rate of S/O between pre and post intervention in the active group, but no difference was observed between pre and post intervention in the control group. The difference between the change in the active and the control groups was not significant.

TABLE 13 World Health Organization checklist completion

Section of WHO checklist	Group, n/N (%)			
	Control		Active	
	Pre intervention	Post intervention	Pre intervention	Post intervention
T/O performed	5/11 (45)	10/10 (100) ^a	11/26 (42)	25/25 (100) ^a
Three components (communication, all team present, active participation)	0/11 (0)	2/10 (20)	4/26 (15)	23/25 (92) ^{a,b}
S/O performed	0/11 (0)	0/10 (0)	2/26 (8)	7/25 (28) ^b

a $p < 0.001$ for before and after change.
b $p = 0.003$ for difference in change before and after the intervention between the active and control groups (two-way ANOVA).

Non-Technical Skills II scale

The mean NOTECHS II scale score increased after the intervention in the active group, while it remained unchanged in the control group (Table 14). The difference between the change in the active and the control groups was statistically significant ($p = 0.047$). Subteam analysis revealed differences in mean NOTECHS II scale scores were not significant for surgeons and anaesthetists, while statistically significant for nurses ($p = 0.006$).

Glitch rate

The post-intervention mean glitch rate was significantly higher than the pre-intervention rate in the active group ($p = 0.002$), whereas it remained essentially unchanged in the control group (see Table 14).

Clinical outcomes

There was an increase in the complication rate in the active group after the intervention and a small decrease in the rate in the control group; the difference between these two just reached significance ($p = 0.05$) (see Table 14). There were minor changes in readmission rates and length of stay in both groups, but neither difference reached significance and trends were in opposite directions.

TABLE 14 Intervention results

Section of WHO checklist	Group			
	Control		Active	
	Pre intervention	Post intervention	Pre intervention	Post intervention
NOTECHS II, mean score (SD)	72.09 (3.36)	70.09 (5.70)	71.62 (5.69)	75.44 (5.53)
WHO T/O attempted, n/N (%)	5/11 (45)	10/10 (100)	11/26 (42)	25/25 (100)
WHO T/O satisfactory, n/N (%)	0/11 (0)	2/10 (20)	4/26 (15)	23/25 (92)
WHO S/O attempted, n/N (%)	0/11 (0)	0/10 (0)	2/26 (8)	7/25 (28)
Glitch rate per hour (SD)	10.31 (3.79)	10.79 (4.53)	7.21 (2.73)	10.20 (3.67)
90-day readmissions, n (%)	51 (8.5)	37 (9.0)	72 (13)	74 (11)
Complications, n (%)	162 (27.1)	106 (25.7)	140 (21.5)	185 (26.8) ^a
Length of stay in days, (SD)	4.82 (13.5)	4.93 (11.7)	5.09 (11.1)	5.38 (13.2)

SD, standard deviation.
a $p = 0.05$ for difference in change before and after the intervention between the active and control groups (two-way ANOVA).

Discussion

In this setting, a TT programme was successful in improving non-technical teamwork skills, but no significant improvements were seen in clinical outcomes, and team technical performance, as measured by the glitch rate, actually got worse. The frequency with which the WHO checklist T/O was attempted improved considerably, but this also increased in the control group, and the difference in change between the intervention and the control groups was not significant. However, the quality of the T/O increased significantly only in the active group. Interpretation of these inconsistent results requires attention to the specific challenges and circumstances of this study. Implementation of training was beset by a variety of logistic and organisational difficulties that may have detracted from its effectiveness, but it was ultimately delivered to most of the target staff group and did significantly improve non-technical skills, which was its primary target. WHO compliance would be expected to increase if non-technical skills were improving and did so. As with other studies of TT we have conducted, the nursing subgroup experienced the largest rise in NOTECHS performance. We are therefore reasonably confident that the improvement in non-technical skills was real and was produced by the training programme. The rise in glitches in this group is more difficult to explain. There were no major changes in operation type, staff, resources or other environmental influences on performance that could explain deterioration in the orthopaedic teams in isolation. A large percentage of the increase in glitches was a result of increased distractions. In a situation in which the team have been made more aware of the risks, but have not been given an opportunity or specific training in how to bring about QI and change in the system they are working in, it is possible that their dissatisfaction with working processes was increased by the training, leading to the increase in distractions. This remains speculative, however, and we cannot be sure why glitches increased in this group.

A marked improvement in the WHO checklist attempt rate for T/O was noted in the control group, which suggests a possible contamination problem. Although there was little or no staff transfer between the groups, the control group were aware of the study and the presence of study personnel watching their performance and this may have induced a type of 'Hawthorne effect' in relation to the part of their routine most closely linked in their minds with safety. The fact that attempt rate went up in both groups but the quality of T/O performance increased only in the group receiving training is consistent with this interpretation. The use of parallel controls was a major strength of this study. We could not arrange for these to be matched to the active group in terms of the surgery performed but this was of limited importance, as their principal function was to be sensitive to secular trends affecting the whole operating suite. In an observational study using non-blinded observers and semiobjective end points, there is a good deal of scope for bias due to Hawthorne effects and observer bias. However, the glitch results obtained were in the opposite direction to those expected from observer bias and it seems unlikely that this would have been operating for NOTECHS but not for glitch counts. The findings relating to clinical outcomes are also difficult to interpret. As the data on operations observed show, these were actually a small percentage of all cases passing through these busy operating theatres. The staff mix present in the large majority of procedures that were not observed is unknown and it may be that too few trained staff were present to make any difference to performance. Equally, it may be true, as indeed the observational data suggest that improved non-technical skills do not translate into improved technical performance without the addition of a further element, such as systems improvement training.

We can therefore conclude that the TT intervention used in this study appeared to be effective in improving non-technical skills and compliance with the WHO checklist procedures but did not improve technical performance or clinical outcomes. We have had discordant results in regard to the effect on technical performance in other settings and it may be that the effects here are context specific, relying, for example, on the scope and ability the team have to change their working environment. These results do not recommend TT as a single instrument to improve theatre team performance, but the effects on non-technical skills suggest that it may be useful if integrated with other approaches.

Chapter 6 Lean process improvement: study C

Introduction

Lean is a term used to describe an approach to systems redesign and continuous improvement which originated in the Toyota car company, and has been adopted in many settings worldwide.⁵⁰ It is characterised by a focus on reducing waste and inefficiency, using an approach that requires the deep engagement of frontline staff in both problem analysis and generation of solutions. Lean approaches have been applied frequently in health care, both to improve efficiency and to reduce risks and improve patient safety. Most recently, the NHS has seen a number of initiatives that are based on the principles of lean utilised by other high-reliability industries.^{60,61} The NHS Institute for Innovation and Improvement developed two major lean-based programmes, The Productive Ward and The Productive Operating Theatre, which attracted considerable interest and support; however, rigorous scientific evaluation of their effects was unfortunately lacking, which is a common occurrence in QI initiatives.^{62,63} This surprising dearth of valid evidence explains why a recent review concluded that, although there appears to be potential for lean approaches to QI in health care,⁶⁴ currently there is a need for further evaluation of the efficacy of such approaches in this setting, particularly when applied to the problem of patient safety.^{65,66} The appeal of lean is clear, in that analytical work on risk and error in health care has clearly identified that poorly designed systems and processes play a major role in creating risk. However, lean approaches are often difficult to operationally define. Furthermore, lean encompasses many different practices and often holds a variety of meaning to different individuals.⁶⁶ Lean is often seen as a 'top-down' initiative requiring senior management to make a decision to radically revise all work processes, although the basic concepts of lean in fact emphasise the importance of involving the production line staff. Another approach frequently taken by consulting firms is to emphasise the measurement and iterative improvement aspects of lean such as PDCA cycles. Thus a number of variants on the original lean idea have developed. The aim of this intervention was to examine the effectiveness of applying a 'light-touch' lean approach influenced by the work of Spear and Bowen⁵⁰ to surgical safety in theatre. This approach was defined as an experimental approach that is participative, with an emphasis on bottom-up improvement. The aim was to prioritise staff ownership of all improvement work⁶⁷ encompassing lean principles of Kaizen and Gemba, encouraging continuous improvement through the co-operation of all staff and a detailed analysis of the working environment and processes. This was the third intervention study of the S3, although the same approach to lean was taken in the subsequent study in which it was integrated with TT.

Methods

Design

The study design was a controlled interrupted time series. It consisted of three phases: pre intervention (baseline data collection), intervention (active only) and post intervention (follow-up data collection). Each of the phases lasted 6 months, with a total study duration of 18 months.

Setting

This intervention was set in theatres performing orthopaedic trauma (active) and elective orthopaedic (control) surgery in a large acute teaching hospital in the UK. Some elective surgery was performed at the main hospital site, in a very large theatre complex where the trauma theatre was also situated. Other elective procedures were performed in a small specialist hospital some miles away, in which there was a focus on high-throughput elective hip and knee replacement surgery. As trauma surgery duties rotated among surgical teams according to day of the week, it was initially difficult to identify a stable team who could be trained to carry out the intervention. The rota system meant that the staffing of the trauma

theatre was constantly changing. However, we identified that an intervention focused on the same day each week would allow us to work with the most stable team available, although some flux in the team remained. A control team in the small specialist hospital was selected so as to ensure that there would be no crossover between the two services and to eliminate chances of contamination. Operations conducted in the two settings were typical for each of the services in a UK setting, generally consisting of fixation of common fractures in the trauma theatre and replacement of arthritic hip and knee joints in the elective orthopaedic theatre.

Interventions

The intervention in this study was the application of lean to the systems problems of the trauma theatre; however, it quickly became clear that this needed to be viewed in terms of a primary and a secondary intervention. The primary intervention consisted of the S3 training in lean improvement principles and any additional support provided by the S3 research team. The secondary intervention was the project or suite of projects which the staff decided to pursue in the subsequent staff-led improvement project. The strong emphasis of the lean philosophy adopted was that staff should be allowed to identify their key problem and develop solutions without outside direction, which is felt to reduce engagement and ownership of the subsequent project by staff. This included any improvement activity or projects conducted by the staff who were trained as part of the primary intervention.

Primary intervention

The primary intervention consisted of training in lean improvement principles and follow-up support from the research team for the duration of the intervention period. Training consisted of a 1-day course and an additional half-day course (for those who could not attend the former) covering lean principles and experimental approaches to improvement and delivered by an expert in the field. All staff were invited to attend. Those who attended the training included surgeons, nurses, anaesthetists and administrators. The training focused on the central concepts of lean: the elimination of waste (*muda*), the process of continuous improvement (*kaizen*) and the design of systems in which problems are made visible (*jidoka*). The discussion also drew out the need for experimental, participative approaches to process improvement. Following training, staff were asked to identify areas which they believed would benefit from improvement. An action plan was then designed to take this forward. Follow-up support for staff was provided by the research team and included facilitation and encouragement, along with additional support for some data collection and analysis.

Secondary intervention

The improvement areas chosen for focus by frontline staff were patient and information flows, with a specific concern in terms of delays in starting the operating list. It was deemed important because of the fact that it impacted on other theatre lists being delayed or postponed, potentially increasing the risk of harm to patients. Staff aimed to reduce the delay and implemented changes via PDCA cycles to allow for iterative development. The improvement project is presented in *Results*.

Evaluation

Primary intervention

The impact of the primary intervention was evaluated using both process measures and clinical outcomes. The process measures used were team non-technical skills (Oxford NOTECHS II), intraoperative process reliability (glitch count) and WHO checklist compliance. These measures were chosen as they reliably quantify process and safety measures that may have been affected by the initial training programme. Briefly, NOTECHS II assesses team non-technical skills by scoring each of three subteams (anaesthetic, surgical and nursing) for each of four dimensions of leadership and management; teamwork and co-operation; problem-solving and decision-making; and situation awareness. An 8-point Likert scale tied to exemplars is used to give a theoretical maximum possible score of 96. The glitch count recorded all process deviations, whether or not consequential, and categorised these into 13 types. WHO checklist

performance was evaluated first by recording whether or not the T/O and S/O procedures were attempted, where they were performed and how well they were performed. We used a simple yes or no three-point assessment for this, gauging whether or not all team members were present, whether or not all the points on the checklist were covered and whether or not active participation by the team was seen (see *Chapter 3* for further information on these measures). The clinical impact was evaluated by measuring mortality, readmissions, complications and length of stay.

Observational data in theatre were collected 3 months pre and post intervention for evaluation purposes. Observations were conducted by two members of the team, one with a surgical background and one with a HF/psychology background. Each of these observers was oriented to the other field and both were provided training in data collection methodologies prior to the study.

Secondary intervention

Frontline staff identified information and patient flow as areas for improvement, with the focus of their work aiming to reduce delays to the start of the operating list. A project group, including trainee doctors and specialist nurses, was formed with a consultant surgeon as the lead. Initially, data were collected in theatre to establish a better understanding of the process and current problems. A number of issues were identified as causing potential delays. These included the manner in which patients were/were not prepared for transfer to theatre, the sequence in which anaesthetists completed their preoperative tasks, and communication delays/failures between the pre-list meeting, the operating theatre and the ward. Further data were collected in the pre-list meeting room, in theatre and on the wards. The order in which events occurred was recorded, commencing at the beginning of the day and ending at the start of the second operation in theatre. Timings were also collected to generate time charts. These timings established the duration of tasks and provided a real-time overview of the process before intervention. These data were then used by the project group to inform iterative changes to the process. Follow-up data were also collected after the project, to allow for comparison before and after the project.

Data analysis

For the primary intervention, a two-way ANOVA was used to examine differences across time (pre vs. post intervention) and condition (active vs. control). Differences between groups were assessed by the group–time interaction. This statistical analysis was carried out using R version 3.01. Mean age was compared pre and post intervention using *t*-tests. Chi-squared tests were used to compare sex distribution. Three of the clinical outcome variables were binary in nature (mortality within 30 days, readmissions within 90 days and having at least one complication). These variables were compared before and after the intervention using 95% CIs and odds ratios from a logistic regression, adjusted for age and sex. Linear regression was used to compare mean length of stay across time, controlling for sex and age. The significance level was set at 1%, based on the number of comparisons made. This statistical analysis was conducted in Stata version 12. Statistical analyses conducted for secondary interventions were carried out in Microsoft Excel® 2010 (Microsoft Corporation, Redmond, WA, USA).

Results

Primary intervention

Overview

In total, 17 active and 21 control operations were observed pre intervention. Post intervention, 25 active and 16 control operations, were observed. The records of patients operated on were collected pre intervention (active, 224; control, 352) and post intervention (active, 292; control, 173). The mean operating time was similar in both groups (active, 2 hours; control, 1 hour 45 minutes). Little change (> 5 minutes) was observed in operation duration post intervention.

Non-Technical Skills II scale

An increase was observed in mean NOTECHS scale score post intervention for both active (73 to 77.84) and control groups (71.31 to 78.06), but the difference between these changes was not statistically significant. Analysis at a staff group level showed no significant change in post-intervention NOTECHS scale scores for surgeons, anaesthetists or nurses.

Glitch rate

Mean glitch rate per hour was 7.85 before the intervention in the active group. This decreased to 6.59 after the intervention. The mean hourly glitch rate in the control group was 6.52 before the intervention, which increased to 7.94 post intervention. However, the difference in glitch rates from pre to post intervention was not statistically significant between the two groups.

World Health Organization compliance

Individual increases were observed across all three components of the T/O post intervention for the active group (*Table 15*). Increases were observed in two out of three components for the control group. Results demonstrated an 18% increase in completion of all three components of T/O in the active group post intervention. An increase of 10% in the same was observed in the control group post intervention.

Clinical outcomes

No change in readmissions was observed in either the active or the control groups (*Table 16*). Complication rates in the active group increased very slightly post intervention, but decreased somewhat in the control group. Length of stay in days declined in both groups. Differences between the pre- and the post-intervention values were not significantly different between the active and control groups for any of these outcomes.

Secondary intervention

Changes introduced as part of the improvement project included new patterns of communication after the pre-list meeting, new protocols for how this meeting was co-ordinated and a set of guidelines regarding the sequencing of tasks. The results of these changes in terms of their impact on the team process and, specifically, the process of getting patients to theatre in a timely fashion, are presented in *Table 17*. There was an increase in the presence of anaesthetists at the trauma meeting (from 33% to 66%). A decrease in the number of changes to the theatre list was also observed. The time the first patient of the day arrived in the anaesthetic room decreased post intervention by about 20 minutes, a clinically significant improvement. The changes did not, however, lead to improvement in the average time of 'knife to skin' for the first patient (the primary target of the intervention), owing to an increase in the time patients spent in the anaesthetic room.

TABLE 15 World Health Organization checklist completion

Section of WHO checklist	Group, n/N (%)			
	Control		Active	
	Pre intervention	Post intervention	Pre intervention	Post intervention
T/O performed	20/21 (95)	16/16 (100)	17/17 (100)	24/25 (91)
Communication	16/21 (76)	12/16 (75)	7/17 (41)	13/25 (52)
All team present	15/21 (71)	15/16 (94)	9/17 (53)	17/25 (68)
Active participation	18/21 (86)	15/16 (94)	11/17 (65)	20/25 (80)
S/O performed	0/21 (0)	1/16 (6)	0/17 (0)	1/25 (4)

No significant differences in change before and after the intervention between the active and control groups (two-way ANOVA).

TABLE 16 Primary intervention results

Section of WHO checklist	Group			
	Control		Active	
	Pre intervention	Post intervention	Pre intervention	Post intervention
NOTECHS II, mean score (SD)	71.81 (7.72)	78.06 (6.57)	73 (7.1)	77.84 (11.59)
WHO T/O, n/N (%)	20/21 (95)	16/16 (100)	17/17 (100)	24/25 (96)
WHO S/O, n/N (%)	0/21 (0)	1/16 (6)	0/17 (0)	24/25 (96)
Glitch rate per hour (SD)	6.52 (3.06)	7.94 (4.01)	7.85 (2.69)	6.59 (3.95)
90-day readmissions, n (%)	130 (19)	55 (18)	94 (20)	102 (18)
Complications, n (%)	95 (14)	32 (10)	47 (10)	70 (12)
Length of stay in days, (SD)	10.2 (20)	7.6 (16)	10.3 (25)	7.7 (15) ^a

SD, standard deviation.

^a $p < 0.01$ for reduction in length of stay (linear regression with age and sex adjustment): no significant difference in change before and after the intervention between the active and control groups (two-way ANOVA).

TABLE 17 Secondary intervention results

Evaluation metric	Pre intervention	Post intervention
Mean PTWR start time (hours and minutes)	8 hours 12 minutes	7 hours 59 minutes
Mean PTWR first patient review time (hours and minutes)	8 hours 16 minutes	8 hours 5 minutes
Mean PTWR second patient review time (hours and minutes)	8 hours 41 minutes	8 hours 8 minutes
Any communication with anaesthetist? (yes)	30%	44%
% of days on which operating plan was changed after PTWR	20	0
Patient reviewed by anaesthetist/other?	40%	100%
Did first patient on printed list remain first on list? (yes)	50%	100%
Did the plan change from the printed plan during the trauma meeting?	70%	22%
Mean time porters/staff arrive to collect first patient from ward (hours and minutes)	8 hours 18 minutes	8 hours 10 minutes
Mean time first patient left ward (hours and minutes)	8 hours 29 minutes	8 hours 15 minutes
Mean time patient arrived to anaesthetic room (hours and minutes)	8 hours 48 minutes	8 hours 25 minutes
Mean time patient entered theatre (hours and minutes)	9 hours 11 minutes	9 hours 22 minutes
Mean time surgeon in theatre for case (hours and minutes)	8 hours 46 minutes	8 hours 38 minutes
Mean time preparation started (hours and minutes)	9 hours 22 minutes	9 hours 29 minutes
Mean time knife to skin (hours and minutes)	9 hours 37 minutes	9 hours 37 minutes

PTWR, post-take ward round.

Discussion

We hypothesised that using a 'light-touch' approach to training and supporting staff in lean QI methods would have an impact on the quality and safety of the surgical service they provided, but the results of this study do not support this hypothesis. No statistically significant improvements were observed across either the process or the outcome measures for the primary intervention. The results of the secondary intervention show a very marked improvement in part of the process to which most of the effort was directed, but not in the predetermined measure of success. On the face of it, the results therefore demonstrate little or no impact of this intervention on quality and safety. Closer examination, however, shows that the intervention did have some success and the differences between what was achieved and what was intended teach some important lessons about the value and limitations of lean approaches to patient safety.

There are a number of reasons which can be offered as to why the intervention had such little impact. It is probable that the key issue was the approach we adopted in ensuring that the improvement work was staff led, which presented us with a number of issues. As improvements were primarily led by frontline staff, engagement and motivation of this group was vital for success. This was not a simple task, with many staff reluctant to get involved. The term lean may have evoked scepticism in staff and resistance to participation resulting from previous failed initiatives. Even for those highly motivated and engaged staff members, the demands of clinical work made it difficult to ensure all staff were able to attend training. We therefore felt it was important to allow staff considerable freedom in choosing their projects and objectives. The project scope chosen for the secondary intervention, targeting the delays in getting the first patient to theatre, was not likely to address the biggest risks to patient safety in the process and was arguably more efficiency oriented. It is also a very complex project requiring co-operation from many different actors for success. Arguably, it may have been best to give some sort of guidance initially as to what types of projects were (1) more closely related to safety and (2) easy to measure/demonstrate if there is an impact. However, engaging staff by allowing them to identify issues and design their own solutions is clearly one of the major strengths of this type of lean approach and appears to be associated with strong sustainability once changes are implemented. An important lesson was therefore that lean methodology, at least using this staff engagement-based approach, presents a challenge of managing the tension between engagement and control. Too much freedom for staff to decide the direction of the work may result in activity which is not likely to benefit safety (although it may fulfil other agendas). Too much direction, on the other hand, may disempower and disengage staff, leading to lack of enthusiasm and poor sustainability.

Another factor contributing to the negative results could be the 'light-touch' approach to training during the primary intervention. This approach covered only basic lean principles and may have been inadequate to provide the skills necessary for improvement work. It also may have not been enough to challenge some existing cultural beliefs. Following the brief training we provided on-site coaching and support during the project, although most of this was supplied by more junior team members and not the specialist in lean who led the project, for logistic reasons. Therefore, it may be that this project suffered from an intervention with too low a 'dose' of expert support in the forms of initial training and subsequent advice on project development. The extent of training was determined by an assessment of the likely maximum access we would have to staff time (which proved fairly accurate) together with reference to the previous literature on successful training interventions in patient safety.⁶⁸⁻⁷⁰ We noted that several successful interventions seemed to have used only a half-day course of classroom training and, therefore, set this as the minimum acceptable target for staff. While it seems plausible that longer training for more people might have provided an improved result, our experience suggests that it would not have been feasible in the current NHS environment, in which releasing any staff time for training is very difficult. Our intervention may therefore be seen as the highest practicable dose and, therefore, a reasonable one to use in an assessment of practical value.

A third issue was the fact that staff were reluctant to make large changes in case it was too costly to the hospital. This represents a form of 'learned helplessness', but also to some degree a realistic recognition by

the staff of the inherent limitations of our approach. Without adoption as a management priority, with appropriate allocation of major resource, it was not realistic to suggest changes which would require major expenditure or potentially interfere with the normal running of other services within the theatre or the surgical department. This observation underlines the limited potential of small-scale change based only on clinical staff within the microsystem. To make a major change requires a more systemic approach with strong and active management support.

Another perspective on our results would be that the secondary intervention was in fact remarkably successful at doing what it intended to do, but, because of the divergence of the project direction from the original intentions, the predefined measures were not appropriate to demonstrate any impact. The measures of theatre process show that a remarkable improvement in timely arrival of the first patient to the anaesthetic room was in fact achieved, but all of this advantage was then lost by an unexplained increase in the time spent in the anaesthetic room. We did not have time to explore the reasons for this and develop ways of eradicating this increase, but there was clearly potential to do so. This is in effect an argument that the duration of the intervention may have been too short to make an impact; experts in lean would tend to agree that the timescale we worked to was very ambitious in terms of making a change which could be measured and sustained, and for which the PDCA cycle process was considered complete. Again, our choice of timing was based on logistic necessity and an understanding of the limits within which improvement needs to be conducted in busy hospitals. It may be that a longer interval is essential to yield reliable effects, and that, if lean is to fulfil its potential as an instrument for change, negotiations and arguments will need to be pursued to persuade management and clinicians to allow projects to continue for much longer before evaluating them.

Logistical issues were a major problem during this study and should not be neglected as a possible reason for reduced efficacy for this type of intervention. One challenge was that staff were never routinely in the same place at the same time. This made it often very difficult to get all the relevant people together for improvement work to occur. The nature of clinical work meant that often people could not show up to meetings or would have to cancel or leave at the last minute. This resulted in some miscommunications and the loss of momentum, which took a great deal of effort from one single consultant to revive. This lack of time and space in which staff can meet and plan improvement work is a significant challenge to this type of work.

Although the use of a lean approach in this study was not successful, other studies have demonstrated considerable improvements using lean in short time frames.⁶⁸ This suggests that the lean approach taken in this paper may not have been suitable for the context in which it was applied. It also indicates that further research is needed to determine which specific elements of a lean approach are appropriate for use in health-care improvement.

The intervention trialled in this study did not solve the issue targeted, although the overall intervention was successful in encouraging QI work to occur. It prompted staff to come together, identify issues, generate ideas for solutions and implement changes. Future work should consider how to balance this participative approach with guiding the direction of the improvement work. Further research should also look to creating better conditions for improvement, such as creating the time and space in which staff can do this work.

Chapter 7 Standard operating procedures and teamwork training: study A and study B

Introduction

Much attention has been devoted to unintentional error in health care and the subsequent patient harm occurring as a result.⁷ Therefore, the need to develop strategies and interventions to increase safety and mitigate risk to patients has been widely acknowledged. Standardisation or SOPs have been used across many other industries and high-reliability organisations, including manufacturing and aviation, to improve safety and increase reliability. SOPs involve developing a detailed, written framework which makes actions, functions and tasks more uniform and clear. The aim of standardisation in doing so is to reduce reliance on human cognitive functions such as memory and attention. SOPs have been used often in health care in many forms (e.g. checklists or protocols);⁴¹ however, many barriers to effective implementation have been identified and rigorous objective evaluations of such approaches are rarely conducted. Another approach which has been utilised within the health-care context is TT,^{58,59,71} often drawing specifically from the aviation industry, incorporating concepts based on CRM. CRM focuses on threat and error management, leadership, communication, situational awareness and decision-making among the team as a whole. There have been some positive findings in terms of CRM on teamwork, but there has been less evidence in terms of clinical outcomes. This study aimed to evaluate an integrated intervention of SOPs and TT. By doing so it wished to test the clinical effectiveness of such an integrated intervention in improving quality and safety in surgery by targeting both culture and system simultaneously.

Methods

Design

The study design was a controlled interrupted time series. It consisted of three phases: pre intervention (baseline data collection), intervention (active only) and post intervention (follow-up data collection). Each of the phases lasted 6 months, with a total study duration of 18 months.

Setting

This intervention was set across two university hospital sites (one active and one control) specialising in lower limb orthopaedic surgery. Typical procedures included operations such as knee arthroscopic procedures and knee and hip primary and revision arthroplasty. Care was taken to prevent contamination between active and control teams.

Interventions

This study was split into two types of interventions: the primary intervention and the secondary intervention. The primary intervention consisted of training in teamwork (based on CRM) and SOPs, in addition to any additional support provided by the research team. The secondary intervention was the staff-led improvement project. This included any improvement activity or project conducted by the staff who were trained as part of the primary intervention.

Primary intervention

The primary intervention consisted of training in the principles of SOPs, TT and follow-up support from the research team for the duration of the intervention period. All theatre staff were invited to attend training sessions. SOP training focused on improving standardisation and visibility of processes. The TT was an interactive session, focused on awareness raising, followed up by in-theatre coaching on effective teamwork. The interactive TT session was delivered first, followed by the SOP training. All training was

held off site, with the exception of in-theatre coaching. Those who attended the training included surgeons, nurses, anaesthetists and administrators.

Secondary intervention

Following training, staff were asked to identify areas that they believed would benefit from standardisation. The teams identified areas which they believed required improvement and an action plan was formed; these included the morning briefing, sending for a patient and handing the patient over to recovery. Follow-up support was then provided by the research team which included facilitation and encouragement, along with additional support for some data collection and analysis.

Evaluation

Primary intervention

The impact of the primary intervention was evaluated using both process measures and clinical outcomes. The process measures used were team non-technical skills (as measured by the Oxford NOTECHS II scale), intraoperative process reliability (as measured by the glitch count) and WHO checklist compliance. These measures were chosen as they reliably quantify process and safety measures which may have been affected by the initial training programme. (See *Chapter 3* for further information on these measures.) The clinical impact was evaluated by measuring mortality, readmissions, complications and length of stay.

Observational data in theatre were collected 3 months pre and post intervention for evaluation purposes. Observations were conducted by two members of the team, one with a surgical background and one with a HF/psychology background. Each of these observers was oriented to the other field and both were provided training in data collection methodologies prior to the study.

Secondary intervention

Owing to the nature of the secondary interventions chosen by staff, the measures used for evaluation of the primary intervention were also relevant for evaluation of the secondary intervention, for example WHO compliance, and so no additional measures were used.

Data analysis

A two-way ANOVA was used to examine differences across time (pre vs. post intervention) and condition (active vs. control). Differences between groups were assessed by the group–time interaction. This statistical analysis was carried out using R, version 3.01. Mean age was compared pre and post intervention using *t*-tests. Chi-squared tests were used to compare sex distribution. Three of the clinical outcome variables were binary in nature (mortality within 30 days, readmissions within 90 days and having at least one complication). These variables were compared before and after the intervention using 95% CIs and odds ratios from a logistic regression and were adjusted for age and sex. Linear regression was used to compare mean length of stay across time, controlling for sex and age. The significance level was set at 1% based on the number of comparisons made. This statistical analysis was conducted in Stata, version 12.

Results

Primary intervention

Overview

In total, 101 operations were observed. Seventeen operations were observed in the control group pre intervention and 21 were observed post intervention. Thirty cases were observed pre intervention for the active group and 33 were observed post intervention.

World Health Organization compliance

Time-out was attempted in 94 out of 101 operations observed, with no significant differences observed between pre and post intervention for either the active or the control groups (*Table 18*). For the S/O, a significant difference was observed between pre and post intervention for the active group ($p < 0.001$). This was not observed for the control group. The difference between the change in the active and control groups was significant ($p < 0.001$).

Completion of all three components of T/O (communication, all team present and active participation) increased from pre to post intervention for the active group. A decrease was observed for the control group. The difference between the change in the active and control groups was statistically significant ($p < 0.001$).

Non-Technical Skills II

An increase in NOTECHS scale score was observed post intervention in the active group, while a decrease was observed in the control group (*Table 19*). The difference between the change in the groups was statistically significant ($p = 0.002$). Analysis at a team level revealed differences in mean NOTECHS scale scores that were non-significant for surgeons ($p = 0.067$) and nurses ($p = 0.093$), but statistically significant for anaesthetists ($p < 0.001$).

Glitch rate

Glitch rate decreased in both the active and control groups post intervention (see *Table 19*). The difference between the two groups was not statistically significant. The rate of planning and preparation glitches decreased by almost 50% in the active group, but remained unchanged in the control theatre. A reduction in process deviations was also observed in the active group. Distractions fell in the control theatre, but absences and process deviations increased.

Clinical outcomes

Readmissions were almost the same pre and post intervention in the active group (see *Table 19*). A decrease was observed in the control group. An increase in complications was observed in both groups post intervention. Length of stay remained similar pre and post intervention for both. No statistically significant differences were observed for changes in any of the clinical outcomes between groups.

Secondary intervention

To address issues identified, staff developed improvement projects utilising SOPs focusing on the morning briefing and the conduct of the WHO's surgical safety checklist. A project team was formed which led on the improvement activities, linking back with the wider group when relevant. The existing briefing sheet

TABLE 18 World Health Organization checklist completion

Section of WHO checklist	Group, n/N (%)			
	Control		Active	
	Pre intervention	Post intervention	Pre intervention	Post intervention
T/O performed	17/17 (100)	20/21 (95)	27/30 (90)	30/33 (91)
Three components (communication, all team present and active participation)	11/17 (65)	11/21 (52)	3/30 (10)	20/33 (61) ^a
S/O performed	2/17 (12)	0/21 (0)	0/30 (0)	21/33 (64) ^a

a $p < 0.001$ for difference in change before and after the intervention between the active and control groups (two-way ANOVA).

TABLE 19 Primary intervention results

Section of WHO checklist	Group			
	Control		Active	
	Pre intervention	Post intervention	Pre intervention	Post intervention
NOTECHS II, mean score (SD)	76.00 (4.21)	71.81 (7.72)	74.83 (3.72)	79.27 (8.28) ^a
WHO T/O attempted, <i>n/N</i> (%)	17/17 (100)	20/21 (95)	27/30 (90)	30/33 (91)
WHO T/O satisfactory, <i>n/N</i> (%)	11/17 (65)	11/21 (52)	3/30 (10)	20/33 (61)
WHO S/O attempted, <i>n/N</i> (%)	2/17 (12)	0/21 (0)	0/30 (0)	21/33 (64)
Glitch rate per hour (SD)	7.67 (4.45)	6.52 (3.06)	6.84 (3.05)	5.25 (3.17)
90-day readmissions, <i>n/N</i> (%)	112 (20)	84 (20)	72 (13)	74 (11)
Complications, <i>n/N</i> (%)	58 (11)	64 (15)	81 (14)	123 (18)
Length of stay in days, (SD)	10.6 (19)	10.5 (20)	3.6 (11)	3.0 (7.2)

SD, standard deviation.
^a $p = 0.002$ for difference in change before and after the intervention between the active and control groups (two-way ANOVA).

and WHO checklist were modified iteratively based on staff feedback and workshops. WHO checklist changes included:

- specification of team member responsible for leading WHO checklist sections
- modification of questions for clearer terminology
- specification of which team member each question is directed to and who should answer
- directions were written for which checklist sections should be read out loud
- specific text was added to confirm when the checklist section was complete
- overall guidelines were created, detailing modifications and how the checklist process should be carried out.

The morning briefing document was adjusted to include documentation of:

- essential team members to be present
- whether or not a copy of the theatre list is present and verified
- a dedicated section on the briefing sheet for each patient on the list, specifying:
 - their place in the list order for operation
 - the confirmed operation description including site and side
 - details of specialised instruments, blood tests required, patient allergies and any other specific requirements
- an area for general comments
- an area for potential hazards
- overall guidelines were created, detailing adjustments and how the briefing document should be completed.

Discussion

This study evaluated an integrated approach to improving quality and safety in surgery. The intervention demonstrated an improvement in theatre team performance, but no direct effect on clinical outcomes was observed. WHO checklist completion and non-technical scores, as measured by the NOTECHS II scale,

showed improvement in comparison with the control group. No statistically significant improvements were observed in terms of technical performance across either group. An improvement in clinical outcomes was not observed for the active or control group. There are a few plausible reasons as to why an improvement was not observed. It is likely that an improvement in such outcomes would require some improvement in both technical and non-technical performance. Although an improvement in technical performance was achieved, the parallel improvements in the control group means that this cannot be attributed solely to the intervention, highlighting the importance of using a control group. Another is the fact that the study was relatively small, to allow for observations of whole operations and, therefore, there is a possibility that the study was not powerful enough to detect any significant effect.

Similar to previous research which has looked at challenges to QI and contextual issues impacting intervention success, a number of issues were experienced during implementation and probably affected outcomes. Staff attitudes towards the training and improvement activity varied across the group from very engaged and supportive to extremely negative and sceptical. Issues with staff engagement and negativity towards improvement have been identified as barriers to QI previously.⁵⁶ An additional challenge was the fact that it was extremely difficult to have all staff members at the training days. As a result of significant management support, theatres were closed for a whole day to maximise attendance.

The staff-led approach to improvement was used in this study based on research suggesting that it is more likely to generate enthusiasm and thereby increase the sustainability of changes made. However, this approach is not without its challenges. First, the fact that staff do not choose their area of improvement until the intervention phase makes it difficult for researchers to pre-select measures that are likely to demonstrate the benefits of any changes which occur. Second, the role of the experienced practitioners in training and supporting staff facilitated the improvement. This is a difficult model to scale up and even more difficult for the NHS to replicate. An additional challenge is the potential for staff to choose projects which would not appear to improve their work processes; however, in the case of this study, this did not appear to be the case and the staff projects were appropriate.

The existing literature does not provide substantial support for the efficacy of either TT or SOPs, particularly in terms of clinical outcomes. There are many studies which report the effectiveness of CRM-based TT in health care; however, this is mostly in terms of non-technical performance, and many do not show any effect in terms of improving clinical outcomes.⁵³ The findings of this intervention are consistent with this literature. Prior SOP literature is mixed; while much of it is poor quality, there are some studies which have demonstrated improvements in clinical outcomes. Some such studies have suggested significant improvements as a result of implementing the WHO's surgical safety checklist.⁷² There is still a need, however, for more rigorous evaluation of SOPs including controlled and blinded studies. Follow-up studies are also required to determine the sustainability of such approaches.

There were a number of strengths and weaknesses to this study. First is the use of a control group, which prevents misinterpretation of results and the misattribution of certain findings to the intervention rather than secular trends. Data collection in theatre was conducted by two observers to increase reliability. Although observers could not be blinded to the group conditions and there is a potential for bias here, they were blinded in the case of clinical outcome data. Measures utilised in this study were well designed and validated. A final weakness is that there is no lengthy follow-up to determine whether or not changes were sustained in the long term.

This study has provided an evaluation of an integrated intervention of SOP plus TT, which generated staff-led engagement and the implementation of subsequent QI projects. Although some benefits were observed in terms of non-technical skills and in WHO checklist completion, no effect was observed in terms of technical performance or clinical outcomes. This is potentially related to the power of the study. Similar to other research, a number of implementation issues were observed, which likely had an impact on the improvement activity. In conclusion, this intervention has resulted in the recognition of effective strategies for improving team processes, which are considered to be related to clinical outcomes.

Chapter 8 Teamwork training and lean process improvement: study B and study C

Introduction

The prevalence of injury and harm in health care has led to many initiatives to improve patient safety.^{41,73} Recently health care has seen many attempts to translate lessons of safety and resilience from other high-reliability organisations such as those in the aviation or nuclear power industries. TT based on CRM has been applied in health-care contexts.^{53,74} This approach trains staff in communication and teamwork skills in theatre to use in difficult or stressful circumstances. Some evidence has shown such training to be beneficial in improving teamwork and technical performance,²³ although studies have indicated further research is needed to explain some variability in results. Lean principles have also been utilised in various different approaches in health care. Lessons have been taken from the lean manufacturing industry, in particular the Toyota production system.⁵⁰ This approach involves looking at how to reduce system-based errors and puts the emphasis on frontline staff taking responsibility for improvement activity. Both approaches have been used individually in health care; however, there is little evidence as to the benefits of combining such interventions. This intervention aimed to evaluate the effects of combining a double intervention, TT plus lean improvement methods, in improving safety in surgery.

Methods

Design

The study design was a controlled interrupted time series. It consisted of three phases: pre intervention (baseline data collection), intervention (active only) and post intervention (follow-up data collection). Each of these phases lasted 3 months, with a total study duration of 9 months.

Setting

This intervention was set in a tertiary referral centre specialising in orthopaedic and reconstructive surgery. The hospital was based on one site, with six operating theatres. Both active and control teams were based in this hospital. The active intervention was carried out with the reconstructive/plastic surgical teams. The control group was surgical teams delivering elective lower limb orthopaedic surgery. Operations that were carried out were typical for this type of unit. Reconstructive and plastic surgery largely consisted of elective upper limb surgery, osteomyelitis surgery and sarcoma surgery. The orthopaedic surgery mostly consisted of primary and revision hip and knee arthroplasty, and arthroscopic knee surgery.

Interventions

This study was delivered as two types of interventions: the primary intervention and the secondary intervention. A combination of training in lean process improvement and TT was the primary intervention in the study. The secondary intervention was any staff-led improvement activity which occurred as a result.

Primary intervention

The primary intervention consisted of a combination of training in lean process engineering and TT based on aviation-style CRM. Frontline staff were invited to attend three half-day sessions of training, which was delivered by external consultants. Staff then identified areas of their work which they wanted to target. Project groups were formed and the research team provided support in terms of facilitation and project management. An additional 5 days of teamwork coaching in theatre around the area of the WHO's surgical safety checklist was also provided.

Secondary intervention

Frontline staff identified a number of areas which they believed could benefit from improvement. The areas chosen for improvement were pre-list communication, WHO checklist use, debriefing, equipment procurement, preparation room layout, list design and awareness of sterile services. Staff worked to improve each of these areas, developing their projects and changes iteratively using PDCA cycles.

Evaluation

Primary intervention

The three main outcome measures used were team non-technical skills (as measured by the Oxford NOTECHS II scale), intraoperative process reliability (via the glitch count) and WHO checklist compliance (see *Chapter 3* for further information on these measures). These data were collected in theatre for both active and control groups pre and post intervention. Observational data collection was carried out by two members of the research team. The observers had either a surgical or a HFs/psychology background and each was oriented to the other field. Observers were also provided with training in data collection methodologies prior to the study commencing. Four additional outcome measures were also used to evaluate the impact on patient clinical outcomes. These comprised mortality, readmissions, complications and length of stay. Duration of operation was also recorded.

Secondary intervention

Evaluation of secondary interventions was based on data collected during the intervention phases. This consisted of PDCA data, in addition to some clinical and process data collected, depending on the nature of the secondary intervention.

Data analysis

A two-way ANOVA was used to assess group differences, with condition (control vs. active) and time (pre intervention vs. post intervention) as factors. Differences between groups were assessed by the group–time interaction. Differences in the results between pre and post intervention are reported as 95% CIs. All above statistical analyses were carried out in R, version 3.0.1.

Results

Primary intervention

Overview

In total, 96 operations were observed: 51 operations were observed in the control arm (26 pre intervention and 25 post intervention) and 45 were observed in the active arm (21 pre intervention and 24 post intervention). Mean operating time was somewhat longer in the active theatres (2 hours 17 minutes) than in the control theatres (1 hour 36 minutes), with no significant change occurring in time post intervention.

World Health Organization compliance

The results from data collected on compliance with the WHO's surgical safety checklist can be seen in *Table 20*. Overall, 88% of total operations attempted the WHO T/O and 17% attempted S/O. Differences observed pre and post intervention were not statistically significant in either the active or the control group. Quality of the WHO T/O compliance increased in the active group, with improvements observed in performance of all T/O elements (communication, all team members being present and active participation).

Completion of all T/O components increased from 14% to 71% in the active arm of the intervention. It also increased minimally in the control group, from 19% to 28%. The difference in change between the

TABLE 20 World Health Organization surgical safety checklist compliance

Section of WHO checklist	Group, n/N (%)			
	Control		Active	
	Pre intervention	Post intervention	Pre intervention	Post intervention
T/O performed	23/26 (88)	23/25 (92)	18/21 (86)	21/24 (88)
Communication	11/26 (42)	15/25 (60)	6/21 (29)	19/24 (79)
Active participation	13/26 (50)	10/25 (40)	10/21 (48)	18/24 (75)
Three components (communication, all team present and active participation)	5/26 (19)	7/25 (28)	3/21 (14)	17/24 (71) ^{a,b}
S/O performed	1/26 (4)	3/25 (12)	0/21 (0)	12/24 (50) ^a

a $p < 0.001$ for the change between before the intervention and after the intervention in the active group.
b $p = 0.032$ for the difference in change from before the intervention to after the intervention between the active and control groups (two-way ANOVA).

active and the control arms was statistically significant ($p = 0.032$). Increases in S/O were also observed post intervention in the active group (0% to 50%), and this change was statistically significant ($p < 0.001$). A small change was observed in the control group for S/O post intervention (from 4% to 12%), but this was not statistically significant ($p = 0.574$). No statistically significant difference in pre- to post-intervention change was found between the active and control arms ($p = 0.093$).

Non-Technical Skills II

In the active group, the NOTECHS II score increased from 69.81 before the intervention to 75.56 after the intervention. In the control, the NOTECHS II score remained unchanged (Table 21). Although the difference between the groups was not statistically significant, it was close, with a value of $p = 0.058$.

TABLE 21 Summary of outcome measures

Section of WHO checklist	Group				p-value
	Control		Active		
	Pre intervention	Post intervention	Pre intervention	Post intervention	
NOTECHS II, mean score (SD)	72.88 (8.65)	72.54 (4.78)	69.81 (7.52)	75.56 (9.33)	0.058
T/O, n/N (%)	23/26 (88)	23/25 (92)	18/21 (86)	21/24 (88)	1
S/O, n/N (%)	1/26 (4)	3/25 (12)	0/21 (0)	12/24 (50)	0.093
T/O success, n/N (%)	5/26 (19)	7/25 (28)	3/21 (14)	17/24 (71)	0.032
Glitch rate per hour (SD)	9.79 (4.12)	13.20 (5.37)	10.48 (2.68)	4.38 (2.50)	0.001
90-day readmissions, n (%)	3 (1)	7 (2)	3 (2)	12 (9)	0.33
Complications, n (%)	77 (18)	98 (28)	14 (9)	9 (7)	0.08
Length of stay in days, (SD)	5.2 (7.2)	6.0 (7.9)	2.5 (6.3)	1.6 (4.0)	0.095

SD, standard deviation.
p-values refer to difference in change before and after the intervention between the active and control groups (two-way ANOVA).

Analysis of NOTECHS scale score at a subteam level (surgeon, anaesthetist and nurse) revealed significant differences only for nurses ($p = 0.016$).

Glitch rate

A decrease in the number of glitches per hour was observed in the active group post intervention (see *Table 21*). An increase in glitches per hour was observed in the control arm. The difference between the two groups was statistically significant ($p < 0.001$).

Clinical outcomes

The results demonstrated an increase in readmissions in both control and active groups; however, this difference was not significant (see *Table 21*). A decrease in length of stay was observed in the active group, but an increase was observed in the control group. Similarly, complications fell in the active group but rose in the control group. Differences in complications and length of stay from pre to post intervention were non-significant.

Secondary intervention

Based on training, frontline staff identified areas for improvement and set up project groups to target these areas for change. Staff initiated seven improvement projects (*Table 22*). The projects varied in levels of success and completeness, with some being discontinued (e.g. list design). The projects, however, were a demonstration of active engagement of staff in improvement activity, which is a vital part of the primary intervention. Staff involved considered the target areas and the issues. The staff collected data to highlight and better understand the problems and, consequently, implemented a number of changes using PDCA cycles until the final solution was achieved.

Discussion

This study demonstrates the successful use of an integrated intervention of lean and TT in improving process and clinical outcomes in surgery. Statistically significant improvements were observed across a number of outcome measures including glitch rate per hour, quality of the T/O, nursing team NOTECHS scale score and completion of S/O. Improvements were observed in length of stay and complications in the active group, which worsened in the control group. However, similar to improvements observed in overall

TABLE 22 Projects arising from the intervention

Project title	Project aim	Outcome
Briefing	Begin a pre-list briefing to understand what work is planned for the whole list	Institution of routine pre-list briefing
WHO checklist use	Improve reliability of WHO T/O and S/O	Standardisation of process: when it should occur and who should be involved
Debriefing	Feeding information back in to the system to reduce repetition of error and glitches	Production of a de-briefing feedback reporting system
Equipment procurement	Reduce waste of stock drift and improve financial planning	Development of new intraoperative recording sheets to reduce loss of stock
Preparation room organisation	Standardisation of layout of prep rooms to reduce waste	Standardisation of prep rooms
List design	Improve the amount of information available on the operating lists	Unable to change because of impending new computer system
Awareness of sterile services function	Reveal previously hidden processes involved in sterile services	Presentation and a question and answer session with TSSU staff

TSSU, theatre sterile surgical services unit.

NOTECHS score and T/O attempt rate, the differences here were not statistically significant, which is probably because of a lack of statistical power. Many changes observed in the active group were not observed in the control group (e.g. length of stay or complications), therefore suggesting that such improvements were not likely to be a result of any bias or secular trend.

This study provides evidence for the effectiveness of an integrated lean-style intervention, in which the emphasis is on staff-led QI. It supports other research which indicates that such improvement is more successful when led by staff. However, it was recognised during the study that staff-led approaches can also have difficulties. Staff may not always choose the projects which are most closely related to safety or may choose projects that are likely to fail. This results in tension for the research team, as there becomes an issue of whether or not the team should help and provide some guidance or if they should risk losing staff engagement by interfering. It was noted as a significant issue during the study.

In terms of TT, it is possible that this intervention experienced a ceiling effect because non-technical scores were already quite high across both active and control groups. However, the improvement projects chosen were mostly quite appropriate for trying to improve surgical safety. It is possible that this was in part a result of the TT, which focused on patient safety issues in the workplace.

The use of a control group is a major strength of this study. Doing such work as a controlled study is difficult, particularly because of the complex and dynamic nature of health care. The fact that this study successfully utilised a contemporary control, examined concurrently with the intervention group, which was similar enough for comparison but disparate enough to avoid contamination, demonstrates the feasibility of such robust research methods in this setting.

There were a number of challenges during this study. One of these challenges was the issue in having all frontline staff present at the training days. It proved difficult for staff to be released from their clinical work to have time to attend training and also to do the project work. Although the research team went to great lengths to make training and meetings as convenient as possible for staff, this still proved to be a major barrier to improvement activity. In other industries it is more typical for staff to be allocated specific time in which they can engage in improvement work regularly. This time protection should be considered at senior levels in health care for improvement work to be feasible.

Another challenge during this study was the learned helplessness experienced by staff. Staff did not appear to believe that they could make change and this made securing staff engagement in the study difficult. Addressing this required significant training, support and encouragement.

One limitation of this study was that observers were not blinded as to which staff were in the active or control groups. Therefore, there was a potential for observer bias when data were collected post intervention in theatre. However, not all data were collected by direct observation. Outcome measures such as complications and length of stay were not collected by observers and, therefore, improvements observed here cannot be accounted for by observer bias.

Another limitation is perhaps the amount of external support required during the intervention. The research team were present to provide the additional resources and this extra support; however, it is unlikely that the NHS would be able to do so at a widespread scale. Therefore, further work is needed to determine how to best scale up such interventions.

This study has demonstrated the combination of lean process engineering and TT to be a relatively sound intervention for improvement in health care. The use of the integrated approach addresses both system and human issues, and in this study has provided evidence of improvement across both process and clinical outcome measures. The results of this study support the hypothesis that integrated interventions targeting both culture and system are more effective, although further comparative studies are needed to support this.

Chapter 9 Integrated interventions: study A, study B and study C

Introduction

Previous chapters have outlined the limited amount of credible research regarding the effectiveness of any one single approach to improving safety and quality in health-care delivery and, in particular, in surgery. Our review of the literature in *Chapter 1* identified two strong themes in the research on interventions to improve safety and quality, namely culture improvement and system improvement approaches. The rationale for our research strategy in this programme has been based on the lack of research we identified evaluating the use of integrated approaches. This chapter outlines the final intervention study in the S3 programme. The first three studies focused on the evaluation of the single interventions alone: SOPs (A); TT (B) and lean improvement methods (C). The fourth study (A and C) and fifth study (B and C) focused on double-component interventions. This sixth, and final, intervention study evaluated the clinical and process effectiveness of combining all three approaches (TT, lean improvement methods and SOPs) in improving safety in surgery. We also attempted to integrate into our intervention strategy some of the lessons that we had learned during the earlier studies, and which are described in detail in other chapters. At the same time we extended our studies to investigate the combination approach across a wide range of hospital activities covering the patient pathway for a specialty, rather than restricting our focus to the operating theatre. We had already identified that one problem of a narrow focus on theatre activity was the extent to which episodes of actual patient harm were multicomponent. Therefore, reduction in harm was likely to require attention to the entire patient pathway rather than to the operation alone.

This final study attempted to integrate the learning from previous studies and carry out a pilot investigation of their value in a wider surgical environment. We did not attempt to carry out any form of comparison with the single- or double-component intervention approaches, as we felt this would have been impractical and premature; the main purpose of this study was to investigate the feasibility of the integrated approach in a whole-system context, and to gain some insights into its potential effectiveness. For this reason, and for practical reasons which will become clear as the project is described, we did not carry out contemporary observations on a control unit. In view of the greatly increased complexity of the environment, and taking into account the lessons learned from the earlier studies, we extended the intervention period in this final study to 9 months.

Methods

Design

This study design was an uncontrolled interrupted time series evaluation of an integrated intervention including SOPs, lean and TT components. There were three phases: pre intervention (baseline data collection at 3 months), intervention (9 months) and post intervention (follow-up data collection at 3 months).

Setting

The study was set on the neuroscience ward of an acute teaching NHS trust. The neuroscience ward consisted of 75 beds and included both neurosurgical and neurology patients. The unit is a tertiary referral centre. It has a high-volume practice in the surgical and interventional management of subarachnoid haemorrhage (SAH), brain tumours and neurotrauma. It also provides specialist spinal surgery and deep-brain stimulation services. The unit had 12 consultant neurosurgeons and includes a 75-bed specialist

neurology/neurosurgery intensive therapy unit (ITU). As the ward was a mixed surgical–medical ward, physicians were also invited to participate in improvement activity but declined.

Interventions

As with all of our interventions with a lean component, we used a staff-led approach in which training in the principles of systems improvement was integrated with discussion and analysis of the most important problems perceived by staff in delivering a safe and high-quality service in their environment. This led on to the selection of a set of target problems defined by the clinical staff and to the instigation of a large number of different projects. Using our adopted descriptive approach (see *Chapter 6*), we have referred to the training programme as the primary intervention and subsequent individual projects as the secondary interventions.

Primary intervention

The intervention consisted of providing staff training in the integrated approach and providing follow-up support for staff to lead improvement work themselves. All staff from the department were invited to attend training. Several half-day sessions were hosted to maximise opportunities for attendance. In addition, a full 2-day course was offered to individuals with specific interest in leading a project. The content of the component parts of the training provided remained the same as in our previous one- and two-component intervention studies, although some reduction was necessary to allow delivery of the whole package in the time available. Those in attendance at the training included surgeons, nurses, anaesthetists, managers, clinical support workers and administrators.

Secondary intervention

Following a process of group discussions involving staff and S3 personnel, a number of improvement projects were agreed on and commenced. An action plan was designed by staff to take each project forward. Interested staff were asked to volunteer as leaders and staff formed project groups based on the areas they were interested in improving. Changes introduced as part of the projects were carried out in small experimental steps, using PDCA cycles. The research team provided support in the form of facilitation, encouragement, data collection and statistical analysis.

Evaluation

Primary intervention

Unlike the previous S3 studies, this study did not collect intraoperative data such as WHO checklist compliance or NOTECHS II scores (see *Chapter 3*). The secondary intervention was largely set on the ward because of the focus that staff chose for the majority of their projects and because the theatre-based project attempted was of a nature that would have made these team performance measures inappropriate as outcome measures. We therefore used a wider variety of outcome measures in the final project than in other parts of S3. Overall clinical outcome data were collected on the same basis as in the previous studies, covering 30-day mortality, 90-day readmission, complications and length of hospital stay; however, each project within the overall intervention had its own more proximate and appropriate outcome measure. In addition, we collected PROMs (EQ-5D) for patients involved in this study.

Patients were provided with research information and asked for consent by a member of the research team on the ward or during the pre-assessment clinic (for surgical patients). At this point the patients also completed an initial baseline EQ-5D questionnaire and received a second follow-up EQ-5D questionnaire in the post 10 weeks after discharge, which they were requested to complete and return. The EQ-5D questionnaire asks five simple questions that capture important aspects of quality of life including mobility, self-care, activities, pain and anxiety. A large convenience sample was considered appropriate, as patients could be admitted at any time of day or night, and full coverage of all admissions was therefore impossible with our limited staff. We compared the change in EQ-5D score from baseline to follow-up in the cohort of patients studied before and after the 9-month intervention period.

Secondary intervention

Evaluation of secondary interventions varied based on the type of improvement activity chosen by the staff. It typically consisted of comparison of ward-based clinical/process data collected pre and post intervention relevant to the nature of the project in question (e.g. falls incidence for a falls project, glitch rate for a ward rounds project). PDCA cycles were conducted during the intervention phases of each improvement project. Data recorded for these small tests of change were often different from those used to evaluate outcomes. Clinical and process measures for evaluation of the secondary intervention could not be predetermined, as they were selected once staff had identified their improvement activities.

Data analysis

All statistical analyses were conducted using IBM Statistical Product and Service Solutions (SPSS) version 20 (IBM Corporation, Armonk, NY, USA). Differences in EQ-5D scales between the baseline and follow-up phases, pre and post intervention, were assessed using two-way ANOVA (EQ-5D stage \times time), with EQ-5D stage (baseline vs. follow-up) and time (pre intervention vs. post intervention) as factors. Differences between groups were assessed by the group–time interaction. Differences from pre to post intervention are reported as 95% CIs.

Results

Primary intervention

The outcome measures common to all of the preceding projects were the clinical outcomes and data. At the time of writing, analysis of the clinical data sets is incomplete, so we have deferred presentation of these results until this is accomplished.

European Quality of Life-5 Dimensions data

In the pre-intervention phase, 266 patients completed a baseline EQ-5D questionnaire, out of which 159 patients returned the 10-week follow-up questionnaire. In the post-intervention phase, 204 patients completed the questionnaire at baseline, out of which 151 follow-up questionnaires were returned. Although mean baseline and follow-up scores differed across pre- and post-intervention phases, there was little difference observed in mean score changes between baseline and the 10-week follow-up across the five domains of mobility, self-care, activities, pain and anxiety (*Table 23*). A greater improvement between mean baseline and follow-up was observed in the post-intervention phase than in the pre-intervention phase. The difference between the change in pre- and post-intervention scores was not statistically significant for any of the scales (see *Table 23*).

TABLE 23 European Quality of Life-5 Dimensions scores

Domain	Pre-intervention scores		Post-intervention scores		<i>p</i> -value
	Baseline	Follow-up	Baseline	Follow-up	
Mobility	0.60	0.46	0.74	0.54	0.071
Self-care	0.38	0.25	0.51	0.32	0.324
Activities	0.87	0.80	1.00	0.76	0.065
Pain	1.01	0.81	0.88	0.70	0.736
Anxiety	0.46	0.41	0.48	0.38	0.270
Overall health	63.45	68.83	58.53	67.63	0.246

p-values are for the difference in the change before and after the intervention between the active and control groups (two-way ANOVA).

Secondary interventions

As indicated, our approach led to the enthusiastic generation of improvement projects for a variety of aspects of the patient management journey from admission to discharge. A considerable number of these projects did not succeed for a variety of reasons, which are discussed in more detail in *Chapter 10*. We therefore report here on the projects that progressed far enough to produce some post-intervention data and follow this with a summary of the aims of those projects that did not get that far.

Secondary intervention: falls project

Background

Falls are a common cause of patient harm in all hospital environments, but the dangers in neurosurgery are unique and are a very challenging high-risk population when it comes to avoiding harm from falls. Various attempts had been made on the ward in the past to address patient falls, all of which had been unsuccessful. The nursing staff were keen to find an effective way of tackling this issue. Data sourced from the incident reporting database Datix® (Datix, London, UK) supported nursing claims that patient falls were a major problem experienced on the ward.

Methods

The project was conducted across three phases: pre intervention (2 months), intervention (6 months) and post intervention (2 months). During the pre-intervention phase initial reviews of falls data (from Datix) were used to identify the most common cause of patient falls. This turned out to be making trips to the toilet. Intentional rounding, which can be described as a systematic approach to visiting patients every hour or every 2 hours to ensure that their basic care needs are met, is an intervention which had been suggested in the literature as potentially effective in reducing falls. Nursing staff therefore decided to trial intentional rounding but with modification or redesign to suit the needs of their ward. In the intervention phase, intentional rounding was implemented, based on each patient having a rounding sheet displayed outside of their room, with each hour and staff member that visited visible, to allow easy visualisation and communication of when the patient had last been seen by a staff member. The system was then modified iteratively over successive weeks based on small samples of PDCA data collected during the intervention phase (*Table 24*). As *Table 24* shows, the first cycle results suggested a need for further staff training and a set of prompt cards to act as a template for a SOP on rounds. The second cycle indicated that an area log

TABLE 24 Plan–do–check–act cycles

Cycle	Changes introduced
1	<ul style="list-style-type: none"> ● Rounding sheet introduced with stamps for staff to record their round ● Staff training provided in intentional rounding via drop-in sessions with video and researcher ● Trialled on a 26-bed area of the ward
2	<ul style="list-style-type: none"> ● Further training provided via drop-in sessions ● Prompt cards introduced (clear set of tasks for rounding) ● Trialled on an 11-bed area and a 17-bed area of the ward
3	<ul style="list-style-type: none"> ● New intentional rounding sheet introduced ● Rounding sheet per ward area rather than per patient (except for high-care area) with a tick used for documentation not a stamp ● Alternating rounding hours for nurses and care support workers ● Trialled on the entire 75-bed ward
4	<ul style="list-style-type: none"> ● Staff provided with the choice of the individual rounding sheet or a ward area rounding sheet (see <i>Appendices 1 and 2</i>) ● Staff decided to use individual patient rounding sheets for high-care area, but to use area rounding sheets for the rest of the ward ● Trialled on the entire 75-bed ward

Source: reproduced from Morgan L, Flynn L, Robertson E, New S, Forde-Johnston C, McCulloch P. Intentional rounding: a staff led quality improvement intervention in the prevention of patient falls. *J Clin Nurs* in press; 2016,⁷⁵ with permission from John Wiley & Sons.

sheet may be more useful than individual log sheets and these were then introduced. This area log sheet was further modified in the third cycle, at which point it was decided to alternate between rounds by registered nurses and clinical support workers, allowing more efficient use of staff time. At the same point, the system was expanded from the pilot area to the entire ward. During this cycle and the final cycle, there was experimentation around whether or not to deal with recording on an area or an individual basis, which was settled in the fourth and final cycle.

The PDCA data included staff feedback from surveys. Evaluation was based on a comparison of falls data (from Datix) before and after intentional rounding was implemented. Falls data from the rest of the hospital (38 wards) were used as a control, excluding those areas which were obviously inappropriate (e.g. neonatal). Pre- and post-intervention observational snapshots were also recorded for comparison. The snapshots were either 1.5 or 2 hours in length and were matched pre and post intervention based on room number and patient diagnosis to reduce bias. During these sessions a researcher recorded who visited the patient, for how long and the tasks completed during the visit. In the post-intervention observations they also recorded whether or not intentional rounding was completed, whether or not the rounding sheet was presented and whether or not the intentional rounding was documented.

Findings

A *t*-test demonstrated patients observed were visited more frequently post intervention than pre intervention ($p = 0.007$). The mean number of visits per hour was 1.47 before and 3.32 after the intervention (Figure 3).

Ward observations during the post-intervention phases found that intentional rounding was completed 100% of the time (Figure 4). During these observational snapshots the intentional rounding log sheets were present on the patient door for 90% of the patients observed, with the round documented 50% of the time.

A 50% decrease in the total incidence of inpatient falls on the intervention ward was observed post intervention (Figures 5 and 6). Falls decreased from 44 in the pre-intervention period to 22 in the post-intervention period. A *t*-test found this difference to be statistically significant ($p = 0.024$; Table 25). Data from the rest of the trust demonstrated an increase of 3.28% in total incidence of falls between pre and post intervention. This difference was not significant ($p = 0.748$).

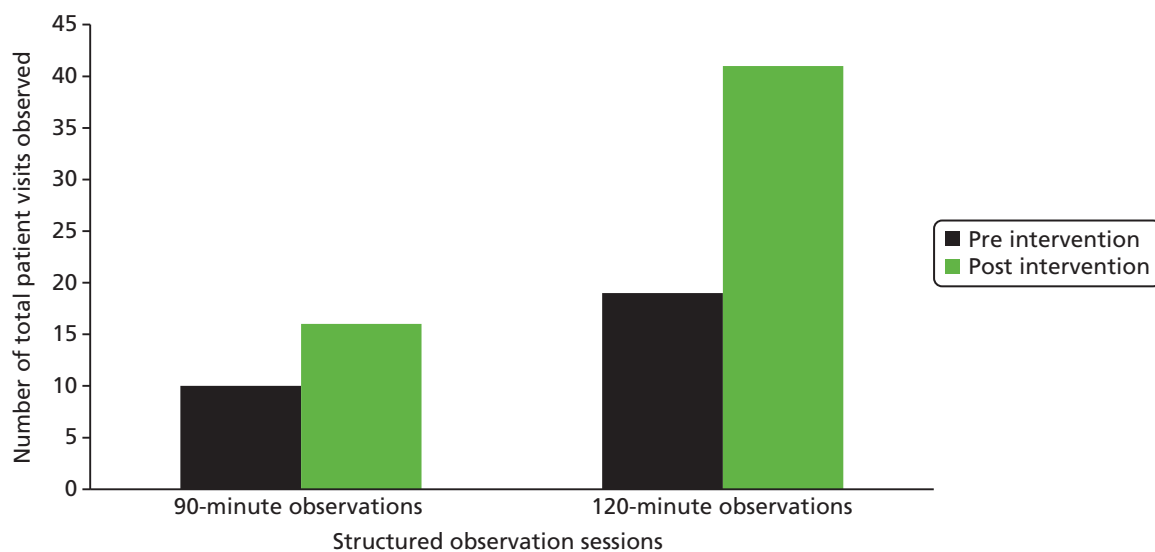


FIGURE 3 Nurse/care support workers visits to patients observed. Source: reproduced from Morgan L, Flynn L, Robertson E, New S, Forde-Johnston C, McCulloch P. Intentional rounding: a staff led quality improvement intervention in the prevention of patient falls. *J Clin Nurs* in press; 2016,⁷⁵ with permission from John Wiley & Sons. This article is protected by copyright. All rights reserved.

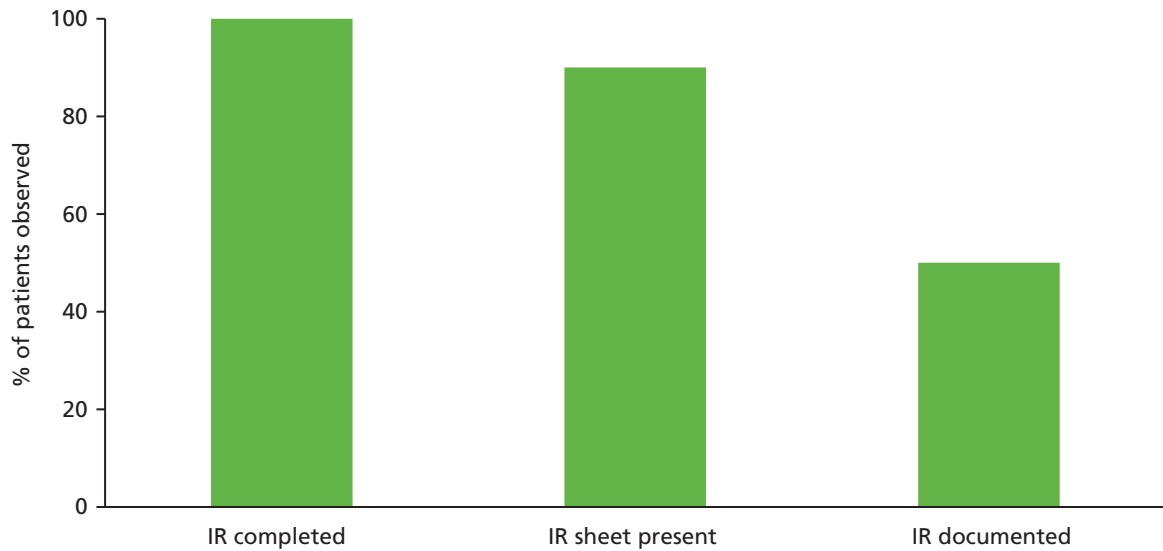


FIGURE 4 Completion and documentation of intentional rounding. IR, intentional rounding. Source: reproduced from Morgan L, Flynn L, Robertson E, New S, Forde-Johnston C, McCulloch P. Intentional rounding: a staff led quality improvement intervention in the prevention of patient falls. *J Clin Nurs* in press; 2016,⁷⁵ with permission from John Wiley & Sons.

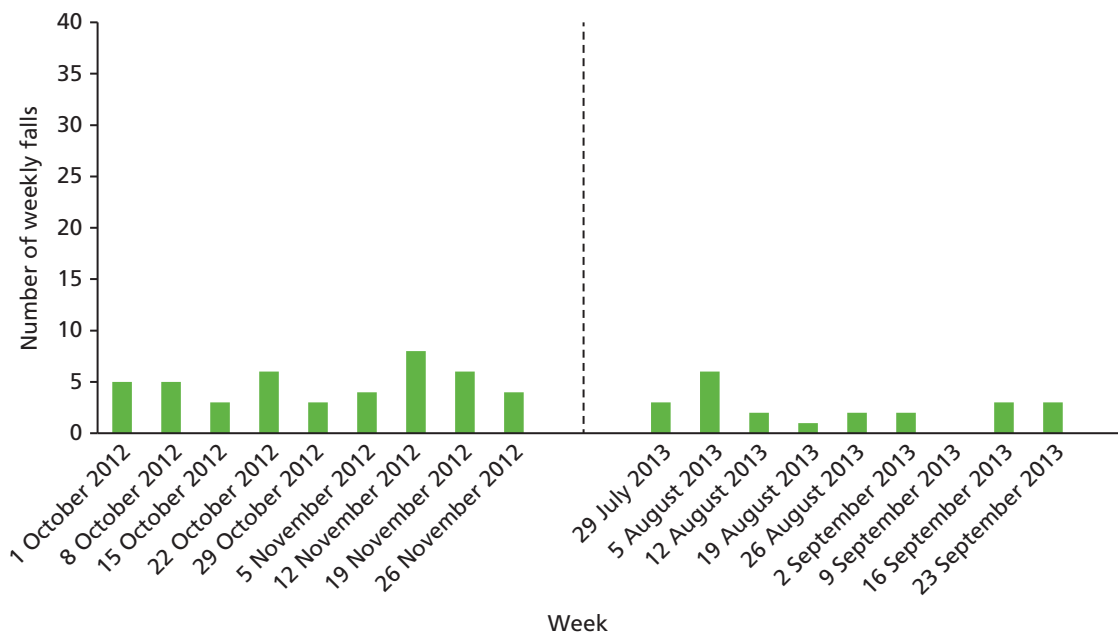


FIGURE 5 Weekly incidence of falls on the intervention ward. Source: reproduced from Morgan L, Flynn L, Robertson E, New S, Forde-Johnston C, McCulloch P. Intentional rounding: a staff led quality improvement intervention in the prevention of patient falls. *J Clin Nurs* in press; 2016,⁷⁵ with permission from John Wiley & Sons.

Comment

This project was a good example of the positive effects of an integrated approach to patient safety intervention, using both cultural and systems improvement. The project was relatively straightforward in that the aims were clear-cut and limited, and implementing the change did not require outside help, co-operation or resources.

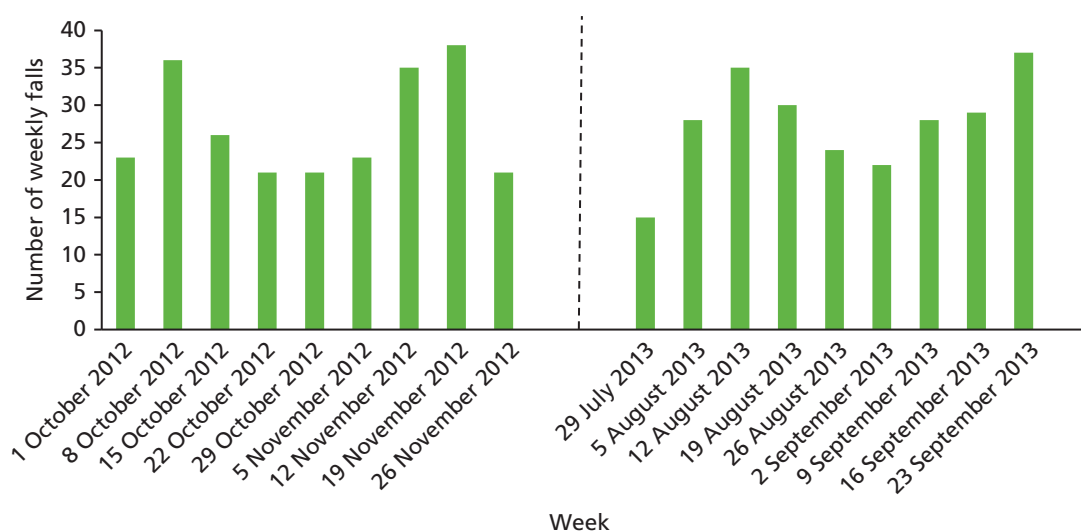


FIGURE 6 Weekly incidence of falls on the control wards ($n = 38$ wards). Source: reproduced from Morgan L, Flynn L, Robertson E, New S, Forde-Johnston C, McCulloch P. Intentional rounding: a staff led quality improvement intervention in the prevention of patient falls. *J Clin Nurs* in press; 2016,⁷⁵ with permission from John Wiley & Sons.

TABLE 25 Total number of falls per intervention

Wards	Pre intervention	Post intervention	% difference
Intervention ($n = 1$)	44	22	-50
Control ($n = 38$)	244	252	3.28

Secondary intervention: ward rounds project

Background

Frontline staff suggested that their ward round was inconsistent and wished to make it more systematic, reliable and safer. The ward round system in place at the start of the project was traditional and dysfunctional. The system was based on two teams of surgeons, derived from a historical standpoint when there were only two consultant neurosurgeons. Separation of patients into these teams was arbitrary and unhelpful. Communication between the wider MDT, including the nursing staff looking after the patients, tended to be poor. A focus group of multidisciplinary staff summarised the key existing issues with their ward rounds as being poor communication, poor filing of notes and lack of a clear plan.

Methods

The project was carried out across three phases: pre intervention (3 months), intervention (8 months) and post intervention (3 months).

The pre-intervention phase consisted of project planning and initial data collection. The data collected included observational data recorded during ward rounds and the results of a case notes audit. There is little evidence-based guidance on the properties of the ideal ward round and the best-known model is based on internal medicine ward rounds. It was felt that this model might need significant modification. An observational data collection pro forma for ward rounds was therefore created based on ergonomic theory about efficient transfer of task-essential information and clinical experience (see *Appendix 3*) and data were collected to help define what staff defined as the surgical ward round. The data collection pro forma retained many of the items in the ward safety checklist⁷⁶ and also included a glitch count as a measure of reliability. Other features of the round measured included staff and patient presence/input during the ward round, whether or not there was explicit consideration of a plan, whether or not charts/

folders were present, whether or not these were reviewed, and what recordings were made of observations and decisions during the round. Completion of specified clinical tasks was also recorded, including specifically whether or not medications were reviewed and whether or not a neurological or wound examination was completed.

Following baseline data collection, analysis and discussion with the clinical staff, changes were implemented and developed via PDCA cycles in the intervention phase. Staff decided to introduce three changes to improve the ward round process. Staff felt time could be saved and information could be communicated in a more structured, reliable way if a thorough process was instituted to hand over information to the new staff team at the beginning of the working day. This would clarify which patients needed to be seen, which team was responsible for them and the plan for the day (e.g. discharge). The first change was the introduction of a 10-minute pre-ward-round morning meeting. The second change focused on the storing of medical notes. The physical separation of nursing and medical notes in separate folders held in different locations appeared to be an obvious barrier to effective communication. The third and final change was focused on the structure of the actual ward round. To standardise work on the ward round and to ensure that all important observations and decisions were made and properly disseminated, a ward round pro forma was designed. Each page represented the visit of the team to the patient that day and included recording of observations and tests, examination findings, diagnosis/assessment, medication and treatment reviews and the plan for the next 24 hours. These three modifications were developed iteratively based on data collected in small audits as part of the PDCA process and staff feedback provided. These changes can be seen in *Table 26*.

Post-intervention observational data collection used the data collection pro forma designed and used for baseline observation. This was to allow for evaluation based on a comparison of data before and after the three changes were implemented. Observational data were also collected during the morning meeting post intervention; no pre-intervention counterpart could be collected here, as the meeting was itself an innovation introduced as part of the intervention. Data collected at these meetings were timings, which staff were present, what patients were discussed and what information was discussed for each of the patients. A follow-up notes audit was also conducted in the post-intervention phase to compare the completeness and accuracy of documentation after introduction of the ward round pro forma with conventional case note documentation prior to the introduction of the pro forma.

TABLE 26 Ward round changes

Project	Aim	Change
Morning meeting	<ul style="list-style-type: none"> • Effective communication of key patient information between MDT • Structured handover early in day to ensure key patient information prioritised 	<ul style="list-style-type: none"> • 10-minute morning meeting pre ward round with MDT • Standardisation of information communicated: diagnosis, theatre plans, discharge plans and therapy needs
Novel filing system	<ul style="list-style-type: none"> • Ease task of documentation • Reduce risk of missing notes • Increase access to doctors notes for other staff 	<ul style="list-style-type: none"> • Moved clinical notes from patient notes to nursing binder (end of patient's bed)^a
Ward round pro forma	<ul style="list-style-type: none"> • Improve communication and quality of information documented from ward round 	<ul style="list-style-type: none"> • New ward round pro forma (see <i>Appendix 3</i>) including recording of observations and tests, examination findings, diagnosis/assessment, medication and treatment reviews and the plan for the next 24 hours

^a Although many staff felt it was beneficial, because of concerns regarding filing inconsistencies and patient confidentiality the new filing system was discontinued.

Findings

We collected data from eight meetings over a 3-week period. Post-intervention observations at the morning meeting revealed that a median of 100% of patients were identified and 98% of neurosurgical patients were discussed appropriately in accordance with the SOP agreed in advance (name, diagnosis, theatre plans, discharge plans and therapy needs). The median attendance at the meeting was six doctors, 17 nurses, one sister/charge, one occupational therapist and one physiotherapist. The average time taken for the meeting was 11 minutes 38 seconds.

Comparison of pre-intervention ($n = 16$ rounds, 220 patients) and post-intervention ($n = 22$ rounds, 249 patients) ward rounds demonstrated improvements across a number of measures of the quality of the ward round. There was a substantial increase in the completeness of documentation after the introduction of the pro forma compared with the pre-intervention situation (Figure 7).

When the detail of the documentation during the ward round was reviewed, the greatest improvement was observed in the review and recording of information in the medical case notes (Figure 8); the review of information in the notes increased from 11.36% of cases to 49%, which was statistically significant ($p = 0.000$), while recording of new information increased significantly from 27.73% to 68.27% of cases ($p = 0.000$). There were smaller improvements in the availability and use of the observations charts and drug charts.

Statistically significant differences in the attendance (presence) of doctors [registrar/fellow, $p = 0.001$; senior house officer (SHO)/foundation doctor year 1, $p = 0.006$], nurses ($p = 0.022$) and patients ($p = 0.006$) were observed on rounds post intervention (Figure 9).

We prespecified a number of key data items that we expected would be considered when reviewing any post-operative neurosurgery patient (Figure 10). Results showed a low frequency of overt discussion or decision-making on these items at baseline and no major improvement after the intervention.

A reduction in a number of glitches (described using the 'glitch count' methodology developed in the earlier part of the programme) or deviation categories from the ward round process was observed post intervention (Figure 11). Statistically significant differences were found for concurrent conversations ($p = 0.000$), concurrent tasks ($p = 0.000$), distractions ($p = 0.000$), multitasking ($p = 0.000$), searching for people ($p = 0.000$) and any deviation ($p = 0.000$).

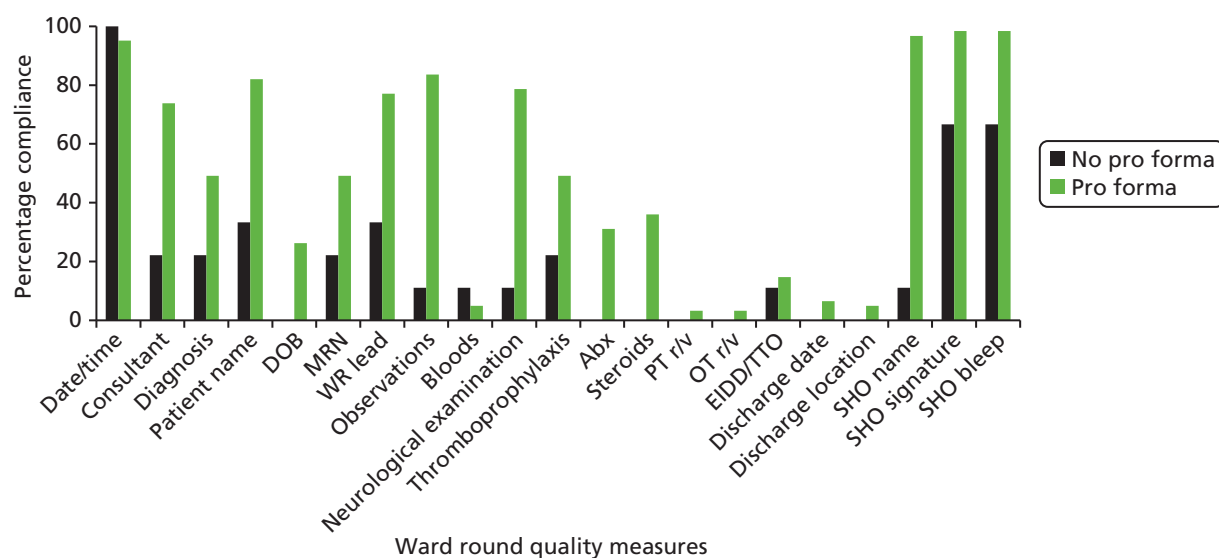


FIGURE 7 Ward round documentation with/without pro forma. ABx, antibiotics; DOB, date of birth; eIDD, electronic inpatient discharge documentation; MRN, medical records number; OT r/v, occupational therapist review; PT r/v, physical therapist review; SHO, senior house officer; TTO, to take out; WR, ward round.

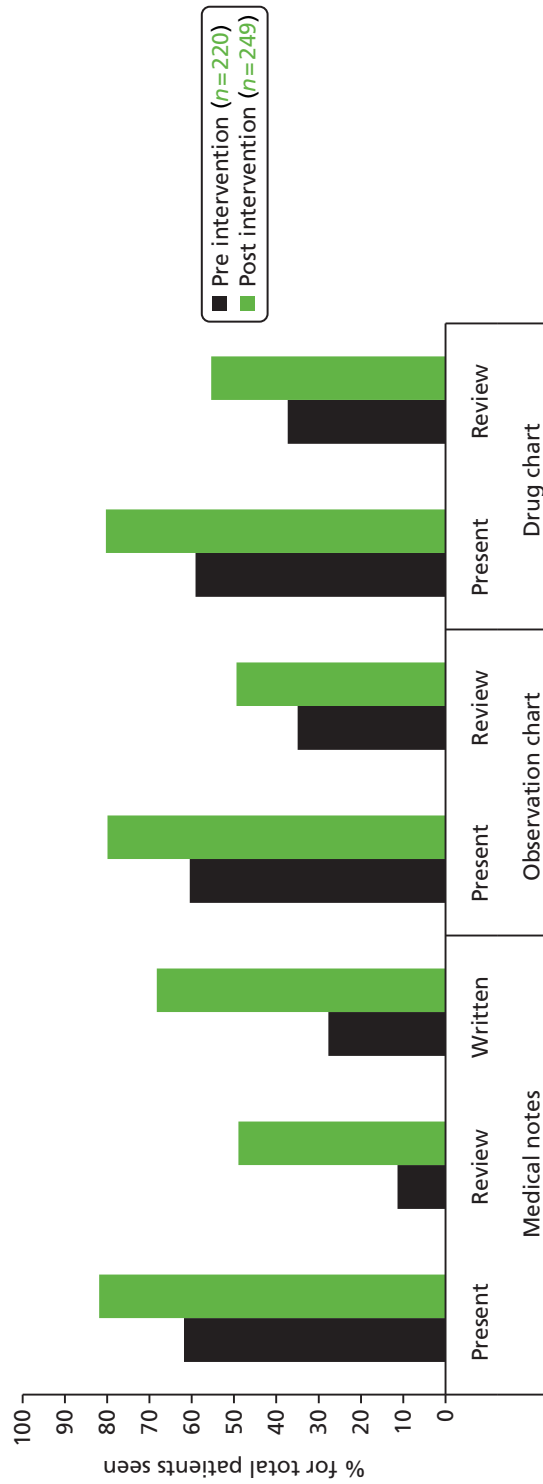


FIGURE 8 Documentation on ward round.

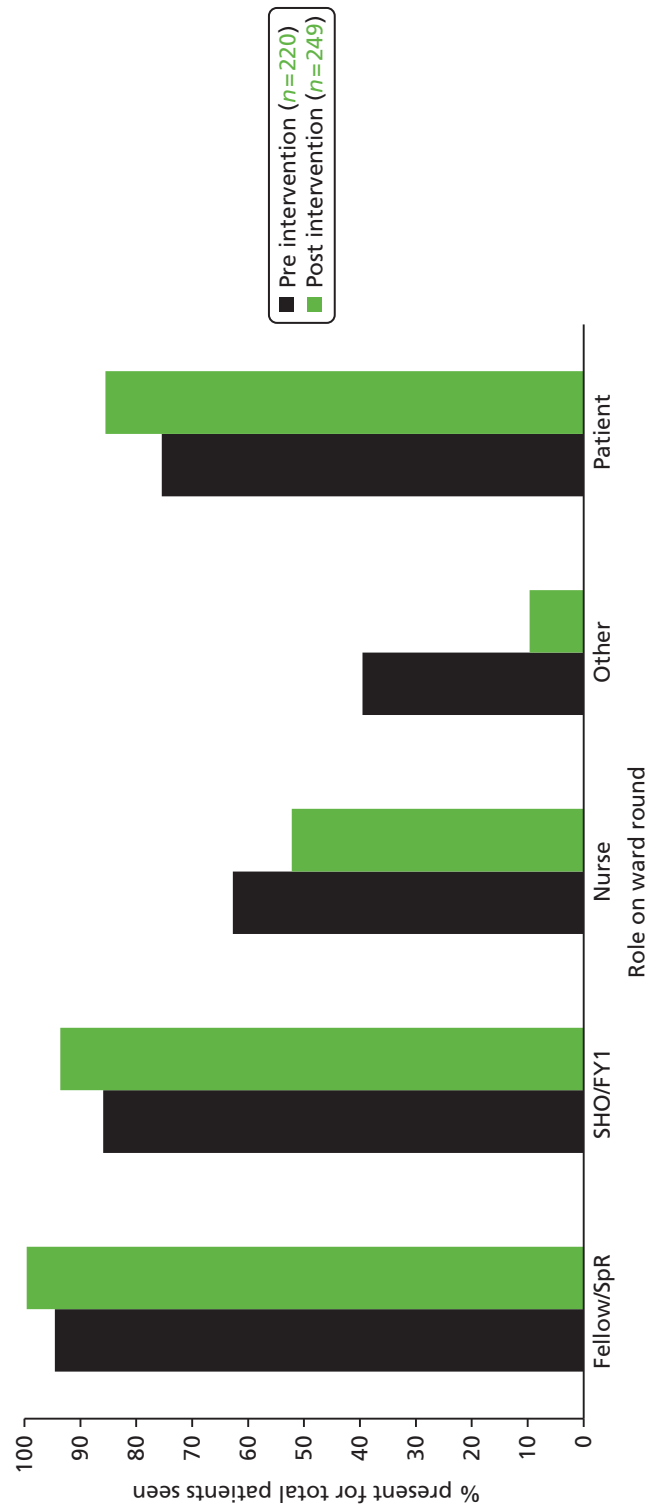


FIGURE 9 Roles present during ward round. FY1, foundation doctor year 1, SpR, specialist registrar.

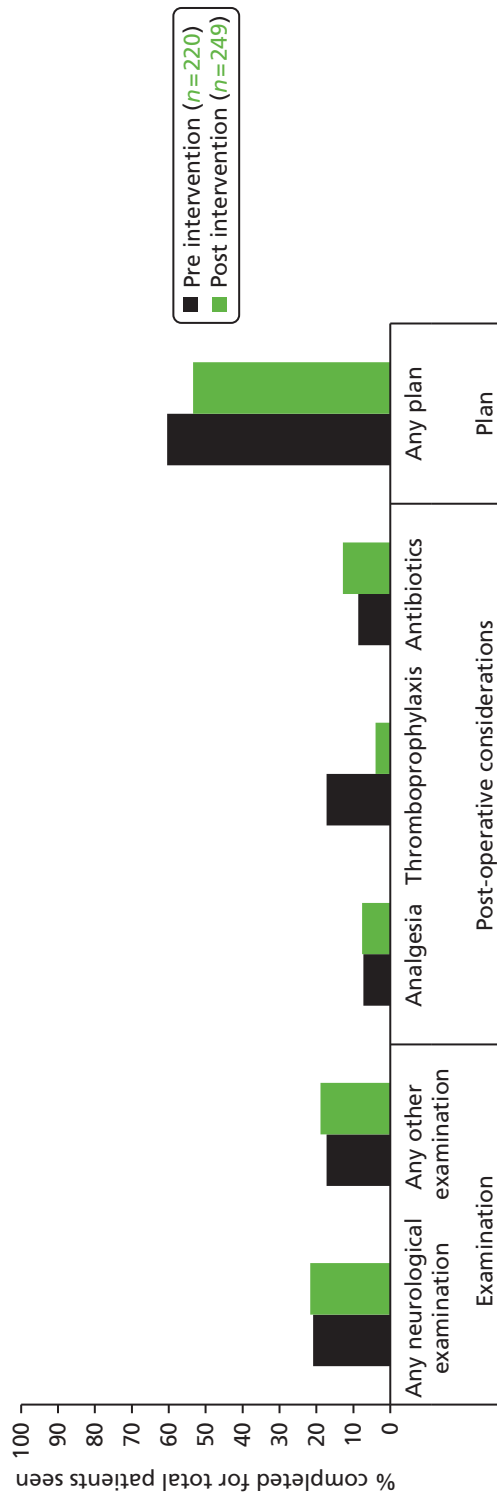


FIGURE 10 Post-operative consideration during ward round.

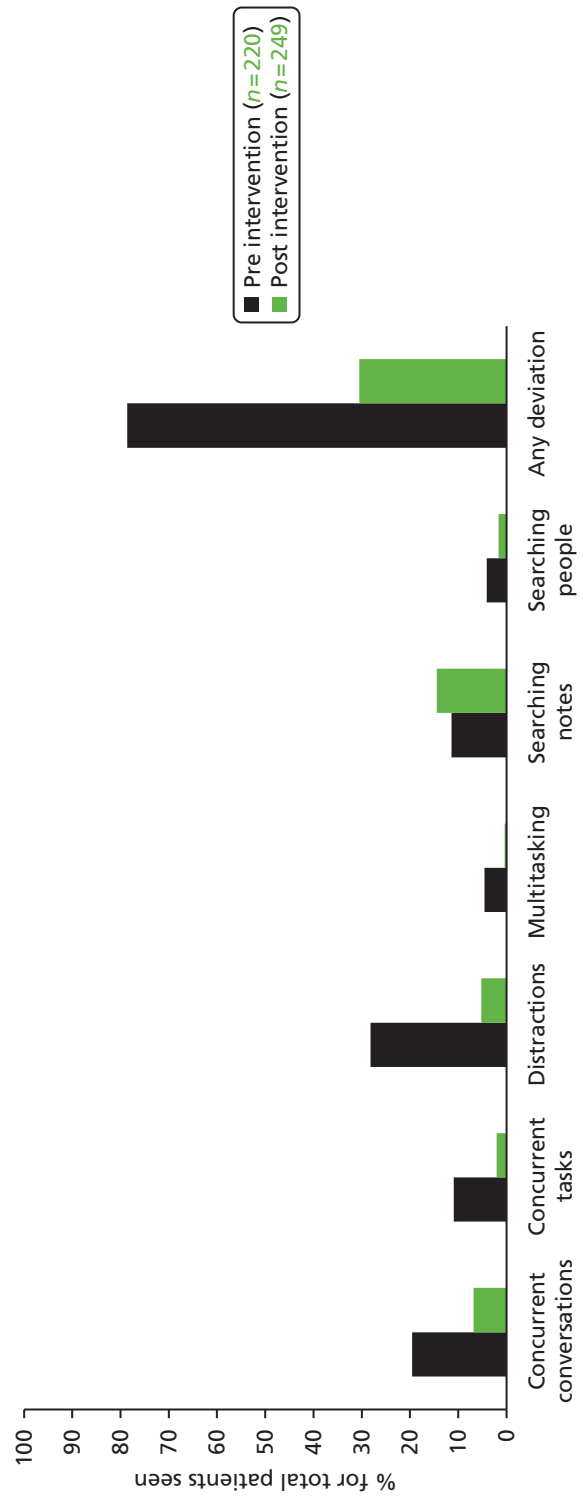


FIGURE 11 Glitches observed during ward round.

The median number of patients observed pre and post intervention were similar (13 and 11.5, respectively). The ward round start time did not change; however, the duration of the ward round decreased from a median of 1 hour 5 minutes to 38 minutes. The time spent per patient decreased from 4 minutes 45 seconds to 4 minutes 19 seconds.

Comment

Our intervention led to major improvements in some aspects of ward round performance but made no impact on others. In some respects this heterogeneous pattern was intuitive. Specific actions were taken to improve data recording and review, and this duly improved. There was no focus on the actions we thought appropriate for post-operative care (see *Figure 10*) and we observed little improvement in their behaviour. The early-morning meeting proved extremely successful in improving the transfer of information between shifts and it was felt by staff to have contributed importantly to reducing the time spent on rounds.

An important barrier to greater progress in this study was reluctance to change established practice on the part of senior medical staff. Consultant concerns prevented the unification of nursing and medical case notes, which would probably have had a number of benefits in terms of allowing access to full information about patients. We would have preferred to involve the consultants more in the process of redesign, which might have avoided these two issues, but found them difficult to engage.

Secondary intervention: referrals project

Background

The system used for patient referrals into the department was an area identified for improvement by frontline staff in the neurosciences service. Staff identified patient referral as a weak point in the care pathway, subjecting patients to potentially poor-quality care and inefficient use of resources. The unit receives approximately 20–30 new acute patient referrals every day. These referrals are mainly for urgent or emergency care of serious conditions, which require time-urgent surgical or radiological management. Each referral is made by telephone and the details are manually written on a blue card. It was expected that the referring clinician should record the advice they are given by the unit team in their own health-care records. No facility existed within electronic patient record (EPR) systems to document and manage this process. The paper-based system in place was deemed entirely unfit for purpose as it is inefficient, inaccessible and error-prone and poses barriers to audit, clinical governance and QI. There was therefore a significant risk of inappropriate delay in accepting referrals because of difficulty in properly assessing patients using the information available via telephone referral systems.

Objective

The objective is to improve the reliability and timeliness of referrals in order to improve patient safety. Staff proposed the development of an electronic system that will improve and document the communication of referral information between referring hospitals and specialist centres. In addition, the system will have the function to link referrals to the EPR. The system will be deployed to district general hospitals and, in the first instance, the neuroscience specialist centre in the John Radcliffe Hospital.

Methods

The project was initiated by developing the idea and communicating the concept with key stakeholders including clinicians, managers and hospital information management and technology staff.

Getting support for the change required negotiation and discussion at multiple levels and with many stakeholders. These included consultants, senior nursing staff, junior surgical staff, key individuals at some of the main referring hospitals and those responsible for IT, clinical governance and ultimately senior trust management who could allocate funding. For this reason the change could not be implemented within the time scale of the research programme but is continuing.

Following this initial phase of gaining support, a task/process analysis was undertaken with clinician involvement to understand and document the current process for making referrals, the decision points and the roles involved.

Then, based on a set of user stories, an initial specification was drawn up and the development of user-interface mock-ups was contracted to a software company. The development of the mock-ups involved two formal user evaluation sessions and was conducted in close co-operation with the John Radcliffe Hospital project team.

A more complete specification has been drawn up and additional analysis undertaken to define the status of referrals and how this relates to clinical decision-making. The project is currently under way with a second phase of software development to produce a functional web-based application for deployment later in the year.

Progress and findings

The new system will document and manage the referrals process. Referring clinicians will use a secure system to submit their referrals and view the responses from specialists, entering the key information which specialists need for decision-making in a carefully designed and tested template. A record of the entire referral process will be visible to all relevant parties, reducing the potential for communication errors and information loss.

Clinicians and managers will not be the only specified end users of the system. Built into the proposed design is the ability to provide information on the referral to the patient, family and/or carers. Patients/relatives/carers will receive standard information on the likely interaction with the trust (through both signposting to current information sources and to new information when relevant) and also condition-specific information.

An improved referrals process will result in higher-quality care through timely and better-structured specialist advice; allowing specialties to manage resources better by prioritising and planning patient transfers using objective information. We expect this will lead to improved patient outcomes including reduced morbidity and increased patient and clinician satisfaction.

A data plan has been drafted to identify metrics to measure the intervention and to determine feasible methods for data collection to include outcome (such as length of stay) and process (such as duration of referral process) measures.

Access to structured data also facilitates ongoing measurement that allows continuous process improvement. Using this referral platform will expand the high-quality care provided to inpatients during the referral process. Whether or not this succeeds depends partly on the ability of the clinicians who have learned QI principles through their interaction with us to engage with senior management and persuade them to give this the appropriate priority.

Secondary intervention: subarachnoid haemorrhage care project

Background

Subarachnoid haemorrhage is one of the more common clinical emergencies dealt with by neurosurgery units. Originating from rupture of small aneurysms in the arteries at the base of the brain, they cause very sudden severe headache, loss of consciousness and residual neurological deficits. Mortality is relatively high, but prompt treatment to prevent further haemorrhage can reduce this. Most of these aneurysms are now treated by neuroradiologists rather than neurosurgeons. As these doctors are principally regarded as diagnostics specialists, they have no ward or area of their own for the patients they treat. By convention and tradition, the patients they treat are thereafter cared for by the neurosurgery unit staff, but ultimate responsibility for their care remains unclear and unsatisfactory. Following primary treatment, there is a critical period of a few days or weeks in which secondary complications are common, particularly rebleeding and severe spasm of the cerebral arteries. Post-intervention care is focused on preventing these two risks. A group of multidisciplinary staff believed that the care of patients with SAH could be improved. Poor communication of information regarding such patients was identified as a key issue within this area. Staff also believed that the care among this group of patients was inconsistent.

Methods

In the pre-intervention phase, a case notes audit of 50 patients was conducted using a standardised template (see *Appendix 4*), which recorded key items of data about referral, intervention and post-intervention care. This showed that compliance with the main post-operative care practices considered to be risk mitigators for rebleed or delayed cerebral ischaemia was very variable and not satisfactory overall (*Table 27*).

The care pathway of a SAH patient was process mapped by interviewing staff combined with direct observation. The process map (see *Appendix 5*) demonstrates the complexity of the process. Based on the data collected, the project team decided to develop a SAH pro forma. The aim of this was to drive a process of standardisation of post-intervention care (see *Appendix 6*). This was designed in collaboration with the MDT including nursing staff, the neurosurgeons and the interventional radiologists. The aim of the pro forma was to promote compliance with care guidance for SAH patients. It was made up of four sections: (1) a cover sheet with basic patient information; (2) a clerking section for patient history, initial examination, assessment, plan and consultant review; (3) a radiology section for aneurysm details, operative/procedure notes and post-operative instructions; and (4) a daily review section, with 12 identical pages for ward round assessments and other clinical notes. The intention was for it to be completed by the MDT including surgeons, radiologists and nursing and allied health-care professionals. This pro forma was implemented during the intervention phase. It was developed using PDCA cycles and staff feedback to guide each iteration.

A follow-up audit in the post-intervention phase was planned, to allow comparison of compliance to standards of care.

Findings

The pro forma was adopted rather inconsistently during the intervention period, being used for only 7 out of 25 patients followed post intervention; in two cases it was inserted in the notes but left blank, and in remaining cases it was not in the patients' notes at all. Of the seven cases in which it was used, only the clerking section was completed in five cases, and only the daily review section in one case; in only one case were both sections completed. When the daily review was used, clinical notes had also been made separately in the patients' notes, with no consistency in how the notes were being entered.

Based on the inconsistent use of the pro forma, a follow-up audit beyond the intervention period was not completed. We judged that an impact on clinical outcomes was implausible given the poor uptake of the intervention, and that any differences between pre- and post-intervention states would therefore be more likely to be because of chance or secular trends than as a result of the intervention. A follow-up case study with ethnographic observation and interviews with staff (JH) identified some reasons for the pro forma being discontinued. There was confusion over patient responsibility between neurosurgeons and interventional radiologists, as previously noted. This affected attitudes to responsibility for completing the pro forma. When junior staff were not clear whether or not their seniors supported the introduction of the pro forma they were understandably reluctant to use it. The issues previously described about the location

TABLE 27 Audit results

Standard of care	Compliance
Neurosurgical consultant review	Occurring 11% of days
Interventional radiology review	Occurring 9% of days
3 l of fluids daily	Not completed for 67% of days
4-hourly nimodipine	Not completed for 12% of patients
4-hourly observations	Not completed for 12% of patients, increasing to 25% over days 5–10 when risk for delayed cerebral ischaemia is greatest

of the notes (nursing binders vs. patient notes) impacted on this project, since the pro forma properly belonged in the medical notes but described a set of tasks of which some were performed by nurses. The lack of a formal policy to back up pro forma compliance or of a clear clinical leadership strongly suggesting that it should be used meant that junior staff regarded it as a nuisance rather than an important adjunct to care which could dramatically improve the gap between intended and delivered care.

Secondary intervention: morbidity and mortality meetings project

Background

Morbidity and mortality meetings (MMMs) are designed to allow review of adverse outcomes with a view to learning and systems improvement. They should play a vital role in audit, allowing the identification and analysis of problems and the development of plans for further investigation and practice change to complete audit cycles. However, in practice, such cycles are rarely completed because staff have numerous other priorities, no allocated time or training to allow them to institute a change programme, no institutional incentives to do so and no resources. Although the meetings were supposed to be attended by all neurosurgical consultants, specialty and foundation trainees, attendance at the neurosurgery MMMs was low. Staff believed that the meeting lacked structure and documentation and had no feedback mechanism for deciding on plans for change and reviewing their results. Although some cases were discussed, no actions or follow-up plans were usually completed.

Methods

This project was conducted across four phases: planning (1 month), pre intervention (1 month), intervention (1 month) and post intervention (1 month). As no standardised guidelines or measurement system currently exist for MMMs, frontline staff used the planning phase to design a set of guidelines for the ideal MMM. These were based on a review of literature, with input from a clinical governance representative. Areas identified for targeting included (1) improving the reliability of case reporting associated with the meeting; (2) improving the reliability of workload data based on admissions and theatre numbers, which at baseline were collected from a diary and log; and (3) improving documentation of learning points highlighted at meetings and instituting a review process to determine whether or not any progress was subsequently made as a result of MMM discussions. These objectives were used in conjunction with available guidelines to create a driver diagram (*Figure 12*).

The planning phase was also used to determine the primary and secondary outcome measures to be used, based on the guidelines and the driver diagram (*Box 2*). Documentation of learning points and action plans (when appropriate) for each case presented at MMMs were the primary outcome measures. Secondary outcome measures included (1) change in practice planned; (2) individual-assigned responsibility; and (3) planned method of action plan. Other additional measures included (1) start/finish times; (2) whether or not an attendance register was completed; (3) whether or not minutes were taken; (4) whether or not there was a consultant chairperson; and (5) whether or not there was a review of previous action plans.

Using the driver diagram, staff-led changes were introduced to the MMM during the intervention phase via PDCA cycles. The changes included (1) the introduction of a group e-mail and distribution list; (2) the creation of a definition for cases to be included in the meeting; (3) the introduction of a MMM monitor role which would be rostered so that each member of the junior staff was allocated to perform it in turn (the monitor was responsible for the identification of cases for discussion using the definition, and for MMM communications; for example, ensuring the meeting was organised, disseminating feedback); (4) the creation of a MMM database, recording the data presented at each meeting, the learning points agreed and the decisions about changes to unit processes and protocols; (5) the introduction of a consultant chairperson for every meeting; and (6) the development of IT systems to allow accurate admissions data collection and presentation.

Post-intervention observations were conducted using the data collection pro forma to allow for comparison of pre- and post-intervention data for evaluation purposes.

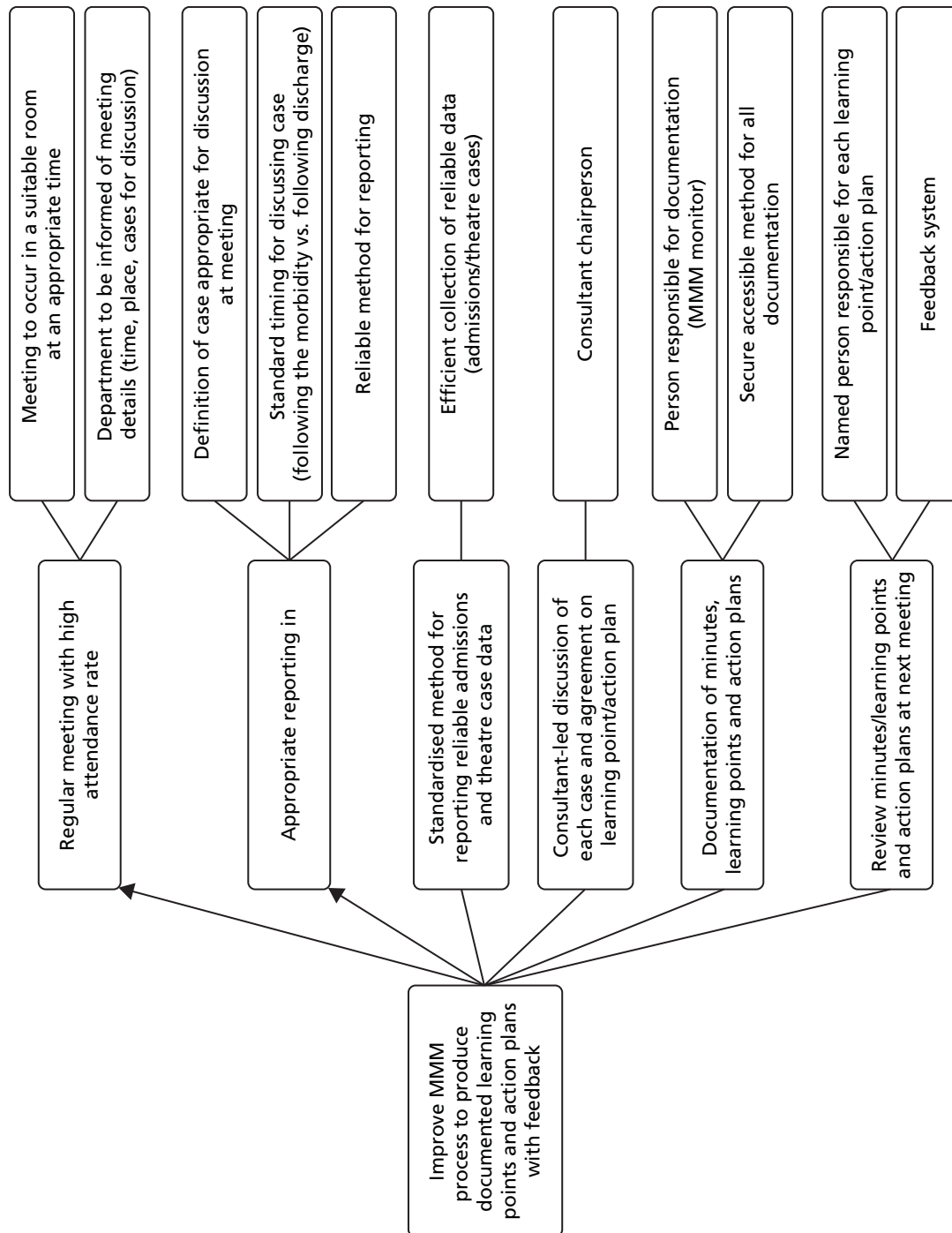


FIGURE 12 Driver diagram for MMMs.

BOX 2 Morbidity and mortality meeting outcome measures**Primary outcome measures**

Documentation of learning point and action plan (when appropriate) for each case presented at MMMs.

Secondary outcome measures

- Change in practice planned.
- Individual assigned responsibility.
- Planned method of action plan.
- Start and finish time.
- Attendance register completed.
- Minutes taken.
- Consultant chairperson.
- Review of previous action plans.

Findings

A number of simple but obvious improvements were noted in the MMM process after the intervention. The attendance register was completed for 100% of cases post intervention compared with 50% pre intervention. A consultant chairperson was observed for 100% of meetings post intervention and minutes were completed for all meetings. Action types decided on following a MMM discussion also expanded to include literature reviews, practice changes and training sessions, whereas prior to the intervention most of the actions were simple reminders to staff to avoid a previously noted error. Methods for instituting change and for determining responsibility for instituting practice change were also determined more frequently (100% and 80%, respectively) in each of the MMMs observed post intervention. Many other improvements were observed against key aims from the driver diagram (*Table 28*).

Secondary interventions: minor/discontinued projects

A considerable number of other secondary interventions were initiated but were not developed into formal improvement projects with PDCA cycles. These are not described in detail, as they never achieved sufficient momentum to merit detailed evaluation. They were either more minor projects in which no formal QI process was envisaged from the outset but for which staff expressed enthusiasm for incorporating a desired change into the general efforts for widespread process improvement (*Table 29*) or the projects which were discontinued through loss of staff engagement, time or resources after an initial enthusiastic launch (*Table 30*). The improvements intended for many of these projects were difficult to measure or, in some cases, insufficient progress towards improvement was made to merit formal analysis of the small data sets acquired.

Discussion

This was the final intervention study of a 3-year programme and was intended to pilot test a new integrated intervention that incorporated learning from the previous five studies. By now the intervention team was experienced and well organised, and our theory of change had undergone major developments since the beginning of the programme as a result of our experiences. We engaged with a larger unit of clinical activity than in previous studies, effectively about 50% of an entire division of a large teaching hospital. In the event, our intervention led to the generation of a multiplicity of projects to improve aspects of care ranging from tracheostomy decannulation through ward rounds to same-day admission for elective surgery. The majority of projects failed fairly rapidly, only five led to sustained activities, and of these only one clearly led to measurable improvement in clinical care during the lifetime of the project. This has to be regarded as a fairly

TABLE 28 Morbidity and mortality post-intervention results

Aim (from driver diagram)	PDCA cycles	Outcome	Compliance
MMMs to occur in suitable room at appropriate time	MMMs already established as weekly meeting with room booked; attendance book introduced to monitor attendance and allow future review of suitability of meeting timing	Attendance record kept at each observed post-interventional MMM	4/4
Department to be informed of MMM details	Creation of hospital group e-mail account to use for all MMM communications, including informing department of meeting details. There was difficulty recalling all members of department to be e-mailed, therefore a distribution list was created in PDCA cycle 2 to improve efficiency	Group e-mail account and distribution list used in each post-interventional MMM	4/4
Definition of case appropriate for discussion at MMMs	Identify themes of morbidity by review of literature and cases discussed at previous MMM; themes incorporated onto database as checkboxes; themes altered in PDCA cycle 2	All cases presented during post-interventional period categorised into themes	16/16
Standard timing for discussing case	Departmental lead requested all cases to be discussed in the meeting following occurrence	Cases discussed contemporaneously in all post-interventional MMM	16/16
Reliable method for reporting into MMMs	Introduction of weekly 'MMM monitor' responsible for MMMs who identified cases and was responsible for the e-mail account (to communicate meeting details to department and confirm cases). The role was worked into the foundation and junior specialty trainee rota in PDCA cycle 2	MMM monitor took responsibility for each post-interventional MMM	4/4
Efficient collection of reliable data (admissions/theatre cases)	Development of IT systems for collection of hospital-level admissions and theatre case data; method of displaying hospital-level data improved in PDCA cycle 2	Use of hospital data in all post-interventional MMM	4/4
Consultant chairperson	Introduction of a consultant chairperson with responsibility for leading discussion, learning points and action plans; standardised system employed for which consultant to chairperson (i.e. on call consultant) in PDCA cycle 2	Consultant chairperson in each post-interventional MMM	4/4
Person responsible for documentation	Develop database to record and review all MMM data. MMM monitor responsible for recording cases onto database, presenting cases during MMMs and contemporaneous documentation of learning points, action plans and named person responsible for action plan. PDCA cycle 2 involved modification of database and storage organised on shared drive with access for all department members	Database used for each post-interventional MMM	4/4
Secure, accessible method for all documentation		Learning points and action plans documented contemporaneously	16/16
Named person responsible for each action plan		Person assigned responsibility for each action plan, when appropriate	4/5
Feedback system	MMM monitor responsible for disseminating summary of MMMs (e-mail and storage of database on shared drive); previous MMM to be reviewed at start of meeting. Modifications to review process in PDCA cycle 2 to present summary of previous cases, learning points and action plans. Projects developed as result of action plan to be presented at 3-monthly hospital audit meeting	Database stored on shared drive with access for all department Weekly review of ongoing work	4/4 1/4

TABLE 29 Minor secondary interventions

Project	Background	Methods	Findings
Patient follow-ups	Staff believed that some follow-ups could be conducted by telephone, reducing the number of follow-up appointments required	A protocol for patient follow-up by telephone was developed and a number of patient follow-ups were conducted by telephone. A satisfaction questionnaire was undertaken by these patients	A majority of patients preferred the telephone appointment to coming to the hospital. Only one patient felt it did not address all of their concerns
Equipment	Staff unaware of what equipment was present on the ward and where it was being stored	An equipment catalogue was produced by a health-care assistant on the ward	The impact was impossible to measure and demonstrate, and therefore no data were collected on this project
ITU outreach	Staff wished to create a business plan for an ITU outreach programme	Staff were supported in the development of a business plan	The aim of the work during S3 was to support the production of a business plan which was completed
Agency nursing	Staff believed that a lack of familiarity with the ward was an issue among agency nurses	Induction training for new agency staff was set up by a senior nurse	The impact of this project was also difficult to measure
Discharge	Staff wished to improve the discharge process	Increased capacity of the discharge lounge by introducing discharge chairs and screens. The time at which doctors completed TTOs was also altered to facilitate a more efficient process. A pharmacist was also provided on the ward to assist in the TTO writing to coincide with the ITU ward round to facilitate this	No post-intervention data were collected for this project

TTO, to take out.

TABLE 30 Discontinued secondary interventions

Project	Background	Methods	Findings
Emergency theatre	The staff wished to target the use of emergency theatre as area for improvement, as many elective cases were being cancelled to facilitate emergencies	Some initial analyses revealed that the number of emergency vs. elective cases was not reflected in the allocation of theatres	Competing agendas within the organisation resulted in this project being discontinued
Tracheostomy care	Care of patients with tracheostomy was an area identified for improvement	The aim of the project was to protocolise the 'cuff down' step in decannulation, which was found to be delayed in most patients, as they wait for a physiotherapist	The staff member who was leading on this project moved during the course of this project and so this work was discontinued
SSIs	Staff wished to reduce the rate of SSIs	Some initial analysis done, but nothing further than this	There were too many competing agendas within the organisation in terms of this project which resulted in it being discontinued
Admissions	Staff wished to improve the admissions process	N/A	This project was discontinued as a result of a lack of interest among staff after initial meetings
Staff rotation	Staff rotation was believed to be an issue on the ward, as it was happening too frequently. Staff wished to reduce the rotation	N/A	This project was discontinued because of issues with the off-duty system which was currently in place on the ward

N/A, not applicable; SSI, surgical site infection.

disappointing outcome, but analysis of the reasons for failure (or in some cases deferred success) reveals a number of very useful lessons, which may help future efforts to improve care quality and safety.

Intervention and implementation strategy

Our approach to intervention in this study was more complex than in the previous studies in two ways. First, we purposely provided training which integrated elements of all three intervention types we had previously used – the ergonomic approach to analysis, rationalisation and standardisation of work, the lean approach to eliminating waste from work process through engaging frontline staff and the CRM staff training approach to make individuals aware of threats to team function caused by unhelpful relationships and attitudes and failure to communicate clearly. Second, we developed and used a carefully thought-through approach to engagement and implementation that had been designed in response to difficulties experienced in previous projects. This involved (1) intensive and carefully structured communication about our intentions with all staff groups; (2) engagement with grass-roots staff and strenuous efforts to involve them in project initiation and design; (3) identification of key leaders within staff groups and attempts to motivate them to take leading roles in study projects; and (4) attempts to design interventions to maximise their potential for uptake by using the principles identified by Everett Rogers.⁷⁷ The key points that we felt were important to safety improvement from his work were to:

- make change attractive to staff by making it quick, simple and cheap (in terms of time and energy) to adopt
- make change attractive by identifying positive benefits for individual staff from adopting it (such as less stress or annoyance at work, simpler procedures, less misunderstanding and confusion, reduced effort and better patient satisfaction or outcomes)
- ensure change was compatible with staff attitudes and beliefs about the roles and behaviour of the different professions
- engage 'early adopters' quickly and use them as role models for other staff
- ensure that change was codesigned and could be modified by staff to suit local needs.

We made an attempt to use the principles of advertising psychology to present our message in the most persuasive fashion. The six principles outlined by Cialdini and Goldstein⁷⁸ (liking, authority, reciprocity, consistency, scarcity and social proof) were considered whenever we developed messages for the staff.

Limitations on success: investigator strategy

The final practical outcomes of these efforts were disappointing, as we have outlined, but there was evidence that it did enhance staff engagement in some areas.

One noticeable feature of this project which points to this conclusion was the strong relationship built up between the intervention team and the clinical staff they worked with. This was commented on by multiple staff members during the qualitative study we completed shortly after the end of the intervention (see *Chapter 10*) as a major motivating factor. Staff clearly felt supported and empowered by the presence of the research team members and commented on the importance of their presence to maintain morale and commitment.

Staff did identify a large number of areas for improvement, a number of which turned into viable projects. These projects were widespread and focused on very different areas within the department. We also saw a varied selection of multidisciplinary staff involved and engaged in frontline improvement, including doctors of varying seniority, nursing, care support workers, ward clerks, management and allied health-care professionals, for example physiotherapists. Some of the improvement work has been sustained and continued beyond the end of the project, and appears likely to result in important practical advances with benefits for patient safety, for example the Oxford Acute Referrals System and the ward rounds (morning meeting) project. These survivor projects tended to occur when individual staff members or small groups became motivated to succeed on their own behalf and seized opportunities presented by their position in the structure of the workplace that allowed them to make change in particular areas.

The description above might give the impression that the reasons for relatively limited success in this study are mysterious. We have mobilised significant, well-trained resources, have thought out our approach carefully based on previous experience and relevant theory and have succeeded in some measure in energising large sections of the staff.

Strategically, an obvious error was to create a situation in which too many improvement projects were simultaneously generated from the enthusiasm of staff attending our training days. Initially, over 20 project areas were suggested, most of which died off rapidly with little to no change achieved. We accepted all suggestions for improvement projects from staff, which could be clearly defined and justified, as we believed this was important to achieve staff buy-in. The multiple projects caused confusion and a diffusion of effort among staff and resulted in S3 team support becoming widely stretched. Staff may have become disengaged in some of the projects because of the lack of progress resulting from this attempt to do too much too quickly.

In hindsight, it would clearly have been better to limit the number of areas targeted. At the time, and influenced by the 'bottom-up' philosophy, which imbued the particular type of lean intervention we employed, we were concerned that taking this approach might have inhibited some staff from becoming involved or may have alienated staff by not targeting an area of their choice. It is, however, worth noting that highly successful interventions (such as that conducted in Michigan,⁵² the SURgical Patient Safety System (SURPASS) study⁷⁹ and a number of projects run by the Institute for Healthcare Improvement organisation) have used staff input as only one of a number of ways of generating projects. These groups have tended to take a more strategic approach to option selection, beginning with the development of a construct of how an ideal system would work based on literature evidence about which components of therapeutic processes are effective. They then codesign a strictly limited number of interventions targeted on those aspects of the system they are working with which seem most in need of change according to baseline data or which are most likely to respond rapidly and dramatically to intervention. It is not clear whether or not these groups experienced any alienation from staff because they were not in complete control of deciding what changes were made, but if so they seem to have compensated for this sufficiently to produce measurable benefits for patients.

Limitations on success: culture and incentives for key staff

A second strategic failing was the failure to adequately engage with senior management and senior clinical staff. This meant that projects had limited official backing and no resources were available from the trust except on a very limited unofficial ad hoc basis. As a result, many projects came up against immovable barriers that caused their collapse. This failing was, however, not an error entirely of our making. We did in fact make strenuous efforts to inform and involve the divisional and directorate managers and the clinical directors at both levels, as well as nursing and theatre management. We established excellent working relationships on specific projects with the divisional manager, the nursing management generally, and with the clinical director for ITU and anaesthetic services. However, we discovered that the diffusion of responsibility for executive action and control in the trust was such that these individuals were not able to ensure that our schemes could be implemented. It was difficult to mobilise resources when they were needed (especially staff time) and managers had to balance the potential of our projects against any disruption that they threatened to the routine running of the unit.

With regard to senior medical staff, we experienced a difficulty in achieving real engagement. The demographic of this was interesting. Doctors who had taken senior management roles in specialties other than surgery, senior surgical trainees and the most junior of the surgical and anaesthetic consultant staff were interested in our work and willing to get involved, but the bulk of the senior surgical staff did not see it as particularly relevant or interesting to them. There was little evidence of active opposition, except when projects directly impacted on consultants' routines. This led to the cancellation of the notes aspect of the ward rounds project and to some difficulties for the junior surgeons involved in the project to improve the MMM.

Other external factors

Other weaknesses in our approach to the intervention should also be considered as potentially important in limiting our success. For reasons related to research funding cycles, we chose to conduct intervention projects which required limited staff involvement and to set ambitious time scales for completion and evaluation. Many practitioners of ergonomics and industrial QI, including those working in health care, take the view that major change cannot be achieved on these time scales, and that it requires persistent effort supported by management over a considerably longer period. The same authorities often take the view that improvement within a division or department of a larger organisation is difficult to achieve without fundamental change in the approach of the whole organisation. We accept that these assertions may be true, but by their nature they cannot be subjected to experimental proof. There is ample evidence that small-scale and short-term interventions can have effects, and in order to understand the mechanisms by which systems and culture interventions improve safety work, small-scale experimental studies of the kind we have conducted here are both justified and necessary.

Of the difficulties imposed upon us by the environment, one of the largest was the challenge of arranging to have all involved staff meet at regular intervals to work on a project. Different staff groups are trained separately and, although they may work together in teams on wards and in theatre, they are managed according to their professional hierarchies in systems largely exclusive of each other. This, together with the structure of their work, means there is little opportunity for them to become familiar or train with each other. It proved very difficult logistically to find times that were convenient for all members of a project team to meet and as a result there were almost always some members of a project group absent. A related issue was the difficulty in communicating information about the projects and changes across all staff because of the number of staff on the ward, the rotating nature of shift work and junior doctor training schedules, with the resulting constant flux of staff members. The difficulties posed by the nature of hospital management structures in the UK have already been highlighted in the sections above, but clearly represented another very major issue.

Conclusions

The final project in the S3 programme did not achieve our original aim of demonstrating enhanced success from an integrated patient safety intervention drawing on the three different approaches used in earlier studies. This provided much of the useful learning from this study, as the staff involved revealed a number of important learning points which may be used to improve future intervention efforts, and highlighted some elements of the hospital environment which pose a significant challenge to any such efforts and which need to be seriously considered in future studies. We recommend providing more direction about the targets of improvement than we used here and restriction of the number and nature of the projects initiated to avoid diffusion or diversion of effort. We remain convinced of the benefits of grass-roots engagement and initiative in the change process but recognise that stronger support from senior management may be needed to ensure that change is implemented. We accept that longer-term, more in-depth, training and support programmes based on larger groups may be needed to make major sustainable change. We recognise the great challenges posed by the working patterns in hospital which make team formation and co-operation on projects difficult, especially lack of time and space to do so. Equally important are the issues that prevent many middle managers and senior clinicians from engaging actively, particularly the diffuse nature of responsibility and authority in hospitals, which makes change management a major logistical challenge. We have developed approaches to implementation of our interventions which address many of these challenges, but more work is required before we can assess their impact.

Chapter 10 Qualitative analysis

Introduction

This chapter focuses on the qualitative phase of the project. This consisted of a qualitative evaluation of all of the intervention arms of the S3 programme, the purpose being to identify and explore factors impacting on feasibility and success of bottom-up staff-led improvement interventions. It aimed to explore the experiences of all six of the S3 studies to gain better understanding of the successes and failings of the programme.

Methods

Design

Semistructured interviews were used for the purpose of this evaluation and an interview guide was developed. The interview guide utilised was flexible throughout the interviews to allow for iterative development built on concurrent analysis using an approach based on the constant comparative method.⁴⁹ For a fuller account of the qualitative methodology, please see *Chapter 3, Qualitative Methods*.

Participants

Semistructured interviews were conducted with 34 individuals: 23 key frontline staff and 11 members of the research team. The interviews were carried out by two members of the research team who had a lesser involvement in the intervention arms of S3 (LF and FC) and were transcribed by the main interviewer (LF). The breakdown of hospital frontline staff interviewed can be seen in *Table 31*.

Analysis

Analysis was based on the constant comparison method. Analysis was conducted partly concurrently with data collection in order to allow for development of the initial interview guide. NVivo software (version 10) was used to assist the coding process. The analytic procedure consisted of creating codes and categories within the data collected, informed both by a priori categories developed prior to data collection and through more inductive analysis of the data. This was conducted by two of the researchers who were independent to the rest of the team (LF and GM).

Results

Analysis of interviews revealed a number of common themes. The findings are organised into these themes under the following headings.

TABLE 31 Hospital frontline staff interviewed

Staff	Number interviewed
Management	4
Doctors	9
Nursing	9
Anaesthetist	1

The importance of senior-level clinician buy-in

Senior-level clinician buy-in seemed extremely important for the success of QI work. Although junior-level engagement was important too, outright hostility towards the project or even simply a lack of engagement from more senior clinicians could result in the halting or complicating of QI work, such that senior staff could effectively veto involvement. Moreover, indifference on the part of senior clinicians was sometimes almost as problematic as outright opposition.

... I never came across a massive opposition and in some ways that wasn't helpful sometimes because people were kind of mildly nice, which is an okay place to work in but when you're trying to get people to rally around something in some ways having some opposition is a good thing because people then come out of the wood work and support it in a meaningful way, but when everyone is sort of saying well this fine, it's a nice idea, it can actually like have some inertia and sort of slow down ...

Researcher 6

Despite the bottom-up intentions of the S3 study, and despite the fact that junior doctors have some leeway to do this work without senior-level buy-in, they felt that leadership and champions from this level were essential if QI work was to have any chance of success:

... you need someone who's passionate about it and somebody who is relatively senior who has the air of people that can make decisions. If you just give it to a junior person then it's a lot more difficult for them because they just won't get anywhere ...

Doctor 1 (lean and TT)

An inability to secure this engagement from senior clinicians proved more of an impediment to change than, for example, a lack of buy-in from non-clinical managers. Although managerial approval was necessary for QI work, more proactive support on the part of managers was not essential. Passive support from managers, without active engagement, appeared adequate for QI work to continue. Provided initial approval was given, management-level engagement became much less important than that of senior-level clinicians. The hierarchy of hospital organisations, of course, is unusual in that senior management tend not directly to manage senior clinicians. In such professional bureaucracies, senior clinicians could strongly influence the behaviour of junior doctors, and similarly senior nurses could allow or prevent the involvement of junior nurses.

... perhaps just the hospital is and the NHS is, but certainly there isn't an overarching controlling type of leadership to the department and it's very much individuals working together as a team, rather than someone, some people, telling other peers what to do, so I think to try and get, whereas further down there is a much more hierarchical structure, you know, so you can be told what to do by people and that sort of cascades through various grades ...

Doctor 1 (all)

There could sometimes therefore be something of a misalignment between the bottom-up approach adopted by the project and the reality of the way hierarchies within clinical professions operated in practice; those with the authority to make things happen often had no interest in the projects, which could scupper their progress, whereas those with the interest sometimes lacked the authority to make them happen.

Ability of staff to undertake quality improvement work

Staff at participating hospitals had received little training in QI approaches or methodologies prior to the S3 study and, currently, many in the English health-care system would have not received it during their clinical training.

... as a doctor you're very much involved in learning a medical body of knowledge and you get on with that but you learn very little about management and processes ...

Doctor 2 (all)

There was also limited knowledge of change management and how to begin making substantial change effectively in quality and safety.

... People don't understand what change is, what improvement culture is, what the tools are, how to use the tools, how to think about things ...

Doctor 1 (lean)

Staff are likely to hold the best knowledge of what the local issues in service quality and patient safety are, and of what needs to be addressed in order to achieve better end results. However, they were often not aware of the best methods to use to achieve this.

... I think probably most people think they're better equipped than they in fact are. You know, that they don't understand the techniques and approaches that are available ...

Researcher 4

Although the S3 study team did provide training in QI, and supported staff throughout the projects, staff still did not seem to believe themselves to be competent to do such work independently. This, of course, raises the issue of what then happens in future in the absence of the support of the S3 study team.

... I was happy with their involvement because we really didn't know what we were doing ...

Nurse 2 (TT)

... Prepared or competent? Probably not to be honest. It would be a bit beyond my ...

Nurse 1 (all)

In addition to having a lack of expertise in the area, staff were restricted by the resources available to them. In an increasingly resource-pressured NHS, factors such as staff shortages and financial cuts impacted upon the feasibility of QI work.

... I mean we have weeks where we're really short staffed and it just makes it absolutely impossible to keep up, you know, so we do our best. So staffing levels definitely an issue ...

Care support worker 1 (all)

Sustaining change

Sustainability of change post intervention in the case of the S3 study intervention appeared reliant on the existence of one of two features of change: (1) a change that was 'systematised', large and difficult to reverse (and which would thus take a great deal of active effort from the staff to change) or (2) a change which involved a great deal of effort from one or two individuals, who worked hard to maintain change through time despite indifference or opposition from their colleagues.

... I think where there is like one key person involved and it is very clear, the outcomes have been very visible and clear and everybody has gotten on board in seeing that it's positive, I think in those cases it will be maintained or in cases where the changes that have been made are too difficult to reverse ...

Nurse 2 (all)

Of course, changes sustained through the efforts of one or two individuals appear fragile, and may be difficult to maintain in the long term. They demanded a lot from individuals and it was not always clear what might happen if that individual moved on or was no longer able to sustain the change almost single-handedly:

... it's just if we can sustain it and it does require quite a lot of, sort of, pushing from one person and I think that's where any of these things will have a weakness, is where there is only one person that's driving it because as soon as they stop driving it, then who's going to take over that role?

Doctor 2 (all)

Incentives

Incentivising participation played an important role in securing staff engagement, but the most effective forms of incentives varied by staff group. Doctors seemed more likely to be attracted to the project by the possibility of career advancement, with specific incentives including improving their curriculum vitae (CV) or having the opportunity to be published as an author on an academic paper.

... I think some of them did it for career benefits; to kind of be involved in research, maybe get their name on a paper or put it on their CV ...

Researcher 10

Other incentives, such as the possibility of reducing workload and improving work life, were also attractive incentives for doctors, particularly junior doctors.

... Probably because I could see that there were areas in the department that were, looked like they could be improved and there didn't seem until S3 came along that there was a mechanism to be involved in doing that, so it seemed like a great opportunity to do that ...

Doctor 1 (all)

Nurses appear more incentivised and motivated to participate by the prospect of improving safety, care and their everyday work, rather than career-related benefits.

... I'm also very passionate about patient safety and anything that will improve patient safety and aid staff to do that, I'm keen to get involved in ...

Nurse 2 (all)

Effort required versus rewards received

Quality improvement projects require ongoing effort and support from staff, particularly from those who become leaders or champions of their project. QI initiatives do not always produce the intended benefits; many fail to be sustained and PDCA methodology explicitly recognises that learning from failure is as important as finding success in ultimately securing improvement. Consequently, the commitment required for QI work often does not provide much in return, particularly in the short term, and in the S3 study this could be off-putting for staff.

... I think I enjoyed it initially and then I lost a bit of interest, because I think you suffer from the fact that you feel that you have to put a lot of work in, and you feel that actually if you don't stay to hold that up, then actually when you leave – not that I'm leaving – but that if you were actually things wouldn't be sustained and it's quite difficult to get that sustainability in the projects ...

Doctor 2 (all)

In addition, for QI work to be successful, staff participating in S3 needed not only to do substantial amounts of work over the long term but also to put in significant effort into overcoming indifference and hostility from their colleagues. The levels of effort required from staff presented a challenge when trying to maintain staff involvement for the duration of a project and, again, may suggest challenges around sustainability of change. Thus there seemed, at times, to be an imbalance in the amount of time and effort staff were required to put in compared with the benefits they received (or perceived) for doing so, which was identified by several participants as a major challenge for those leading change and trying to instil and maintain enthusiasm.

... the lives of staff generally, you know, I think have been made, you know, potentially less rewarding by introducing all these safety and quality measures ...

Researcher 9

It was therefore important to provide staff with feedback and real-time evaluation of progress, in order to make benefits more visible and tangible, and thus to positively reinforce their work.

... I know change takes time and you've got to see if things are better in the long term, but a more active feedback process would be good if you want staff engagement ...

Doctor 1 (lean)

'Learned helplessness'

Interviews with some staff members suggested that they felt they could not make successful changes within their organisations. They also felt that others had little to no chance of succeeding and this seemed again to give rise to antipathy, indifference or nihilism when it came to the projects.

... I'm cynical that within an NHS setting many great changes can be made ...

Doctor 3 (all)

Some staff within the S3 project indicated that they could make change in certain aspects of their work but could not affect whole teams or organisational units. Changes at the team or organisational level they saw as outside their control and this could lead to either disengagement with efforts to make such changes or an inclination to orient projects towards small-scale changes that were undoubtedly within their locus of control but which seemed to hold very limited potential in terms of making meaningful change to quality and safety of care. In addition, staff found themselves working within an environment in which NHS trusts seemed to be constantly rolling out new initiatives and changes. Consequently, staff became overwhelmed by the weight of externally driven, mandatory 'initiatives', implemented from the 'top down', and this impacted notably on both the enthusiasm they expressed for their projects and on the time they said they could devote to 'bottom-up', internally driven, non-mandatory QI work.

... these ever-rolling waves of what's perceived to be management change and new initiatives and lack of funding and constant reorganisation and personnel problems and understaffing ...

Researcher 4

Change fatigue

Similarly, the 'ghosts of past initiatives' haunted many staff. Staff had experienced a number of failed initiatives and poor experiences with external consultancies that they felt did not understand the clinical working environment, and which left them feeling annoyed and resentful.

... we've had previous outside agencies come and look at listening to the staff and identifying issues and that was as far as it went. They'd come up with a long list of problems that was pretty obvious to everybody and actually had no idea about how we were going to resolve them ...

Manager (all)

... They come over, take their money, a large amount of money and then they clear off and they've left a few little tools around the place that no one really uses ...

Doctor 1 (lean)

The status of the S3 study as an academic project helped to counter this fatigue and antipathy, as it was not seen as another top-down initiative or another consultancy brought in to help out with 'empowerment':

... I think they were trusted a bit more by virtue of being part of the university and the trust so they were out of our department but within the institution if you like and I think that was helpful ...

Doctor 3 (all)

However, it was still difficult to avoid a sense of contamination of the S3 study ethos by so many past projects that had come in, stirred things up a bit, and left without a legacy of tangible improvements for staff or patients.

Staff issues with lean

Lean, in particular, held negative connotations for many staff who had previously encountered it. Beyond the general sense of change fatigue, there was a particular sense that hospitals had been plagued by failed lean initiatives, some of which did not seem to conform to the real ethos behind lean:

... we did find a great deal of hostility generated, not to us, but to the idea of quality interventions, particularly things involving the word 'lean' or 'kaizen,' so part of that is a general cynicism about what's going in the area around it ...

Researcher 4

Although apparently straightforward conceptually, the lean approach of iterative development and associated techniques, such as PDCA cycles, appeared to cause substantial confusion and misunderstanding when applied in practice. Many doctors within the S3 study were not comfortable with this experimental methodology, finding it alien to the way in which people in hospital environments usually thought and worked.

... I think actually it was a weird concept for everyone, which was strange because why should it be weird that this week we're trialling something and it's only a trial and if you don't like it, it doesn't matter because it was a trial, if you like it then great. But I actually think we were really bad at that ...

Doctor 2 (all)

Such a 'trial'-based experimental approach was not something they were used to taking in everyday work; rather, on the contrary, the culture described by medically trained participants, in particular, was one in which success was valued but failure was not. The expectation was that staff should be confident in knowing what is right to do the first time around. Being associated with an approach that failed despite promise or needed substantial improvement to work properly, was not an attractive prospect for medical staff.

... for people who are not, have never been involved in research before or have never come into contact with any sort of research context, it's a very, very confusing concept and it's a very unusual way for them to think ...

Researcher 10

Furthermore, PDCA cycles in the S3 study that did not result in improvements were sometimes construed as poor performance by other staff who were less engaged in the projects and thus misunderstood the process.

... I think people who didn't understand or didn't, hadn't been involved in it really, didn't like that and you know, actually it probably backfired, didn't it, in many ways doing that, because it was seen as, you know, a failure because it didn't succeed when it was implemented and therefore, and then the you know, the next time you try and do another cycle, it was like 'well it's failed, why are you doing it again, it should just be consigned to the bin' kind of thing ...

Doctor 1 (all)

In contrast, those who were more heavily involved in the projects appeared to acquire a better understanding of the iterative process through time.

Complacency

Complacency was not an issue among all staff; however, interviewees felt that it was not uncommon to find staff who demonstrated a level of complacency in their work and towards QI.

... it's just people feel that they're doing enough and they're doing a great job and they don't need to improve on their abilities or they just think things are a waste of time because they're already doing such a great job ...

Care support worker 1 (all)

These staff believed that they were doing fine by themselves and that there was little within their work which could be improved. Such a view could present a significant barrier. In some cases within the S3 study, a promising way to counter such complacency seemed to be through the use of data to demonstrate that deficiencies within the system did exist.

... they just need to be highlighted really and as soon as you, it's like Pandora's box I find, as soon as you've highlighted a problem you can't then ignore it, which I think is good because the problem is when people know it's an implicit understanding there's poor care being offered in certain areas, it's very difficult to sort of highlight it but once it's highlighted then you can't say well we didn't know about this, if it's a problem now you can say well we knew about this a while ago so ...

Anaesthetist (all)

Time

Time appeared to be one of the largest barriers to QI. Staff had no time set aside or protected in which they could undertake QI work.

... You know I'm doing a clinical job, OK, that's 110% of my time. Oh you want to do a leadership project as well? Well that's 120% of my time someone's got to give and generally the clinical work takes priority over all the administration stuff, so the administration stuff gets put on the back burner ...

Manager 1 (all)

Staff involved in the S3 study appeared to have completed a substantial amount of QI work in their own time, as they did not have enough time to complete their clinical commitments and the improvement work within their formal working hours.

... the main problem with what we do is that I don't have the time to do stuff that needs to be done out of hours, so I actually think that that is a big thing, because if I choose like to do anything like an audit or a project, we don't get time given to us for that, so it has to either fit in within the work time, which is fine some days or it has to be out of hours, so like the stuff like when I was thinking about the [names project], reading all the [names project] stuff, it was done at home or before coming to work so ...

Doctor 2 (all)

Many staff interviewed believed that without the support of the S3 study team they would not have had time and the work would not have been feasible.

... because of the pressure with time and work; we haven't got any capacity to do it. The advantage of having S3 support is that that's what they're going to do and then you can help them and they will help you to do that, but they've got the resource and a bit of the time ...

Doctor 1 (SOP)

Likewise, members of the research team indicated that they often had to target staff and schedule meetings in staff personal time, such as breaks. It was also suggested that the time requirements was a contributory factor in a number of project drop-outs and in the diminishing levels of engagement in some sites over the course of the projects, and that it may also have been a possible deterrent to initial involvement.

... they've got a lot of other demands on their time, it came to the bottom of the pile of other things ...

Doctor 1 (all)

In addition to this, it is important to note that the time is not just a commitment that is required for setting up a project: ongoing input was also required over extended periods to ensure the progress, success and sustainability of the work. Many staff involved in the S3 study indicated that time set aside for improvement work in the future would be helpful; this currently occurs in many other industries in which staff are expected to participate in QI work as a routine part of their job and development.

... the idea of lean improvement in health care is flawed in that it's based upon the staff having time to do the improvements themselves and until management resource that, and so until they're happy to give people time off to do that improvement work I don't think it helps and by forcing people to do this in their own time, or in breaks, then it reduces it's the feeling of importance to this kind of work ...

Researcher 2

Opportunity

Currently, there is a lack of opportunity for staff to do QI work. Not only do they not have the time in which to it, as described, but also the opportunities to engage in such work are not always available in their work environments. First, staff indicated that within the current climate organisational aims and priorities are primarily financial and not focused on improvement of quality and safety.

... All you're worried about is financial balance and targets, so you don't do anything interesting or fun. It's just chaos management ...

Manager 1 (all)

Second, relating to the importance of senior-level clinician involvement, hierarchical barriers served to block opportunities for junior staff to get involved in QI work. If senior staff did not approve of this work, then junior staff were unable to engage in it without conflict.

... I think [junior doctors] maybe they just feel more controlled by the senior staff anyway, so even if they try to, because when you, if you have an environment and you try to change things and you're not senior it's very difficult to, it's very difficult to and so maybe they just hoped if they went along with it the older ones, the more senior ones rather will actually make it happen and they can be sort of part of it rather than struggling to get there on their own ...

Researcher 5

Although the S3 study demonstrated that an effective research team can work to facilitate the flattening of a hierarchy, there is a risk that this may only be temporary: once the project period ends, it may not be maintained, as the 'oasis' provided by active intervention by an external team is left.

Third, there were few opportunities for staff to engage with one another as a team in QI work. Interviewees suggested that there were few times and places when staff involved in a QI project are all present at the same time. Compounding this issue, teams often rotated, resulting in difficulty in conducting and sustaining QI work with many staff only touching the surface of the work being done.

... it was good to get a bit of time and a little bit of space to get all the members of staff together, to actually have a conversation about what we were doing because normally that doesn't happen ...

Doctor 2 (lean)

... exactly who was on the team would vary from week to week anyway, so it wasn't even on the same Wednesday that you saw the same core people, a pattern. And so in our intervention people we had some people who were really only just touching the surface of what we were doing ...

Researcher 4

The external team

The fact that the S3 study was an external group seemed to be one of its greatest assets. The academic status of the S3 study team, and the fact that they were external to trusts, was very helpful in circumventing 'political' issues relating to hierarchies and permissions needed to proceed, and in helping the team to be seen to be able to generate energy around the need for change, countering some of the nihilism, scepticism and change fatigue noted.

... you guys could find the problems such as I couldn't because it wasn't quite as, you know, personal ...

Anaesthetist 1 (all)

This was helped also by the team's academic legitimacy, as noted: the S3 study team were not perceived as 'quick-solution consultants' whose work would be short-lived but expensive but rather as a group that aimed to improve patient safety for research purposes with potential long-term benefit for the NHS.

... I think they were trusted a bit more by virtue of being part of the university ...

Doctor 3 (all)

They were perceived as having a valuable outsider's perspective because of the academic focus of the work and also perhaps because of the fact that the team included members with clinical experience.

... The fact that they are external and they can give a different, they can be objective about it and they don't have any of the baggage of the way it's always been done so getting somebody external in is brilliant because it's very difficult to have another perspective on something that you're part of ...

Doctor 1 (lean)

Paradoxically, then, working 'outside' the hierarchies and politics of the institutions of the participating trusts was both the greatest asset of the S3 study, in that it bred goodwill and enabled the kind of conjoint work necessary to achieve improvement, and also its greatest liability, in that it lacked the levers to ensure engagement of certain key groups, making it difficult to work beyond small, cohesive groups and address the more complicated problems.

... I think the disadvantages further down the track, you know, to being external, you know, in some people's eyes I'm sure it was perceived as a bit of a threat, so ... and so maybe then didn't get as much longer term engagement ...

Doctor 1 (all)

Team rapport

The other major assets of the S3 study team, and a factor that separated them from previous external consultancies brought in to address quality and safety, were the rapport and the relationships that the team managed to develop with the staff with whom they worked.

... [names researcher] and [names researcher] were so positive and enthusiastic and were always smiling and bubbly ... it will be a shame to not see them around ...

Manager 1 (all)

... It's as if they are an internal component on the ward, which is by far in a way the biggest difference to the outside companies and ideas we've had in the past, is it's really not like working with outside companies ...

Nurse 2 (all)

This allowed them to work closely with staff and establish a trusting working relationship, with some staff stating that they got involved simply because of the enthusiasm and commitment of the members of the S3 study team.

... I think a lot of people did get involved just because of how enthusiastic and positive [names researcher] and [names researcher] were and how just they seen that they were so confident in their work, they were so happy to go and meet staff, they were just so positive and enthusiastic and always there, always willing to meet with people, like there was just a complete continuity of support, that the staff that felt that they were then part of the team, that they were always there for them, that they could go to them with any issues ...

Researcher 10

Team support

The S3 study provided substantial amounts of support for staff during the QI work. This included expert skills input, the facilitation time and the ability to undertake the niggling tasks that are crucial to making improvement happen (e.g. data collection and information management).

... The push from them that they'd say 'well you've done this phase, what's the next phase, we'll set the date for the next meeting and it has to be in this time frame' – that sort of overall admin bit of it, that kept it going because you could have one meeting and then it would just like drift off and then four months later you'd have another one, so that kind of organisation, push ...

Nurse 2 (all)

Through their support, hospital staff indicated that the S3 study team massively increased staff capacity for QI.

... bringing sort of a multidisciplinary group of people together who previously I don't think really had engaged or had a way of engaging together ...

Doctor 1 (all)

... Well just to organise and inform the way of going about this, you know. And fundamentally, so there's two things. I think one is that education about the way to do it and the sort of structure and sort of direction to people of the way to do and the help in doing some of the more mundane aspects of data collection that are never appealing to people, but then the other side of it is just cajoling, you know, and just organising all the stuff that takes forever; trying to pull people together at the same time and you know, continually over an extended period of time, that kind of co-ordination is, I mean, seems relatively unsophisticated if you like, but actually that's the crux of getting it to happen it I think ...

Doctor 1 (all)

What S3 study support had difficulty in addressing was the more structural and cultural aspects of capacity, as these transcended the specific local organisational units of intervention. Arguably, furthermore, the temporary provision of extra capacity may risk being counterproductive, if it removes from staff the responsibility to make things happen, and convinces them of their inability to do such work without external support from a study team like S3.

... I think that a huge amount of time, effort, motivation, communication, facilitation has been done by the research team that I think at times may not be recognised ...

Researcher 3

Conclusion

The qualitative substudy shed a light on how the S3 study team, through their position as an external group, the support they provided, the strategies they utilised and the relationships they formed with staff, increased the capacity for bottom-up QI work to occur. However, there exists a challenge in S3 study-type interventions to ensure that the ability to address the local problems is fostered and that there is not a sense of dependency on the team, whose involvement is inevitably temporary. At the same time, it is crucial to intervene separately at a higher organisational level to influence culture, rather than pursuing the S3 study in isolation at the 'operational level', and expecting it to exert an upwards influence on factors that transcend the frontline (e.g. wider organisational culture and structural/organisational problems). Another perspective using the same material is provided by an analysis using the ability–motivation–opportunity framework from the organisational behaviour literature,⁸⁰ and it may be useful to read this additional material alongside this chapter when considering the implications for future intervention programmes.

Chapter 11 Health economic analysis

Introduction

Harm caused to patients as part of medical treatment leads to substantial costs incurred by the NHS in treating affected patients, short- and long-term health impacts on patients and litigation costs.⁸¹ Therefore, effective interventions capable of improving patient safety may also save the NHS scarce resources to invest elsewhere. Previous chapters in this report have concentrated on evaluating the impact of SOPs, TT and the lean process on the safety and quality of surgical care, using the Oxford NOTECHS II scale score, counting operative process glitches and WHO checklist completion as the main outcome measures. In this chapter we extend these previous analyses and describe the health economic component of the S3 study, which took the form of a cost analysis of the interventions and the impact of the training programmes on hospital-related resource use.

Objectives

This part of the research sought to conduct a detailed cost analysis of the teamwork interventions and their associated impact on hospital-related resource use at the patient level.

Methods

In this section we present the data and statistical analysis used in the health economics component of the study. We explain that the complexities of the S3 study introduced some challenges that forced the team to introduce some deviations from the original proposal in order to successfully complete the study. These challenges affected mainly the original data collection. The following sections provide details of such changes, the final data available for analysis and the statistical analysis conducted for the health economics component of the S3 study.

Deviations from original proposal and final data sources

The S3 programme had a complex design that involved collecting information from different sources, including individual patient-level data on theatre use and hospital stay. The original plan in the proposal specified that data for the health economics analysis would be collected from hospital records and patient questionnaires. Given that responsibility for and custodianship of each data source was different, obtaining a full analysis sample based on the original plan was a somewhat complex process. The original health economics analysis plan needed four separate components:

1. aggregate-level data – information collected from the S3 data records and expert opinion on the preparation and delivery of the training programmes
2. patient-level data – information collected on hospital-related resource use
3. PROMs – data on health-related quality of life using the generic questionnaire EQ-5D
4. patient perceptions of quality of care – an outcome questionnaire successfully used in previous surveys that documented any illness, injury or impairment considered by the patient to have resulted from medical care they had received, together with information on the perceived severity.

Given that most hospitals had systems in place to identify harm to patients, concerns were expressed that surveying a sample of patients from specific hospitals to ask about their perceptions of quality of care received could prompt concerns among patients and could potentially generate patient complaints that would not otherwise have been made. As a result, the patient perception questionnaire (point 4) had to be withdrawn from the study.

Most of the data required for the S3 study were collected at the theatre level, in which observers collected the information needed for the statistical analyses reported in *Chapters 4–6*. However, for the health economic analysis, patient-level data were required. When the S3 programme commenced, the ethics application requested permission to access hospital records that included personal details (e.g. date of birth and sex, clinical details contained in medical notes, clinic letters and discharge summaries, operation notes, nursing and anaesthetic notes, radiology reports and laboratory reports). However, when discussions with information analysts from each hospital started, this level of access was considered insufficient to permit extraction of hospital record information for patients observed in theatre. The S3 study group consulted the Research Ethics Committees from Bristol and Oxford to clarify the ethical aspects of the project. After several rounds of discussions it became clear that the study had not been granted the appropriate permissions to access records from patients observed in the study. It was suggested then that patient consent was needed in order to access information on hospital-related and PROM data. Given that the primary level of assessment in the S3 study was members of theatre staff and not patients directly, it was concluded that it would be impractical to obtain individual patient consent, for two main reasons: (1) obtaining individual patient consent would have introduced an additional hurdle to an already complex study and would have delayed further the health economics analysis and (2) the observed final in-theatre sample size was considered to be too small to conduct any meaningful before and after comparison. Hence, rather than evaluating clinical outcome data from the individuals observed in the study, we obtained ethics approval to extract non-identifiable individual patient-level data from all patients under the care of the consultants participating in the S3 study. Unfortunately, this decision precluded collection of quality-of-life data using the EQ-5D instrument (point 3), which was also withdrawn from the study.

As a result of these issues, robust data for the health economics component of the project were available for points 1 and 2.

Aggregate-level teamwork training data (point 1)

The study interventions varied in their scope and content, but they had similar resource use aspects in practice, including:

- a preliminary meeting between experts in TT and the clinical surgical team to evaluate the current situation and become familiar with the study intervention
- an 8-hour off-site meeting with training co-ordinators at which the clinical team learned about the intervention in detail and developed a strategy for its implementation in theatre
- follow-up meetings between the clinical teams and trainers to evaluate progress, correct any problems and eliminate any dangers.

These three components differed in terms of the intensity of resource use needed to deliver each intervention, particularly for the follow-up period. *Table 32* provides a summary of the resource-use data needed to estimate the cost of each intervention. Data for this section of the project were provided by the experts who delivered the training programmes complemented with information provided by the S3 study group.

Patient-level hospital-related resource data (point 2)

Data on readmissions within 90 days, whether or not a complication had occurred and length of stay were extracted from hospital records in the three sites participating in the study [Nuffield Orthopaedic Centre, University Hospital Coventry and Warwick (including Rugby), and Kettering]. Baseline demographic information on age and sex was also extracted. For each consultant participating in the study in the intervention or control group, data were obtained for all his or her patients 6 months before and 6 months after the intervention was delivered. We contacted the senior information analysts in each hospital with the appropriate ethics documentation, a covering letter and a Microsoft Excel template with the required formatting for the data. The covering letter specified the variables and the periods of time for each consultant over which the data were needed. In order to further ensure anonymity and to avoid linking consultants to a particular case, consultants were combined in groups. These groups were then used to identify cases for a

TABLE 32 Resource-use data needed to calculate the costs of the interventions

Intervention component	Intervention study
TT	
Preliminary meeting	Duration of meeting between trainers and the theatre teams
Off-site meeting	Allocated time of training co-ordinators preparing and delivering the material for the course and number and type of team clinical staff attending the meeting
Follow-up	Trainer's time twice weekly over 2 months
Lean process improvement	
Preliminary meeting	Duration of meeting between experts and the theatre teams
Off-site meeting	Allocated time of training co-ordinators preparing the material for the course and number and type of team clinical staff attending the meeting
Follow-up	Number of full days experts visit operating theatres
SOPs	
Preliminary meeting	Duration of meeting between experts and the theatre teams
Off-site meeting	Allocated time of training co-ordinators preparing and delivering the material for the course, and number and type of team clinical staff attending the meeting
Follow-up	Number of times team and experts met to discuss the SOPs progress. Number of times observers attended operating theatres to evaluate teamwork

particular intervention for the before and after intervention periods and the intervention and control groups in the extracted data set. All hospitals provided the data using the Microsoft Excel template circulated, which facilitated the merging, cleaning and preparation of the working file for final analysis.

Cost analysis of training programmes

Having identified the main types of resources needed to successfully conduct each intervention (preliminary meetings, off-site meetings and follow-up), the frequency and duration of each item were identified and summarised using descriptive statistics, when appropriate. Resource volumes were then multiplied by appropriate unit costs to calculate the cost of the intervention. This analysis was conducted in Microsoft Excel 2013. Unit costs per hour expressed in 2012 prices of orthopaedic operating theatre staff, and sources of information, are given in *Table 33*.

TABLE 33 Standard UK orthopaedic operating theatre staff and invested training time

Staff member	Number	Cost (2012 prices) per hour ^{a,b} (£)	Source
Consultant surgeon	1	172	PSSRU ⁸²
Consultant anaesthetist	1	172	PSSRU ⁸²
Surgical registrar/fellow	2	71	PSSRU ⁸²
Senior house officer	2	20	DH ⁸³
Team leader	1	58	PSSRU ⁸²
Theatre manager	1	41	DH ⁸³
Theatre nurse	1	14	DH ⁸³
Operating department practitioner	2	14	DH ⁸³

DH, Department of Health; PSSRU, Personal Social Services Research Unit.

a It includes salary, on costs, qualifications, overheads, training and capital overheads.

b It is assumed that each course (TT, lean and SOP) lasted for 8 hours.

Hospital-related data statistical analysis

Baseline demographic information was summarised using descriptive statistics. We used *t*-tests for mean age and chi-squared tests for sex distribution to compare the before and after periods. Binary variables (readmissions within 90 days and whether or not at least one complication has occurred) in the before and after periods were compared using odds ratios and 95% CIs from a logistic regression, adjusted for age and sex. Mean length of stay in the before and after periods was compared using linear regression, controlling for age and sex. Before-and-after differences in the active and control groups were compared using a standard parametric test of interaction. We conducted separate analyses for each intervention (TT, lean, SOP, lean and TT, SOP and TT), but also pooled all interventions in a unique group (integrated) to explore the impact of any intervention in the outcomes assessed. The impact of TT versus lean and SOP active groups and the impact of lean versus SOP active groups were also evaluated in two additional separate analyses. This statistical analysis was conducted using Stata version 12.

Results

Training programme costs

The cost per hour and total cost of time involved by non-research and research staff in setting up the courses for each intervention are reported in *Table 34*. The three interventions required similar setting-up resources with an average cost of £1700.

TABLE 34 Non-research and research staff preparation time^a of course set-up

Staff member	Hours	Cost (2012 prices) per hour ^b (£)	Unit cost source	Total (£)
TT				
<i>Non-research</i>				
Consultant surgeon	0.5	172	PSSRU ⁸²	86
Band 6 nurse	10	16	DH ⁸³	160
S3 study project manager	32	19	Finance Division OU (Nigel Byng, R&D Finance Team, Oxford University Hospitals NHS Foundation Trust, 2014, personal communication)	608
<i>Research</i>				
CRM expert	2	67.5	Atrainability invoices	135
Grade 7: Oxford	8.5	22	Finance Division OU (Nigel Byng, R&D Finance Team, Oxford University Hospitals NHS Foundation Trust, 2014, personal communication)	187
Grade 7: Warwick	8	22	Finance Division WU (Nigel Byng, R&D Finance Team, Oxford University Hospitals NHS Foundation Trust, 2014, personal communication)	176
Grade 6: Warwick	8	17	Finance Division WU (Nigel Byng, R&D Finance Team, Oxford University Hospitals NHS Foundation Trust, 2014, personal communication)	136
Total				1488
Lean				
<i>Non-research</i>				
S3 project manager	24	19	Finance Division OU (Nigel Byng, R&D Finance Team, Oxford University Hospitals NHS Foundation Trust, 2014, personal communication)	456

TABLE 34 Non-research and research staff preparation time^a of course set-up (*continued*)

Staff member	Hours	Cost (2012 prices) per hour ^b (£)	Unit cost source	Total (£)
<i>Research</i>				
Lean expert	8	62.5	Lean expert (Dr Steve New, Associate Professor of Operations Management, Saïd Business School, University of Oxford, 2014, personal communication)	500
Grade 7: Warwick	24	22	Finance Division WU (Nigel Byng, R&D Finance Team, Oxford University Hospitals NHS Foundation Trust, 2014, personal communication)	528
Grade 6: Warwick	24	17	Finance Division WU (Nigel Byng, R&D Finance Team, Oxford University Hospitals NHS Foundation Trust, 2014, personal communication)	408
Total				1892
SOP				
<i>Non-research</i>				
S3 study project manager	24	19	Finance Division OU (Nigel Byng, R&D Finance Team, Oxford University Hospitals NHS Foundation Trust, 2014, personal communication)	456
<i>Research</i>				
SOP expert	12	62.5	SOP expert (Dr Steve New, Associate Professor of Operations Management, Saïd Business School, University of Oxford, 2014, personal communication)	750
Grade 7: Oxford	16	22	Finance Division OU (Nigel Byng, R&D Finance Team, Oxford University Hospitals NHS Foundation Trust, 2014, personal communication)	352
Grade 7: Oxford	16	22	Finance Division OU (Nigel Byng, R&D Finance Team, Oxford University Hospitals NHS Foundation Trust, 2014, personal communication)	352
Total				1910
DH, Department of Health; OU, Oxford University; PSSRU, Personal Social Services Research Unit; WU, Warwick University.				
a Preparation time provided by the S3 study researchers and project manager.				
b It includes salary, on costs, qualifications, overheads, training and capital overheads.				

The estimated costs of the three interventions are reported in *Tables 35–37*. The TT programme cost £18,768 (see *Table 35*), the lean training £14,338 (see *Table 36*) and the SOP programme £10,856 (see *Table 37*); most of these costs related to teaching time by faculty, attendance time by participants and in-theatre follow-up coaching. The TT programme was estimated to be the most expensive of all interventions with an additional cost of £4000 compared with lean training and £8000 compared with SOP.

Hospital-related data

Figure 13 shows the total number of patients in the active and control groups and the before and after periods of each intervention within each participating hospital. *Table 38* reports the number of operations performed by surgeons participating in the S3 study by intervention group: in total 4810 operations were performed in the active groups and 4757 in the control groups.

The mean age in each group in the before and after periods is reported in *Table 39* and the male/female proportions in *Table 40*. No significant differences were apparent in the after and before comparisons. The figures are very similar to those reported in the much smaller sample of patients whose operations were directly observed by researchers to provide the data for our analyses of team process (see *Chapters 3–8*), confirming that the convenience samples used for these studies were representative.

TABLE 35 Resource-use and associated costs related to the TT programme

Category	Resource use	Costs (2012 prices; £)	Source
Course set-up			
• Staff involved	See <i>Table 33</i>	1488	S3 study project manager (Sarah Hills, Beth Baslak or Sam French, Nuffield Department of Surgery, University of Oxford, 2014, personal communication)
• Transport	–	162	–
Off-site training			
• Faculty	2	3650	Invoices
• Participants	11	5323	S3 study project manager (Sarah Hills, Beth Baslak or Sam French, Nuffield Department of Surgery, University of Oxford, 2014, personal communication)
• Administration	1	152	S3 study project manager (Sarah Hills, Beth Baslak or Sam French, Nuffield Department of Surgery, University of Oxford, 2014, personal communication)
In-theatre coaching			
• Faculty ^a	2	7200	Invoices
• Transport	–	545	Invoices
• Venue/catering	–	539	Invoices
Total costs CRM		18,768	
a Over 7 days.			

TABLE 36 Resource-use and associated costs related to the lean training

Category	Resource use	Costs (2012 prices; £)	Source
Course set-up			
• Staff involved	See <i>Table 33</i>	1892	S3 study project manager (Sarah Hills, Beth Baslak or Sam French, Nuffield Department of Surgery, University of Oxford, 2014, personal communication)
Off-site training			
• Faculty	1	1500	Lean expert (Dr Steve New, Associate Professor of Operations Management, Saïd Business School, University of Oxford, 2014, personal communication)
• Participants	11	5323	S3 study project manager (Sarah Hills, Beth Baslak or Sam French, Nuffield Department of Surgery, University of Oxford, 2014, personal communication)
• Administration	1	123	Lean expert (Dr Steve New, Associate Professor of Operations Management, Saïd Business School, University of Oxford, 2014, personal communication)
In-theatre coaching			
• Faculty	1 ^a	5500	Lean expert (Dr Steve New, Associate Professor of Operations Management, Saïd Business School, University of Oxford, 2014, personal communication)
• Transport	–	–	–
• Venue/catering	–	–	–
Total costs lean		14,338	
a Over 11 days.			

TABLE 37 Resource-use and associated costs related to the SOP training programme

Category	Resource use	Costs (2012 prices; £)	Source
Course set-up			
• Staff involved	See <i>Table 33</i>	1910	S3 project manager (Sarah Hills, Beth Baslak or Sam French, Nuffield Department of Surgery, University of Oxford, 2014, personal communication)
Off-site training			
• Faculty	1	1000	SOP expert (Dr Steve New, Associate Professor of Operations Management, Saïd Business School, University of Oxford, 2014, personal communication)
• Participants	11	5323	S3 project manager (Sarah Hills, Beth Baslak or Sam French, Nuffield Department of Surgery, University of Oxford, 2014, personal communication)
• Administration	1	123	S3 project manager (Sarah Hills, Beth Baslak or Sam French, Nuffield Department of Surgery, University of Oxford, 2014, personal communication)
In-theatre coaching			
• Faculty	1 ^a	2500	SOP expert (Dr Steve New, Associate Professor of Operations Management, Saïd Business School, University of Oxford, 2014, personal communication)
• Transport	–	–	–
• Venue/catering	–	–	–
Total costs SOP		10,856	
a Over 5 days.			

Table 41 reports the frequency and proportion of patients readmitted within 90 days across interventions, before and after the intervention and in the active and control groups, for each intervention and for the active versus control groups integrated. The before and after adjusted odds ratio is reported with 95% CIs, and this is also shown graphically in *Figure 14*: the only comparison attaining statistical significance is in the lean plus TT active group, in which readmissions after the intervention were significantly higher (odds ratio 4.8, 95% CI 1.32 to 17.4). However, this is based on small absolute numbers: three in the before phase and 12 in the after phase. The table also shows that there were no differences in the frequency of readmissions for the before and after analysis in the TT compared with the other interventions or for the lean process compared with SOP. Tests for interaction between the active and control groups were uniformly non-significant.

Table 42 shows a similar analysis for complications and, again, odds ratios are shown graphically in *Figure 15*. In this comparison several of the before and after comparisons are significant, with higher reported complications following the intervention in the lean and TT control group, the SOP control group, and the integrated active (odds ratio 1.32, 95% CI 1.12 to 1.55) and control (odds ratio 1.34, 95% CI 1.15 to 1.56) groups. A significantly higher number of complications were observed in the before and after analyses in any TT active and lean and SOP active interventions. A higher number of reported complications were also found in the any SOP interaction before and after group (odds ratio 1.63, 95% CI 1.35 to 1.97). A marginal significant interaction between any lean active and any SOP active ($p = 0.04$) was detected.

Table 43 and *Figure 16* report the mean frequency of complications by intervention, before and after the intervention and in the active and control groups, adjusting for age and sex. These results confirm those in *Table 43*, with significantly higher rates in the after than the before phase in the lean and TT control group, the SOP control group, the lean and SOP active group, the any SOP active group, and in the integrated active and control groups. In this analysis, the interaction test of active compared with control

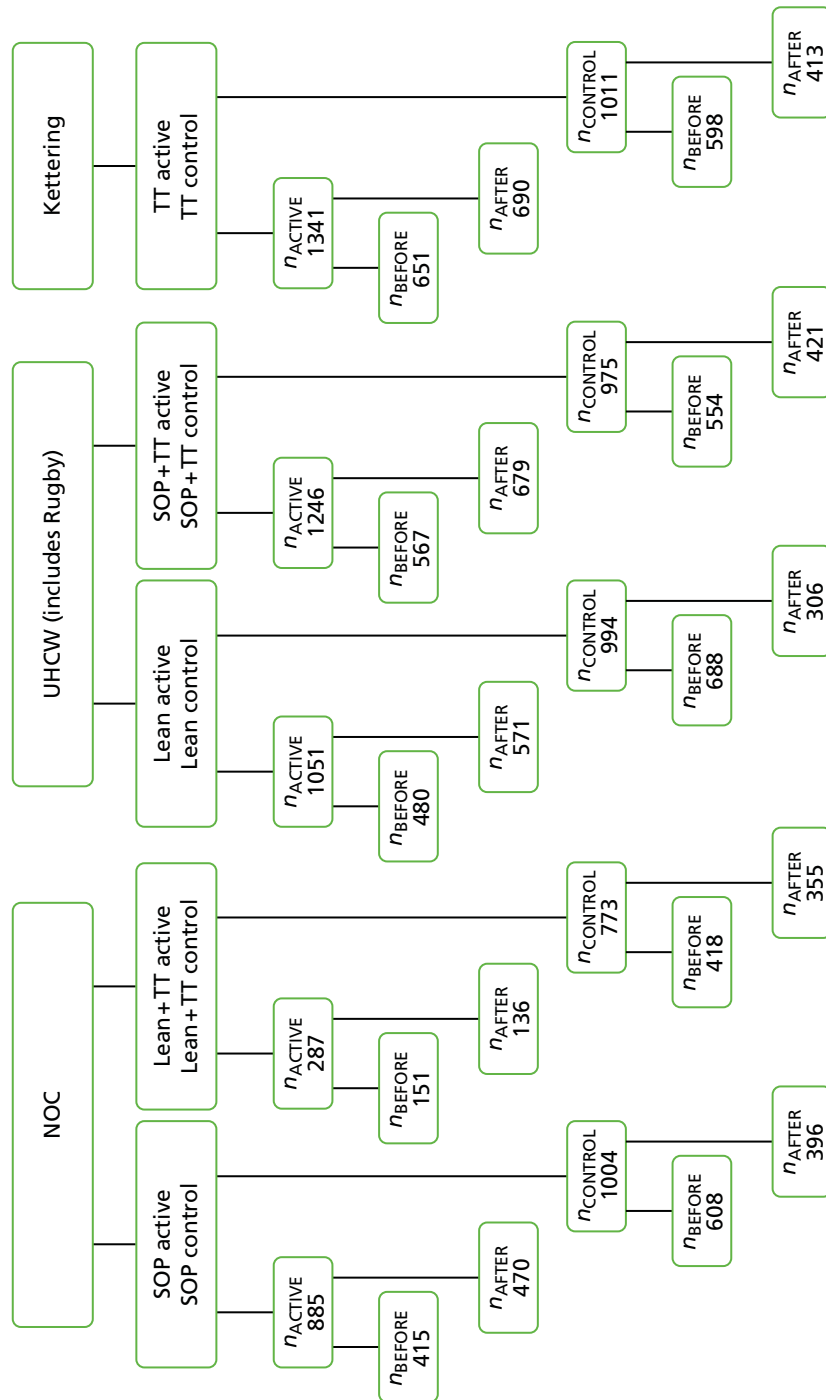


FIGURE 13 Flow diagram of hospital-related data available at the individual patient level available in the S3 study. NOC, Nuffield Orthopaedic Centre; UHCW, University Hospitals Coventry and Warwick.

TABLE 38 Frequency (percentage) of operations performed by consultants participating in the S3 project during the study period by intervention before and after change and in the active and control groups

Intervention	Before, frequency (%)	After, frequency (%)	Total
Integrated active	2264 (47)	2546 (53)	4810
Integrated control	2866 (60)	1891 (40)	4757
TT active	651 (49)	690 (51)	1341
TT control	598 (59)	413 (41)	1011
Lean active	480 (46)	571 (54)	1051
Lean control	688 (69)	306 (31)	994
Lean + TT active	151 (53)	136 (47)	287
Lean + TT control	418 (54)	355 (46)	773
SOP active	415 (47)	470 (53)	885
SOP control	608 (61)	396 (39)	1004
SOP + TT active	567 (46)	679 (54)	1246
SOP + TT control	554 (57)	421 (43)	975

TABLE 39 Mean (SD) age (years) across interventions before and after change and in the active and control groups, and parametric *t*-test comparison before and after change

Intervention	Before, mean (SD)	After, mean (SD)	After vs. before, <i>p</i> -value
Integrated active	54.0 (20.9)	53.9 (21.4)	0.79
Combine control	57.6 (21.3)	57.8 (21.0)	0.74
TT active	57.8 (21.8)	57.0 (23.0)	0.53
TT control	51.7 (25.2)	52.2 (24.2)	0.77
Lean active	56 (24)	54 (23)	0.30
Lean control	57 (23)	54 (23)	0.07
Lean + TT active	53 (19)	52 (18)	0.58
Lean + TT control	63 (14)	65 (14)	0.05
SOP active	49 (17)	50 (20)	0.60
SOP control	63 (15)	62 (15)	0.42
SOP + TT active	52 (19)	54 (20)	0.19
SOP + TT control	55 (23)	56 (24)	0.59

SD, standard deviation.

TABLE 40 The frequency (percentage) of males across interventions before and after change and in the active and control groups, and chi-squared test comparison before and after change

Intervention	Before, frequency (%)	After, frequency (%)	After vs. before chi-squared test comparison, <i>p</i> -value
Integrated active	1187 (52)	1301 (51)	0.36
Integrated control	1424 (50)	956 (51)	0.56
TT active	311 (48)	330 (48)	0.98
TT control	311 (52)	212 (51)	0.83
Lean active	224 (47)	292 (51)	0.15
Lean control	352 (51)	173 (57)	0.12
Lean + TT active	73 (48)	61 (45)	0.55
Lean + TT control	180 (43)	175 (49)	0.08
SOP active	279 (67)	286 (61)	0.05
SOP control	291 (48)	184 (47)	0.67
SOP + TT active	300 (53)	332 (49)	0.16
SOP + TT control	290 (52)	212 (50)	0.54

TABLE 41 Frequency (percentage) of patients readmitted within 90 days across interventions before and after change and active and control groups. Before and after comparison using adjusted odds ratio (95% CI)

Intervention	Before, frequency (%)	After, frequency (%)	Adjusted odds ratio (95% CI) ^a	Test for interaction active vs. control
Integrated active	112 (5.0)	131 (5.2)	1.04 (0.81 to 1.35)	$z = -0.98; p = 0.25$
Integrated control	192 (6.7)	121 (6.4)	0.96 (0.75 to 1.21)	
TT active	21 (3.2)	16 (2.3)	0.71 (0.37 to 1.38)	$z = -0.97; p = 0.25$
TT control	51 (8.5)	37 (9.0)	1.05 (0.68 to 1.64)	
Lean active	46 (10)	58 (10)	1.05 (0.70 to 1.58)	$z = 0.75; p = 0.30$
Lean control	72 (10)	29 (9)	0.83 (0.53 to 1.32)	
Lean + TT active	3 (2)	12 (9)	4.8 (1.32 to 17.4)	$z = 0.62; p = 0.33$
Lean + TT control	3 (1)	7 (2)	2.66 (0.68 to 10.4)	
SOP active	4 (1)	8 (2)	1.92 (0.57 to 6.46)	$z = 1.20; p = 0.19$
SOP control	9 (1)	4 (1)	0.68 (0.21 to 2.22)	
SOP + TT active	38 (7)	37 (5)	0.80 (0.50 to 1.28)	$z = -0.79; p = 0.29$
SOP + TT control	57 (10)	44 (10)	1.03 (0.68 to 1.57)	
Any TT active ^b	62 (5)	65 (4)	0.95 (0.67 to 1.36)	$z = -0.69; p = 0.31$
Lean active + SOP active	50 (6)	66 (6)	1.14 (0.78 to 1.67)	
Any lean active ^c	49 (8)	70 (10)	1.07 (0.81 to 1.40)	$z = 0.65; p = 0.32$
Any SOP active ^d	42 (4)	45 (4)	0.94 (0.71 to 1.25)	

a Adjusted odds ratio from a logistic regression including age and sex as covariates.

b TT active + lean TT active + SOP TT active.

c Lean active + lean TT active.

d SOP active + SOP TT active.

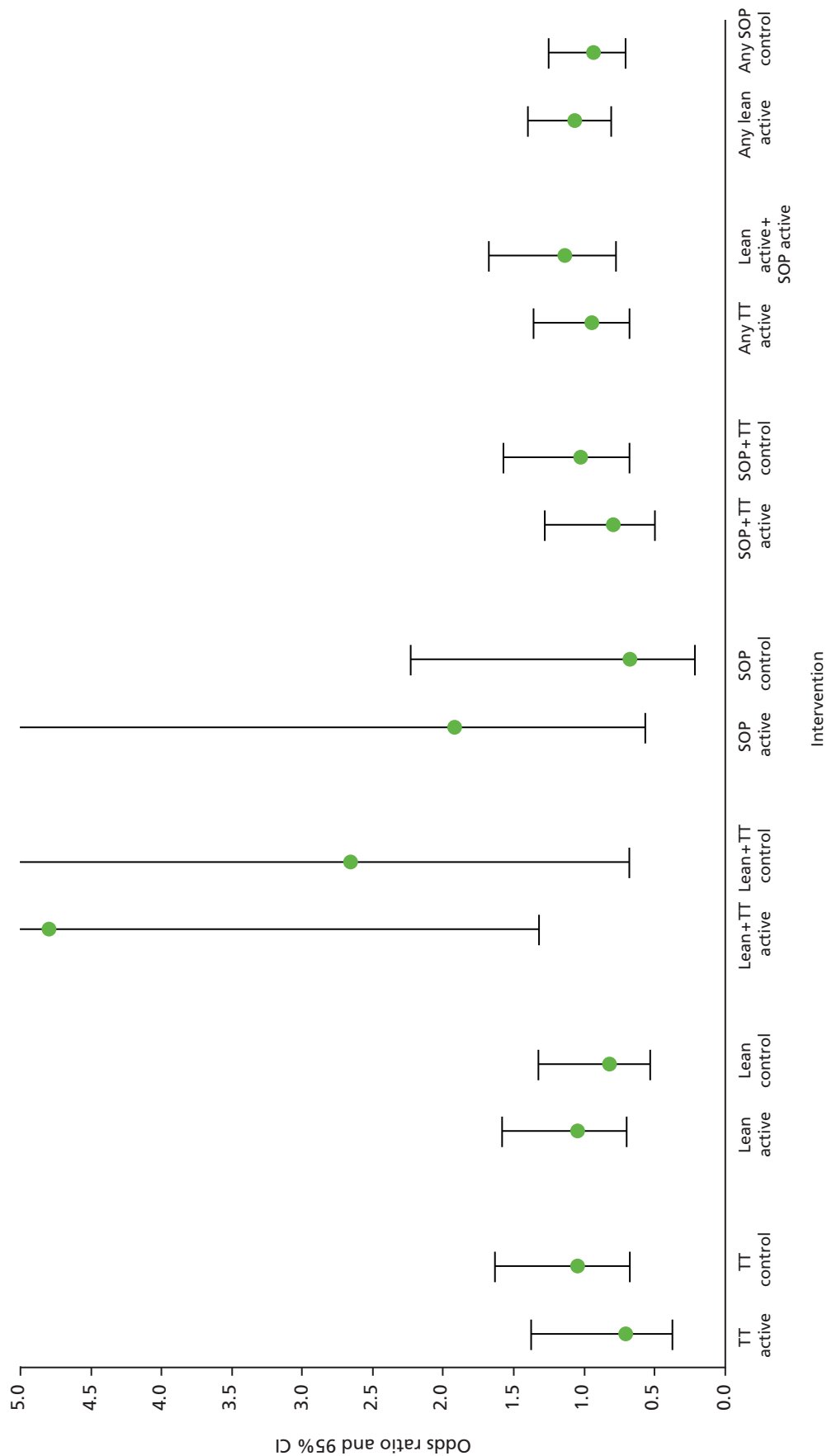


FIGURE 14 The before and after change adjusted odds ratio and associated 95% CI of patients readmitted within 90 days across the active and control groups.

TABLE 42 Frequency (percentage) of patients with at least one complication across interventions before and after change and in the active and control groups. The before and after comparison uses the adjusted odds ratio (95% CI)

Intervention	Before	After	Adjusted odds ratio (95% CI) ^a	Test for interaction active vs. control
Integrated active	299 (13.2)	422 (16.6)	1.32 (1.12 to 1.55)	$z = -1.01; p = 0.24$
Integrated control	453 (15.8)	379 (20.0)	1.34 (1.15 to 1.56)	
TT active	140 (21.5)	185 (26.8)	1.38 (1.06 to 1.80)	$z = 2.01; p = 0.05$
TT control	162 (27.1)	106 (25.7)	0.92 (0.69 to 1.24)	
Lean active	47 (10)	70 (12)	1.35 (0.91 to 2.01)	$z = 1.87; p = 0.07$
Lean control	95 (14)	32 (10)	0.77 (0.50 to 1.19)	
Lean + TT active	14 (9)	9 (7)	0.71 (0.29 to 1.69)	$z = -1.79; p = 0.08$
Lean + TT control	77 (18)	98 (28)	1.68 (1.19 to 2.36)	
SOP active	17 (4)	35 (7)	1.79 (0.98 to 3.26)	$z = -0.68; p = 0.32$
SOP control	61 (10)	79 (20)	2.28 (1.58 to 3.27)	
SOP + TT active	81 (14)	123 (18)	1.30 (0.95 to 1.76)	$z = -0.59; p = 0.33$
SOP + TT control	58 (11)	64 (15)	1.51 (1.03 to 2.23)	
Any TT active ^b	235 (17)	317 (21)	1.30 (1.07 to 1.57)	$z = -0.66; p = 0.32$
Lean active + SOP active	64 (7)	105 (10)	1.48 (1.06 to 2.06)	
Any lean active ^c	61 (10)	79 (11)	1.19 (0.97 to 1.47)	$z = -2.20; p = 0.04$
Any SOP active ^d	98 (10)	158 (14)	1.63 (1.35 to 1.97)	

a Adjusted odds ratio from a logistic regression including age and sex as covariates.

b TT active + lean TT active + SOP TT active.

c Lean active + lean TT active.

d SOP active + SOP TT active.

shows no difference in the integrated analysis, but does indicate a significantly higher complication rate in the SOP control group than in the SOP active group ($p = 0.001$).

In addition to readmissions and complications, it was possible to make similar comparisons for length of stay, and *Table 44* and *Figure 17* report these results. There is no evidence of statistically significant differences in any of the before and after comparisons or in the active and control group comparisons.

Discussion

A cost analysis of the teamwork interventions and their associated impact on hospital-related resource use at the patient level encompassed the health economics analysis conducted alongside the S3 study.

The costing of the training programme showed the TT programme to be more expensive than both the lean process and the SOPs intervention. The additional resources attributable to the teaching faculty and the in-theatre coaching explained the higher costs of the TT intervention. One of the main concerns of hospital managers when organising the training programmes was the substantial preparation time needed from some members of staff. In *Chapter 10* we presented the views of several hospital managers and described the difficulties they encountered in trying to release theatre teams for specific days in a given week to attend the courses. Our analysis suggested that only the TT needed input from consultants and nurses before the actual course. This was not the case for lean and SOPs, which needed more research staff preparation time than non-research staff.

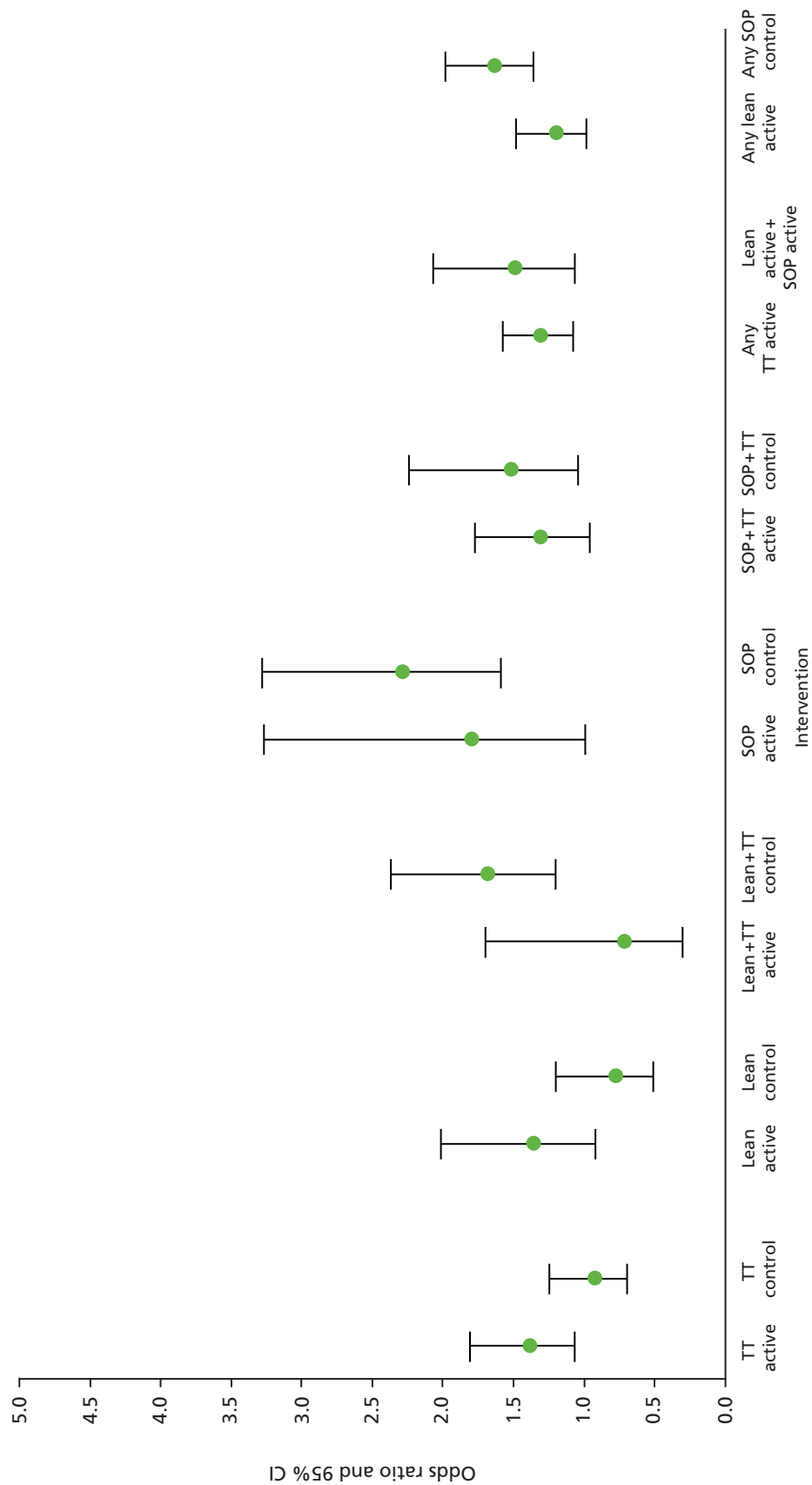


FIGURE 15 The before and after change adjusted odds ratio and associated 95% CI of patients with at least one complication across the active and control groups.

TABLE 43 The mean (SD, minimum, maximum) frequency of complications before and after change and in the active and control groups. The before and after change mean difference (95% CI) was adjusted for age and sex using linear regression

Intervention	Before		After		Adjusted mean difference (95% CI)	Test for interaction active vs. control
	Mean (SD)	Minimum/maximum	Mean (SD)	Minimum/maximum		
Integrated active	0.14 (0.46)	0/5	0.17 (0.47)	0/4	0.03 (0.01 to 0.06)	$z = -1.57; p = 0.117$
Integrated control	0.18 (0.49)	0/5	0.24 (0.58)	0/5	0.06 (0.03 to 0.09)	
TT active	0.25 (0.69)	0/5	0.28 (0.68)	0/4	0.04 (-0.03 to 0.11)	$z = 0.48; p = 0.355$
TT control	0.35 (0.76)	0/5	0.36 (0.86)	0/5	0.01 (-0.09 to 0.11)	
Lean active	0.11 (0.33)	0/2	0.13 (0.4)	0/2	0.03 (-0.01 to 0.07)	$z = -1.84; p = 0.074$
Lean control	0.15 (0.37)	0/2	0.11 (0.32)	0/2	-0.03 (-0.08 to 0.02)	
Lean + TT active	0.093 (0.3)	0/1	0.096 (0.4)	0/3	0.005 (-0.08 to 0.09)	$z = -1.69; p = 0.095$
Lean + TT control	0.21 (0.45)	0/2	0.31 (0.54)	0/3	0.10 (0.03 to 0.17)	
SOP active	0.041 (0.2)	0/1	0.074 (0.3)	0/1	0.03 (0.0001 to 0.06)	$z = -3.36; p = 0.001$
SOP control	0.11 (0.34)	0/3	0.23 (0.5)	0/3	0.13 (0.08 to 0.18)	
SOP + TT active	0.14 (0.36)	0/2	0.18 (0.39)	0/2	0.03 (-0.01 to 0.08)	$z = -0.30; p = 0.381$
SOP + TT control	0.12 (0.36)	0/3	0.16 (0.4)	0/2	0.04 (-0.003 to 0.09)	
Any TT active ^a	0.19 (0.56)	0/5	0.22 (0.54)	0/4	0.03 (-0.006 to 0.07)	$z = -0.00; p = 0.399$
Lean active + SOP active	0.76 (0.28)	0/2	0.11 (0.33)	0/2	0.03 (0.005 to 0.06)	
Any lean active ^b	0.10 (0.32)	0/2	0.13 (0.38)	0/3	0.03 (-0.01 to 0.06)	$z = -0.00; p = 0.399$
Any SOP active ^c	0.10 (0.30)	0/2	0.14 (0.35)	0/2	0.03 (0.006 to 0.06)	

SD, standard deviation.
a TT active + lean TT active + SOP TT active.
b Lean active + lean TT active.
c SOP active + SOP TT active.

Given our estimates, the implementation of any of the training programmes by a complete department at a specific hospital will have substantial costs. However, replicated nationally, these do not appear large in relation to sums made available in recent initiatives to improve patient safety, for example the £150M announced in 2013 by NHS England. Moreover, it is likely that, when distributed across large number of patients who would benefit from the new skills of the teams, these costs will be almost negligible for the NHS because of economies of scale.

The analysis of possible interactions between active and control groups for the hospital-related data showed that, in the case of number of readmissions and length of stay, there was no evidence of effect for the interventions. This was not the case for complications, as our results suggest that there were fewer complications in the lean plus TT and SOPs groups than in the control group, indicating that some of these interventions seemed to have an effect on outcomes. In addition, a significant interaction between any lean intervention and any SOP intervention was also detected, indicating that a lean process may yield better outcomes than SOP-related interventions. These results, however, need to be interpreted carefully and with caution by the decision-maker. The fact that we observed some evidence of effect in only one of the outcomes instead of all of them and the large number of comparisons made for the before and after change and in the active and control groups can suggest these results as spurious. The frequency of complications and number of readmissions are strongly correlated and both of them are strongly correlated

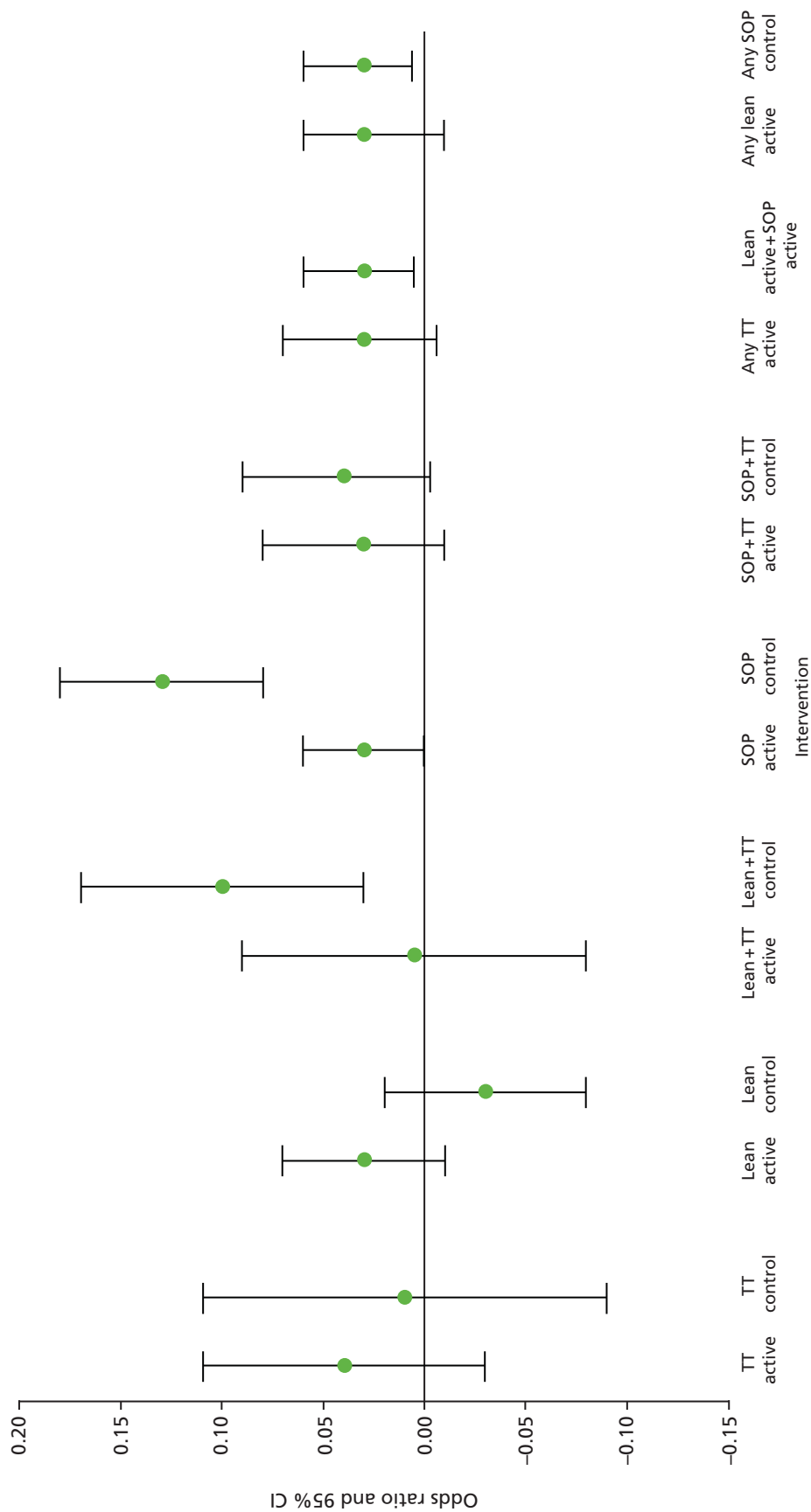


FIGURE 16 The before and after change adjusted mean difference and associated 95% CI of the frequency of complications across active and control groups.

TABLE 44 The mean (SD, minimum, maximum) length of stay (in days) across interventions before and after change and in the active and control groups. The before and after change mean difference (95% CI) was adjusted for age and sex using linear regression

Intervention	Before		After		Adjusted mean difference (95% CI)	Test for interaction active vs. control
	Mean (SD)	Minimum/maximum	Mean (SD)	Minimum/maximum		
Integrated active	5.0 (14.6)	0/232	4.4 (10.8)	0/165	-0.59 (-1.3 to 0.1)	$z = 0.02; p = 0.399$
Integrated control	7.5 (15.2)	0/232	6.9 (13.8)	0/185	-0.60 (-1.4 to 0.2)	
TT active	5.09 (11.1)	0/96	5.38 (13.2)	0/159	0.43 (-0.8 to 1.7)	$z = 0.38; p = 0.371$
TT control	4.82 (13.5)	0/113	4.93 (11.7)	0/102	0.04 (-1.5 to 1.6)	
Lean active	10.3 (25)	0/232	7.7 (15)	0/165	-2.17 (-4.5 to 0.2)	$z = -0.12; p = 0.396$
Lean control	10.2 (20)	0/232	7.6 (16)	0/185	-1.96 (-4.4 to 0.5)	
Lean + TT active	2.5 (6.3)	0/56	1.6 (4.0)	0/29	-0.86 (-2.1 to 0.4)	$z = -1.69; p = 0.095$
Lean + TT control	5.2 (7.2)	0/82	6.0 (7.9)	0/54	0.55 (-0.5 to 1.6)	
SOP active	1.7 (4.3)	0/45	1.8 (3.7)	0/31	0.063 (-0.4 to 0.5)	$z = 0.43; p = 0.363$
SOP control	5.7 (8.4)	0/79	5.4 (7.6)	0/82	-0.18 (-1.2 to 0.8)	
SOP + TT active	3.6 (11)	0/185	3.0 (7.2)	0/86	-0.74 (-1.7 to 0.3)	$z = -0.34; p = 0.376$
SOP + TT control	10.6 (19)	0/149	10.5 (20)	0/135	-0.30 (-2.6 to 2.0)	
Any TT active ^a	4.1 (11)	0/185	4.0 (10)	0/165	-0.2 (-0.95 to 0.52)	$z = 1.29; p = 0.173$
Lean active + SOP active	6.3 (19)	0/232	5.0 (11)	0/165	-1.19 (-2.5 to 0.12)	
Any lean active ^b	8.4 (22)	0/232	6.5 (13)	0/165	-1.59 (-3.46 to 0.28)	$z = -1.19; p = 0.195$
Any SOP active ^c	2.7 (9)	0/185	2.5 (6)	0/86	-0.39 (-1.00 to 0.23)	

SD, standard deviation.
a TT active + lean TT active + SOP TT active.
b Lean active + lean TT active.
c SOP active + SOP TT active.

with length of stay. Therefore, solid evidence of a possible treatment effect would have affected all outcomes and not only complications.

There are some limitations in our analysis. Given the complexity of the study design it was not possible to select a control group based on matching methods and the control was selected based on similarity of interventions observed in orthopaedic theatre. Strictly speaking, active and control groups may have been subject to some differences but patient characteristics for age and sex were balanced in our analyses. Standard test of interactions have also received some criticism in the literature because of their lack of power to detect differences in small and medium sample sizes, but lack of power should have a minor effect in our large sample of anonymised individual patient-level data.

In summary, in this chapter we have reported cost estimates for the training programmes and have provided preliminary indications of a possible effect of some of the interventions on patient-level hospital-related data. Nevertheless, the strength of this effect is currently small and further research is needed to clarify whether or not the results reported in this chapter are real. Finally, future research should also assess the impact of the intervention on the patient's quality of life that can be combined with the treatment costs reported here and any additional necessary cost aspects to evaluate formally the cost-effectiveness of interventions that aim to improve patient safety in surgical theatres.

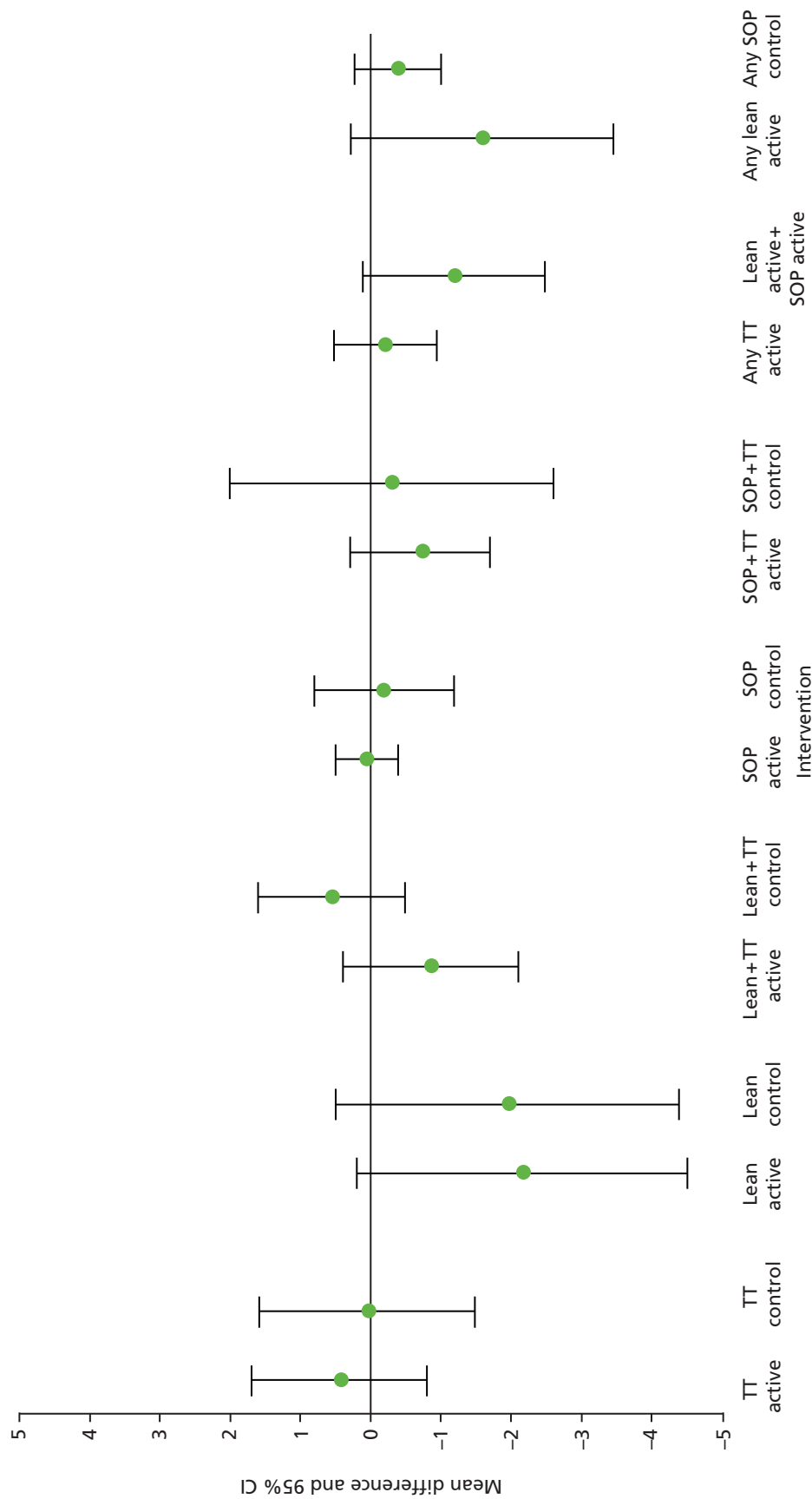


FIGURE 17 Before and after change adjusted mean difference and associated 95% CI of length of stay days across active and control groups.

Chapter 12 Final synthesis and analysis

We set out to develop and refine the tools needed for an examination of team performance in operating theatres and then to use the tools in a set of studies designed to explore the implications of our 3D model of health-care error and harm. This led us to conduct studies to refine and validate the glitch count and the Oxford NOTECHS II methods, and then to carry out our series of identical intervention experiments. We then utilised the learning from these to address a bigger question: do the approaches which appear most successful in the confines of an operating theatre also work effectively once the target is enlarged to encompass an entire clinical pathway in hospital? This led us to develop our suite of studies in the Oxford neurosurgery unit. Qualitative researchers within the group contributed important qualitative information about why things worked, or did not, available from interviews with clinical staff and researchers.

Measurement

We set out to refine our measures of teamwork and team effectiveness in ways which would help us use them in a wider variety of settings and we decided to make the existing scale as responsive and reliable as possible. Our work on the Oxford NOTECHS II scale achieved this to a considerable degree and we believe that this now represents the strongest available measure of the non-technical skills performance of an entire theatre team. It nonetheless still has imperfections. Despite our attempts to increase the sensitivity of the scale to smaller changes in team interactions, most operations continue to produce a summary score within a relatively narrow range, somewhere between 70 and 75 on the 0–96 scale. The scale is a very rich information source, recording in the subscale results the performance on the four dimensions of each of the three subteams, and there appears to be much potential for further analysis of how this information could best be presented and analysed, particularly with reference to the effect of different groups, dimensions or extremes of performance on technical and clinical outcome.

We did not originally plan to revise our scale for technical performance, but when the larger team of researchers began training with the non-operative procedural errors system (NOPES), it quickly became clear that a revision would be required. There was considerable controversy within the team about the definitions used to categorise events in NOPES and about the technicalities of how the scale was, and how it ought to be developed. New team members found it difficult to use the method because they were uncertain how to classify some common process imperfections. An attempt to produce consistency through a training programme directed by Dr Catchpole was unsuccessful, so it was decided to review the categories of process deviations using a battery of observations we had already collected and a consensus methodology. The resulting 'glitch' method appeared intuitive and validity and reliability testing were satisfactory. However, it was clearly different from the preceding NOPES method in a number of important ways. Most importantly, whereas NOPES correlated quite strongly with the Oxford NOTECHS II scale,²⁶ the glitch rate did not show any obvious correlation with the NOTECHS II scale. The scale therefore needs to be considered as quite different from NOPES, despite their similarities. The lack of correlation is counterintuitive at a superficial level; as we ourselves argued in our earlier work, one would expect teams with a better teamwork performance to make fewer mistakes. However, this prediction ignores the complexity of real-life operating theatre work and the relationships between teamwork and error. When a major error has occurred, for example, teams may have the opportunity to demonstrate excellent teamwork skills in dealing with it, which might not have been evident if the error had not happened. We therefore feel reasonably comfortable with the idea that technical error and teamwork quality are not necessarily correlated in a linear fashion. We did not initially plan to evaluate the WHO's surgical safety checklist as a separate component of our study; our plan was to incorporate evaluation of performance of this checklist within our observational measures of technical and non-technical performance. However, the contemporary importance of the checklist became so prominent that we felt it was essential to give readers of our work an idea of how performance of this mandatory safety measure correlated with our other measures. We were able to develop a unique and simple method for evaluating checklist performance which we have since published.²³ In doing so we highlighted some

uncomfortable truths about compliance with the checklist, which resonate with subsequent publications.^{84,85} We showed that real effective compliance, as opposed to recorded compliance, with the checklist T/O procedure averages around 38% of operations, with a very wide range, whereas real compliance with the final S/O was even lower. S/O effective compliance was so low that we concluded it was not being done because it had not been appropriately designed around the actual work patterns of theatre teams, or that the teams had not had enough time or training to adapt these to allow them to perform it. These findings are important and were possible because of the development of our WHO checklist evaluation tool.

Interventions

Our central purpose during the study was to examine the relative effects of interventions to improve process and outcomes, and to determine whether or not there could be synergy between them. We found in our parallel operating theatre studies that TT interventions always induced an improvement in team non-technical skills. As the WHO checklist is so intimately related to team culture, it was not surprising to note that TT also nearly always brought about some improvement in how the T/O and S/O procedures were performed. However, TT itself did not seem to affect technical performance, and for the study in which it was used alone, it was actually associated with deterioration in this. This combination of findings is in contrast to the findings of our initial intervention study on CRM in 2009, but is consistent with the results of our systematic review of studies of CRM/TT internationally. It suggests that TT, while giving clinical staff a better understanding of safety issues, and improving the safety culture in the workplace, may fail to bring about measurable improvement in technical performance and outcome because staff are given motivation but not the practical skills to bring about change. As has been extensively documented, incorporating apparently simple safety measures into working routines, such as the WHO checklist, is in actual practice challenging.⁸⁶ Part of the process recommended for checklist adoption optimisation involves staff training which covers very much the same ground as TT. However, another important part relates to codesigning systems change with participants using QI methods, and there is accumulating evidence that failure to do this results in low compliance with the checklist and clinical ineffectiveness. It therefore appears from our work that TT alone may result in motivated but frustrated staff who recognise the safety defects in their work systems but are not empowered to change them.

We found that the ergonomic approach to systems improvement, which we labelled SOPs, was ineffective when we tried it as a single intervention. It did not cause an improvement in any of the measures of process we had chosen to study. We had very similar findings when we used the lean process improvement method as a single intervention. In both cases, teams decided to concentrate their efforts on projects which we deemed unlikely from the outset to have much effect on our chosen evaluation measures. In one case, they decided to revise their visual guide (the white board) to the work they had to do. In the other, they decided to focus on ensuring that patients arrived in emergency trauma theatre in a timely fashion. None of these was entirely irrelevant to the aim of producing a better-functioning team that was less prone to error, but neither focused directly on it. This experience was important because it led us to realise the importance of the tension between control and engagement when attempting to keep operating theatre teams (or any other staff group) active and positive in pursuit of an agreed goal. The more external direction is supplied, the less motivated staff become, since they come to see the intervention as external force which reduces their autonomy.

We do not deny that the nature of the team and indeed other context factors such as trust management engagement did play a role, but we rely on two features of our experimental design to argue that we may nonetheless have some confidence that our results are meaningful. First, the use of these parallel studies has allowed us to investigate whether or not there are consistencies in effect for specific interventions across studies. Second, the use of parallel control groups in the same trust in every case has protected us to a considerable extent against being misled either in a positive or a negative way by secular trends and random 'noise'.

The amount of resource expended and the amount of training time provided were both significantly less than would be normal in many other sectors. Each intervention involved not just the classroom days, but also intensive work by a skilled team of experts in the relevant areas (QI, HFs, systems redesign and clinical expertise) working for several months with a relatively small staff group, and using principles and practices based on the best evidence available from current scientific literature. This was therefore neither a superficial nor a cheap intervention method and if it was not effective because it was not intensive enough, this must raise questions about the economic feasibility of disseminating a method which is considered sufficiently intensive to entire health systems.

Combining TT with either the SOPs or the lean approach to systems improvement proved more successful than any of the single-intervention approaches. The TT/SOPs combination resulted in significant improvements in non-technical skills and in compliance with the WHO checklist procedures. The combination of TT and lean was the most successful of all, combining improvements in NOTECHS scale score and WHO compliance, with a marked improvement in technical performance. When we performed a pooled analysis of the studies comparing integrated approaches with those using a single approach, the integrated approach was significantly superior in terms of the quality of the WHO checklist and the quality of technical performance (glitch count). Pooled analysis comparing the systems approaches with the data from the study of TT alone showed that TT was better at improving the frequency with which T/O was performed, but the systems approaches were superior in terms of improving S/O and glitch count. Given that the (TT alone) study was small, it is unwise to put too much weight on these findings, but it is consistent with the logic of the situation as we understand it: the TT training will tend to improve the culture (and, therefore, the frequency with which the WHO checklist is followed), while the systems interventions are more likely to improve technical performance by addressing imperfections in the manner in which tasks are performed. Our qualitative studies supported the interpretation suggested by our original hypothesis,²⁴ namely that an integrated intervention allowed staff to understand and accept the safety agenda, and that this helped focus their attention on genuinely safety-enhancing goals when they were challenged to use the systems improvement skills they had acquired.

None of the interventions resulted in a significant change in the clinical outcomes we measured, which was not unexpected, given the small numbers involved in each study. As discussed elsewhere, the trends noted in the outcome measures generally supported the findings of the process measure studies (WHO, glitch and NOTECHS), moving consistently in the same direction.

Implementation strategy

Our experiences conducting these studies convinced us of the need for a well-planned implementation strategy to maximise the uptake and impact of the interventions we planned. Getting the intervention to happen as planned was a major challenge in every study and this experience has been reported by others as outlined in earlier chapters. Our qualitative work looked into the context features which contributed to this 'institutional resistance' and is reported in *Chapter 10*. In order to counter it and achieve maximum intervention effect, we developed a strategy for implementation that addressed some of the barriers we identified. Thought needs to be given to how the strategy can be empirically validated in the future (see *Chapters 12 and 13*). The components of our final strategy comprised:

- communication and engagement
 - with senior leaders: to get support in principle and have this made public
 - with clinical frontline staff: to explain and reassure; to involve early in analysis
 - with middle management: to identify ways in which goals can be aligned
 - with senior clinicians: to persuade that they should lead systems improvement

- recruitment of key clinical leaders
 - identification of individuals with group authority who conform to Rogers' 'early adopter' profile⁷⁷
 - careful consideration of how individual motivation and incentives are likely to play into engagement with leading the project
 - individual discussion and offer
- establishment of strong relationships with leaders and staff
 - early intensive engagement
 - strong efforts to level the playing field and encourage dialogue
 - use of socialising events to break down barriers and establish rapport
- strong staff engagement, encouraging experimentation and leadership of change
 - early devolution of responsibility to staff leaders, with initially strong coaching and support
- ensuring new systems improve working lives
 - consider for every change whether workloads are increased or decreased
 - discuss with frontline staff how to favour the former
 - use PDCA cycles to test the water with staff
- ensuring new systems harmonise with staff cultural beliefs
 - discussion with staff before and after each change
 - use of content experts to predict whether or not staff will accept change comfortably
- ensuring success becomes visible quickly
 - identify targets for improvement with particular importance for staff and which appear likely to be straightforward to change
 - devise simple measures to demonstrate change
- developing crosslinks between units to form a community for change
 - organise meetings, video links and newsletters to share experiences and encourage mutual learning
 - ensure a controlled element of competition by presentation of successes.

Critique

Our study design was intended to protect against effects on the outcome variables arising from secular trends and sudden externally imposed changes in environment or practice, both of which are commonplace in the dynamic environment of a working hospital. Uncontrolled studies are highly vulnerable to such effects and a great deal of the literature on improvement work in hospitals is difficult to interpret because of this. The value of the control groups was illustrated on several occasions, most notably when hospital management decisions on accelerated discharge caused a dramatic change in length of stay. An additional potential benefit of the control groups was to allow to a certain extent for the Hawthorne effect of having work observed. Theatre teams might of course react differently to observers depending on whether they were in an active or a control group, but subjective impressions and the results of interviews and discussions suggest that in practice the reaction was similarly small in both groups. The effect of the control groups on conclusions about intervention effects are illustrated by the

difference between the number of occasions when an outcome measure improved significantly after the intervention¹² and the number for which the measure improved significantly more in the intervention than in the control group.⁸ This damping effect is appropriate and helpful, as falsely optimistic conclusions based on inadequate evidence have been one of the key weaknesses of the science of quality and safety in health care since its inception.

Although we consider them essential, we recognise that control groups in this kind of work are difficult to organise and are never perfect. Contamination is a constant risk: in many situations there is an exchange or flow of staff members between groups, but even when this does not happen, the risk of contamination via informal learning among peer groups is important. In a sense, contamination is less dangerous to validity than some other forms of bias, as it acts to reduce any observed net intervention effect; that is to say it increases the risks of a false-negative result and decreases those of a false-positive result. The more the intervention relies on specific constructs that are difficult to replicate without help, the less serious this threat becomes.

We could not investigate multiple options for improvement side by side in large-scale comparative studies with clinical primary outcome measures in this programme: the number of patients needed would have been prohibitive. We therefore chose to conduct smaller-scale studies with rich observation of process that would allow us to evaluate team performance as the primary outcome to be changed by intervention. The integrated data from the five studies represent, as far as we are aware, the largest direct observation study of theatre process ever attempted. As we were interested in determining how interventions affected process as much as whether or not they could be shown to improve outcome, this was appropriate as a strategy but meant that our studies were inevitably underpowered to detect clinical outcome improvements, which are always of greater interest to a clinical audience. Ultimately, the justification for any methodology has to be whether or not it was effective in providing a valid test of the hypothesis studies. We feel the suite of five identical intervention studies did this and that important information was generated as a result.

Whether or not our choice of direct continuous observation as our principal study method was appropriate is also worth considering. It is difficult to eliminate the risk of observer bias using any method requiring humans to observe behaviour and make judgements about it, but teamwork and technical performance can only be studied in detail by continuous observation, either directly or recorded. Our preliminary work with video recordings convinced us of the difficulties inherent in using these in a live theatre situation. Health-care staff concerns about video recordings (which they feared might be retained and used for disciplinary or medico-legal purposes) were much greater than those they expressed about having observers in theatre and video recording in a live theatre often results in a poor-quality recording which is difficult to interpret afterwards. We therefore think that our chosen solution was reasonable but far from perfect. We deliberately paired up observers with HFs and surgical backgrounds. This worked well and some of our validation work on glitches, for example, shows the significant differences in the starting points of the two observer groups. However, it was very evident that adaptation to the situation took place quickly. Observers learned from each other and quickly reached a stable state where the scoring done independently by the two observers in a pair was very similar. This of course does not eliminate the possibility of observer bias as a result of lack of blinding, the most serious threat to the validity of our findings. This cannot be completely discounted as an explanation for some of our positive findings, but there are a number of reasons for believing that its role may not have been important. Although observers were aware of the study hypothesis, it was not emphasised in team discussions, and we expected to see some improvement from all of the interventions. The fact that several of the studies showed no improvement against control groups, and, in some cases, important measures actually worsened in the intervention group after the intervention, argues against a systematic observer bias in favour of post-intervention scores or active groups. The faithful recording of secular trend effects, which were not part of the expectation of any of the observers (e.g. the increased glitch count post intervention in the TT study), also argues against a strong observer bias effect. Finally, the notable association of positive movements in the clinical outcome measures with movements in the same direction for process measures strongly suggests that the process results are not a result of observer bias. Our outcome data were produced for us by NHS clerical staff who

had no knowledge of the study whatsoever and can therefore be regarded as truly blinded. It is true that none of the clinical outcome effects reached statistical significance (as expected), but the strong association in terms of direction fulfils one of the Bradford Hill criteria for causal inference.⁸⁷

One final critique on our methods concerns the clear mismatch, in some studies, between the intervention undertaken and the outcome measures used. The lean study was perhaps the starkest example; here the team decided to improve efficiency of patient transfer rather than intraoperative process and the lack of any impact of their efforts on our measures of team performance during the actual operation appeared highly predictable. This problem arose because of our strict adherence to a policy of non-directive encouragement of teams in designing their projects to address what they regarded as the most serious safety problem. We were concerned from the outset at the direction the team took, but were unable to interfere because of our own policy. We drew two conclusions from this experience. First, our policy needed to be modified so as to achieve a balance between investigator/intervention lead objectives and team objectives, rather than giving teams complete freedom of action. How to do this without demotivating teams by making them feel that the objectives and projects have been externally imposed remains a difficult challenge. Second, collection of specific secondary outcome data which best reflect the effectiveness of the intervention chosen is essential if staff autonomy is encouraged in such a way as to produce heterogeneity in the projects carried out.

Chapter 13 Learning and development

Introduction

The S3 study provides a number of points for learning and development for a variety of audiences, including clinical staff, NHS managers, academics and practitioners in relevant disciplines such as ergonomics, patients and the general public. These are discussed in this summary and are addressed in terms of what we have learned (both from our results and from the experiential learning involved in conducting the study) and which groups might benefit from the learning.

What have we learned?

Engagement

A key factor in the success of the improvement work (when it was successful) was the active engagement of health-care staff fulfilling a range of clinical roles. Having a champion or a key point of contact is essential for the success and maintenance of improvement efforts, and the profile of the most effective champions is well described by Everett Rogers'⁷⁷ description of the 'early adopter' of new innovations. However, efforts to engage with the most likely candidates were often unsuccessful. The reason may lie in Rogers' description of an early adopter as someone who is better educated, wealthier and more cosmopolitan than his or her peers, but adheres strictly to the conventions and values of society. In the context of a modern hospital, such people (senior clinical staff and managers, for example) often have all that they want in their job and, therefore, lack motivation to change a system that affords them status, comfort and satisfaction. There was some support for this proposition in the qualitative findings. This may be why other staff whose enthusiasm compensated, to some extent, for their lack of status in the existing working culture often became the drivers for individual projects. However, as outlined in *Chapter 10*, the engagement of frontline staff, essential as it is, can be challenging. Staff in health-care environments are generally strongly motivated to deliver a good service for altruistic reasons and as a consequence of their professional persona and self-image. They are driven by a system which is overtly austere, constantly aware of the need to do as much as possible in the most efficient way and to contain costs. Coupled with the ever-increasing demand for health care in a consumerist and ageing society, this results in a situation in which they are generally working beyond their capacity and QI is not their top priority. While many may applaud the objectives, most tend to feel that it is not possible to assign the necessary time and effort to the required work when 'firefighting' the urgent problems which confront them in dealing with individual patients absorbs all their energies. Other barriers to engagement are evident in our experiences and have been reported frequently by others. The experience of working in NHS hospitals leads to the development of shared outlooks among staff groups based on their experience of work. Two factors that were important in our studies were learned helplessness and change fatigue. The former represented a natural reaction to previous experience of attempting to bring about change in their own environment and failing. After several such experiences, staff commonly adopt the attitude that effort expended in trying to change the current order of things is a waste of energy and can only lead to disappointment and frustration. The latter refers to the rather different but equally widespread experience of having change imposed on them by top-down initiatives. These often result in extra work for frontline staff and their perception is that they rarely, if ever, result in real beneficial change in terms of either working lives or improvement for patients. Combined with understandable cynicism about the motivation of external consultants, who often drive these initiatives, this negative experience can result in a formidable barrier to engagement. We quickly learned that it was unwise, for example, to label our systems improvement initiative with the term *lean*, even though this was the origin of the methodology, because (as noted in *Chapter 10*) so many staff members had had negative experiences of 'being leaned' in the past, usually with a view to increasing efficiency rather than directly focusing on improving patient care. Conversely, in units which had a good reputation locally we encountered some staff

who demonstrated significant levels of complacency about their performance and openly expressed doubts about whether or not the service they provided could realistically be improved on. The braking effect of this attitude on engagement in improvement was best demonstrated when reflecting on conversations with staff in such units, alongside similar conversations with staff whose trusts had less prestige or which had been publicly identified as having problems. The greater willingness of staff in the latter institutions to take a fresh look at their own performance and processes was striking. A final, universal and important obstacle to engagement was a lack of time and space for improvement work. Hospitals do not generally abound with suitable environments for group meetings in clinical areas and, more importantly, scheduling of such meetings so that all staff can attend was virtually impossible in nearly every study. This led to a very pressurised atmosphere and strained communications, as fragments of the project team attempted to make progress and contact those who could not attend to ask for their support and update them on what needed to be done. The contrast with normal practice in high-performing commercial and high-reliability public organisations is striking. A regular allocation of staff time, built into job plans, to allow improvement work, is standard in such organisations and would undoubtedly be a tremendous boost to the capacity of the NHS for self-improvement if it could be generally adopted in a fashion sufficiently flexible to allow staff to time their allocation so as to facilitate co-operation on local projects. Addressing such challenges is clearly difficult; however, strategies such as developing good rapport between the research staff and the clinical team, producing evidence of success in other projects by presenting data from previous work, and the provision of practical support helped us to overcome many of these.

Sustainability

Sustainability of improvement is one of the biggest challenges in health-care improvement. Staff engagement and ownership of change is clearly a major factor in encouraging this, but the dynamic nature of hospitals makes it highly likely that even changes which have a good deal of staff ownership are liable to be eroded with time unless they clearly result in systems that staff find easier to use or less stressful to negotiate. Our experiences in the S3 study suggested that the two other key factors that ensure long-term sustainability is more likely to occur are the instigation of changes that are too big or too radical to be reversed without enormous effort (so called disruptive change) or changes which occur when (a) stakeholder(s) has/have invested a lot in the project in terms of time, effort or personal prestige and reputation. The latter influence obviously implies that, when interventions were dependent for their success and sustainability on key individuals, they are inherently unstable, as there is a high risk such individuals could move organisation or role.

Staff-led improvement

Although we have repeatedly emphasised the benefits of the staff-led approach and the high likelihood of failure if staff cannot be actively engaged, staff leadership also throws up significant problems. One of these is that effort put into persuading staff to renounce learned helplessness by emphasising their freedom to improve whatever they feel most strongly needs reform in their environment can easily backfire. Staff may (and in at least two of our studies, did) focus on areas for improvement that are difficult to link to patient safety or to measure effectively. We have adapted our language in important ways in our approach to staff in the light of these experiences. We still emphasise their ability to bring about real change in their environment and the capacity they have to make that change beneficial for their own working lives, but we emphasise strongly the overall objective of improving patient care, and talk about 'win-win' solutions in which this can be accomplished while helping staff with some of their problems at work at the same time. We also feel that one of the major benefits of the TT part of our programme was to focus staff attention on this as their overriding concern and to help them to think in terms of change that would primarily benefit patients. Whereas we initially presented the staff with a blank page in discussion of improvement projects, we have moved to a strategy of framing the problems and thereby the change to be considered more narrowly. In this regard we have moved significantly, but not completely, towards the approach championed by organisations such as the Institute of Healthcare Improvement and Advancing Quality Alliance, in which the objectives of a project and the main interventions to achieve them are predefined and not open to change by the staff. Although we now try to identify the problem fairly precisely, and we do provide evidence and arguments for certain strategies,

we retain willingness to modify or add to the programme of work depending on staff input. An allied problem of major importance is that, although staff are familiar with the clinical working environment, they do not possess the in-depth expertise to carry out improvement, making them reliant on teams such as ours for support. If training in systems improvement and HFs were generally available to staff this would reduce this problem, but it is likely that to achieve lasting success in major projects some external professional expertise will be needed. Trusts need to recognise that a core team of systems improvement professionals is needed to make them more capable of reacting appropriately when it becomes apparent that their systems need reform.

Single compared with integrated interventions

The aim of this programme was to examine the hypothesis that interventions targeting teamwork and those targeting systems reform were synergistic. We have provided significant evidence in support of this hypothesis and in this sense the principal objective has been achieved. It is unfortunate that we were not able to demonstrate clinical benefit, but this was understandable given the study design and the myriad factors that affect clinical outcome apart from theatre team performance. The results form a solid base from which to argue for a larger trial using clinical outcome measures, but the obstacle we perceive to this at present is the need to demonstrate that our programme can be delivered in a scaled-up form which will allow us to achieve comparable results with a much smaller ratio of research/support staff to clinical staff. A study which reproduced the current model would be unfeasibly expensive and liable to produce results that could not be translated into clinical practice because of cost issues. Another aspect of our original model of patient safety is also worthy of consideration. We identified the technology used as another major dimension in risk and safety in clinical environments, but these studies have not dealt with this at all. Further studies to confirm the significance of technology in both engendering and preventing risk to patients would be valuable, and following these further studies to determine whether or not a 3D intervention approach taking into account the need for changes in technology would be more effective than the two-dimensional approach we are currently advocating. The results of the S3 study suggest that mono-component interventions are unlikely to yield great results, possibly because of the fact that such approaches do not tackle all elements of work systems. The integrated studies within the S3 study have indicated that there is likely to be more potential for improvement when utilising an integrated approach and, therefore, targeting more than one aspect of work systems.

How have we developed?

The S3 study team has developed an integrated approach to QI, which has demonstrated improvements both in the operating theatre (lean and TT study) and also across a number of ward-based QI projects (lean, SOP and TT study). The programme has provided descriptions of each study in detail. This is not to be taken for granted, as it is something that is often frequently omitted in the literature, preventing replication, or at the very least an understanding of how the intervention was actually implemented in context. In addition to providing descriptions of each intervention, the S3 study has provided a qualitative insight, further shedding light on the practicalities of the interventions and exploring additional factors which may have interacted with the intervention. In particular, it has examined the role of the research team; the support it provided and the strategies it employed to facilitate the interventions. Through this, the S3 study identified types of support as particularly useful for staff and also strategies that were effective in overcoming barriers to QI. Types of research team support identified as useful included guidance, co-ordination, administration, communication and linking between groups, training, data collection and analysis, motivation and feedback. Strategies utilised by the research team to overcome barriers included building rapport, facilitating, producing evidence and data, setting goals and encouraging a flat hierarchy. We also developed a pragmatic strategy around the development of relationships with management at all levels and the judicious use of these when barriers arose that would otherwise have been insurmountable. These developments and lessons from the S3 study are extremely important for replication and upscaling of this work, and also for further developing QI science in health care. Lessons based on the learning from the S3 study programme are outlined in the following sections.

Lessons: for researchers studying safety and quality intervention strategies

Outlined in the following text are a number of strategies and lessons for other academics looking to engage in similar ventures.

External team status

Staff generally tend not to approve of external management consultants as a result of previous negative experience which transmutes itself into a generally hostile cultural attitude. Having a research title and status that makes it clear that the external team has few ties to the NHS organisation for which staff work and having no commercial incentive has proven to be extremely helpful. Working with members of a research group is much more acceptable to frontline staff than either implementing 'top-down' change imposed by their own management or being required to assist a paid external consultant with a short-term improvement brief. It is therefore extremely important that researchers working in this context make their role very clear, outlining (1) their specific research aims and how staff can be involved and (2) how they are separate to the organisation and will not be reporting back to hospital management on any issues relating to individual staff.

Maintenance of objectivity in evaluation

Academic teams are not usually large enough to allow the separation of evaluation from intervention duties, which is desirable. In order to retain as much internal validity as possible in the evaluation of an intervention project, the evaluation strategy needs to be thought out in advance, giving appropriate consideration to the resources available and the practical constraints of the situation. Intervention projects are, by definition, complex and should therefore be evaluated using a range of independent approaches that complement each other, allow triangulation, and explain different aspects of the process and outcome of the intervention. To minimise effects from the inevitable close interaction between members of the research team and the subject being observed, some of the following tactics may prove useful.

Use of hospital administrative data: although this may have accuracy and completeness issues and be subject to biases related to the method of collection, this is unlikely to be subject to biases related to any expectations of the performance of the intervention. It is therefore particularly valuable in comparing process, and outcome before and after the intervention and should be used as a major component of the evaluation strategy.

Creation of PDCA measures: whenever possible these should be objective and simple, not requiring any judgement on the part of the observer. For example, nurses may be asked to make a mark on a sheet for every time they are interrupted during a medication round.

Direct observation techniques: complex evaluations of process requiring the use of trained observers to conduct relatively prolonged observations of process need to be designed so that they are semiobjective, using a scoring scheme closely tied to explicit behavioural markers. The use of independent dual observers is costly but a useful guard against individual observer bias. If possible, teammates should be allocated to score teams or subteams with whom they are not familiar.

Prospective audits: data collection schemes for evaluating outcomes and processes need to be discussed in advance with the clinical team and designed so that judgement required is minimal. This involves the use of very clear explicit definitions of what is being studied or sought. When possible, independent confirmation of findings by a second observer is desirable.

Practicalities of data collection and evaluation

The requirements of data collection for research purposes are considerably greater than for service improvement projects and are subject to more stringent regulation. Given the need to ensure objectivity, accuracy and completeness in collection of all data (with the exception of PDCA data for which the

requirement for such stringency can be relaxed) it is necessary to develop a well-organised system for data collection, cleaning and storage that is as near real time as possible. This will require a significant amount of dedicated time from either a specific member of the team or a small subteam sharing the duties. The pressures of intervention and data collection tend to overshadow psychologically the equally important task of ensuring that the data are complete and accurate, and this should be recognised and problems forestalled by insisting on a disciplined approach to this vital task from the outset.

Logistics and communication

It is very important that the research team retains a coherent sense of purpose and does not get so absorbed in the work of developing individual projects that objectivity is compromised or effort for other necessary projects is sacrificed. This requires excellent time management and regular team briefings and discussions to retain a communal understanding of the process and problems of the programme as a whole and where the individual parts of it fit. Development of a project management system which supports group understanding, while not imposing a requirement for too many meetings, is an important component of a successful implementation strategy for this kind of intervention research project.

Lessons: for independent quality improvement groups

Organisational context

Practitioners involved in running improvement interventions with frontline staff need to be aware of the organisational context in order to maximise the chances of success. Improvement efforts need to bridge gaps between different groups and departments, helping them to understand what the other does and how they can help each other. The organisational aims need to be considered because if the improvement aims are misaligned with these, the chances of success are greatly reduced. Furthermore, those running the intervention need to take into account the nature of the clinical working environment, recognising that it will need to be flexible and deal with great variability, often coming up against challenges they cannot fix such as staff shortages and resource constraints. Furthermore, the fact that there is a lack of time, opportunity and space in which health-care staff can do QI work should be considered.

Relationships with senior management

Improvement interventions result in changes in practice for frontline staff and so their knowledge and involvement is highly important. However, without buy-in at all levels the chance of success is reduced. How the top-level involvement is managed, however, is extremely important. The involvement of hospital management and senior clinicians should be utilised but in different ways. It is important that improvement efforts have the overall approval from management first and then the leadership from senior-level clinicians second. For less wide-scale improvements management do not need to be heavily involved in the work; their implicit support allows frontline staff to become more involved on the ground floor. Managers, however, should be visible and loud in their support of the work and should be responsible for checking the improvement goals are not out of line with organisational goals. They should also be responsible for determining how the organisation can contribute to creating opportunities and conditions for improvement, in addition to determining what changes can be made at the management level to improve safety.

Using evidence and data

The use of data to engage staff and overcome barriers, such as complacency, is extremely useful at the start of a project. At the next stage, during and after process mapping, data can be used to help staff determine where the biggest issues actually exist in their system. This is often revelatory for them and can be extremely motivating when they recognise that long-standing problems are not as diffuse and incomprehensible as they had believed and that a potential solution can be identified. The presentation of evidence from previous successful work in improvement in similar contexts is a strategy which can be utilised in overcoming initial staff resistance, which is often based on a belief that improvement work is likely to be a waste of effort or doomed to fail.

Building rapport

Building rapport between the team driving the improvement programme and the frontline staff was identified as one of the most effective strategies within this study. We did this largely through informal social contacts when opportunities arose during necessary engagement with staff. Mixing formal training and support with opportunities to chat proved so useful that we incorporated it into our programme as a deliberate strategy, and ensured that there were frequent opportunities for social discussion in as comfortable an environment as we could contrive in a ward or theatre environment. This 'cake and coffee' strategy aligns well with the work of Cialdini and Goldstein,⁷⁸ who emphasise 'liking' for the salesman as an effective driver of success in persuasion. Not only did building relationships with staff result in increases in staff engagement and involvement in the S3 study work, but also it was effective in overcoming challenges such as resentment towards outsiders and change fatigue. It is important that groups working with frontline staff take the time to establish this rapport to maximise engagement and overcome some of the barriers to improvement.

Technical support from intervention team

Support from the expert team driving the programme is of course important in creating conditions for successful improvement work, with hospital staff rarely having all the skills and knowledge required, and with strictly limited opportunities for training them in large numbers. In this situation training can create motivation and a basic understanding of the principles of systems improvement work but, in order to deliver a successful project and empower the staff to make further progress independently in future, a considerable amount of ongoing coaching and tuition during the improvement projects is required. Adequate support can empower staff to lead QI efforts in a relatively short period of time. However, it is very important that this involvement is monitored carefully and this requires some difficult judgements. Adopting a 'staff-led' model in the context of modern hospital work means accepting that, initially, staff will require considerable support, but that this must be withdrawn gradually so that they do not fail through its sudden withdrawal or remain dependent on the external team because it is continued for too long. External teams must therefore be careful not to absolve staff of responsibility for the tasks required to make a project a success; arranging meetings and collecting data are two key elements of the 'nuts and bolts' of this kind of programme, which need to be given to staff as early as possible to foster a sense of ownership and responsibility for the success of the project. The balance of improvement project ownership should always favour the staff and external teams should strive to move to an advisory role as soon as they feel the staff can cope. Therefore, the support provided should be factored and planned carefully within the intervention, and plans for fading out the team effectively should be considered so that the staff does not become reliant on them.

Incentives

Incentives should be used carefully. Although it is important to identify and utilise personal incentives to get staff on board, particularly those identified as potential leaders (e.g. career development) and to increase extrinsic motivation for them, it can result in projects lacking ownership and investment if this is the only reason for becoming involved for individuals or if it becomes apparent to other staff that they are involved in an individual's 'vanity project'. It is therefore important to ensure that such motivational factors are secondary and that improvement efforts are led by staff who are intrinsically motivated by a more general desire to improve patient care, working lives or general effectiveness separate from the personal gains they may identify through being involved. To build this group motivation, strategies such as producing data (both historical data demonstrating the possibility of success and early pilot data from their own project demonstrating signs of change) and creating rapport by social bonding should be used to overcome group motivational issues based on historical factors such as change fatigue and learned helplessness. Strategies should also focus on the provision of opportunity; relatively small allocations of dedicated staff time, or the use of rooms and other facilities can create a sense of progress and capacity and act as a powerful motivational boost, and are well worth negotiating with local management when this is possible. The provision of external support and resource should be used to increase this intrinsic motivation to improve, bearing in mind the cautionary notes above about judgement in its gradual withdrawal.

Making benefits visible

Most health-care staff with any length of experience in NHS hospitals will have been involved in improvement initiatives that failed or demonstrated no lasting benefits. Furthermore, involvement in this work requires significant commitment, sometimes in the face of resistance from others. It is therefore important throughout the project to continuously communicate on-going changes and provide real-time feedback to all staff. To maintain staff engagement and enthusiasm it is important that they see progress for their efforts and that the benefits are visible. In situations in which changes or PDCA cycles fail, it is extremely important for researchers to help frontline staff understand why and to understand that this is a temporary setback that can be learned from and overcome, communicating at the same time how to use this information effectively and informatively in making the next change. It is vital that this type of 'good news' communication is done for all staff on the unit where the project is occurring and, if possible, with those external units which are affected by it, so as to prevent confusion regarding changes and encourage support for the work.

Multiple projects

Because of the philosophy adopted at the outset of the programme, we have acquired some practical experience of the consequences of setting up a large number of projects simultaneously. While it may seem important to newly motivated and sometimes excited staff to tackle a whole range of issues at once (and this strategy was indeed advocated in some early versions of the lean approach), caution should be taken in setting up multiple improvement projects simultaneously, and it is a strategy best avoided unless resource and experience levels are high, and senior management support is very strong. If there are an excessive number of projects, it may result in many dropping off or being discontinued relatively quickly, which not only wastes considerable effort but also has negative effects on staff morale and self-belief. Therefore, in the case of staff-led improvement, it may be wise to limit the number of improvement projects suggested. This can be a tricky negotiation, as particular individuals often develop a strong attachment to their project ideas, even when it seems evident that they have little chance of success because of lack of consensus on their importance, or other factors. It is generally better to try to garner enthusiasm for projects that already have a broad swell of support, while offering those with 'pet' projects the possibility that these might be taken on by the group in the future once the problems communally seen as most urgent have been dealt with and the group has acquired more experience. A sensible and helpful guide for slimming down the project list is to propose that projects should be based on specific interventions for which there is evidence of feasibility and effectiveness. For example, evidence that frequent door opening increases bacterial counts in theatre might be used to promote reduced passage of staff in and out of operating theatres.

In any substantial project in which there is significant staff ownership, it is unlikely that there will be only one project. Therefore, external teams will be faced with the task of managing a small number of simultaneous projects. Care should be taken when running these so that there is co-ordination between them and that side effects of one project on another, redundancy or contradictory efforts can be avoided. If too many changes are occurring simultaneously, this can result in difficulty communicating the changes and their rationales to all frontline staff. Furthermore, the impact of a change can be misattributed to other changes occurring at the same time. To avoid a sense of confusion and loss of motivation, it is important that all of the projects and changes can very clearly be linked to a single, simply stated objective, for example 'to reduce the time taken to get emergency patients into theatre for urgent surgery'.

Project focus

The question of linking projects to an overarching objective leads on to consideration of another common problem. The nature of staff-led QI in health care generally leads to individual projects being run with a narrow scope. While this has benefits in that it allows efforts to be focused, it also brings challenges. Staff may choose to focus on areas that are difficult to link to safety or that are difficult to measure and therefore demonstrate an impact. This should be avoided if possible. Being unable to demonstrate an impact can result in a project losing validity in the eyes of others, particularly the staff tasked with implementing it, and thus losing support. Even when changes are measurable, having a narrow project

focus can result in staff missing wider implications of changes. For example, reducing the amount of time it takes for discharge prescriptions to be prepared could free up beds, preventing elective operations being cancelled. The wider implications of improvements may need to be emphasised and supported by convincing data to persuade staff that, by supporting an apparently mundane change, they can potentially impact the overall process.

Lessons: trust management

Staff-led improvement compared with external consultants

Frontline clinical staff, of all types, hold invaluable knowledge about the systems in which they work. While it is important to recognise their ability to provide a unique insight, it is also vital to acknowledge that staff usually do not hold the specialist knowledge required for systems improvement work. Expertise in QI must necessarily be supplied for them to contribute in a leadership role to the development of the systems they work with. This can be done either by employing external consultants who carry and retain all the necessary skills and knowledge or by providing training and support to clinical staff as part of a project designed not only to achieve specified improvements but also to upskill and empower staff to make future improvements with less or no assistance. The potential benefits of the second course are self-evident, but its complexity means that the first is often preferred as less risky, albeit often more costly. Although frontline staff-led improvement is observed across many industries, health care appears to be the only major employment sector in which staff do not routinely have expert support of the second type.

External consultants are generally perceived very negatively by frontline staff and bring with them a number of issues. The hiring of external consultants is generally a non-standard top-down initiative which is usually viewed by frontline staff as a signal that they are failing. Their cost to the NHS is well known to clinical staff, who as a group have a strong loyalty to the NHS concept and, therefore, resent the idea of private companies earning substantial profits from work that they perceive could be done internally. In addition, consultants commonly lack familiarity with the clinical working environment and organisational structure locally, which is particularly important as modern hospital work is particularly complex in comparison with many other businesses, with every hospital process interlinked with multiple other processes. Although external consultants can bring learning from other areas and may possess skills that the clinicians do not, when they are brought in they tend to focus on very specific parts of the process. This narrow focus, together with the fact that they have no brief to train or empower staff, means that learning does not get translated to the frontline staff or to other parts of the organisation.

We propose that a better approach would be one similar to many other industries: change that is led by frontline staff but with support from an expert group. The model provided by the S3 study research team provided this type of expertise and, thereby, greatly increased staff capacity to do the improvement work; however, this support was not continued, as the team were no longer available to staff after the S3 study for both scientific and logistic reasons. However, if there was a similar team with the expertise in the field permanently embedded in the health-care system, working alongside clinicians, this support could be provided on an ongoing basis.

Management role

The role of senior management in facilitating successful staff-led improvement is complex, but two areas are particularly important: public support and deconstruction of intraorganisational barriers. It is useful for management to provide explicit support for frontline improvement work, but this needs to be done in such a way that the credit for the initiative and its leadership are clearly attributed to the unit or area of the organisation that they arise from, to enhance feelings of ownership among staff. It is important that any support given is visible not only to the workers in the area concerned but also to other staff in the trust. The knowledge that a project is approved by management can have considerable power in modifying attitudes to interdepartmental co-operation and working across professional silos.

Health-care delivery is made up of numerous different teams and departments, all of which are interconnected through patient pathways. Owing to the nature of clinical work, clinicians tend to be responsible for one specific area of the patient pathway, and to have neither responsibility nor power over others. So, for example, a consultant surgeon can direct the care of patients he or she operates on but, if they request a radiological investigation for the patient, all decisions about how and when this will be performed are the responsibility of a radiology professional. As a result of this, when improvement work is ongoing it can be very difficult for staff to target the entire pathway, as it usually involves sectors which sit within an entirely different management structure with different priorities and objectives. This frequently results in various different small projects being set up to address different parts of the patient pathway without assessing the impact on the overall process. In the example quoted above, if the surgeon were leading a project to improve the management of patients suspected of having complications after major surgery, they might wish for the capacity to get urgent computerised tomography for certain categories of patient, but achieving this would likely prove difficult or impossible if it did not fit with the objectives of the radiology department or if it increased their costs or pressure of work. It should be management's role to ensure that improvement work is not fragmented and frustrated by this kind of silo thinking, and that changes should take account of the bigger picture, as well as fitting into the organisational aims.

Opportunity

Currently, opportunities within the clinical working environment for staff to do QI work are limited by logistics as the time and space required for staff to perform improvement-related tasks are lacking. The organisational reasons for this are discussed earlier in this chapter. No time is allocated for improvement work in NHS organisations generally and staff are frequently working beyond their capacity, which often forces them to do any improvement work they have taken on in their own personal time. This makes improvement efforts challenging. Health-care staff groups, for example, theatre nurses, anaesthetists and surgeons, also frequently rotate and complete their tasks largely independently of each other. This means staff groups generally tend not to meet regularly, which makes improvement efforts more difficult. Trying to get all the relevant staff together in one place at the same time is incredibly difficult and often involves rescheduling clinical work, for example cancelling theatre lists or asking staff to participate outside their regular work hours. Physical space in which staff can do this work or meet is also lacking in most hospitals, as meeting or staff rooms either do not exist or have been commandeered for clinical use. The time and space to engage in QI work with colleagues therefore just does not exist in health care.

This is very different from most other industries, in which frontline staff have time protected in which to carry out improvement. Health care could consider the allocation of time to allow frontline staff to become fully involved and engaged in QI work, preferably arranged so that the work can be done in combination with some expert groups (e.g. researchers). The logistics of freeing up time in this way should be considered carefully so that this time does not get abused or channelled back into service work.

Resourcing

Specific resources are needed in order for QI projects to be successful. Such resources can vary from IT facilities, to specific equipment or access to hospital data; many of these resources are not easily accessible. It is very often difficult to access accurate and reliable data within the hospital. This often leads to several forms of data being collected about the same factor because staff do not trust the reliability of the sources (e.g. in the MMM project). As one of the more common resources which will be needed among most improvement projects, efforts should be made to make clinical and administrative data more easily accessible to those involved in the improvement work. Furthermore, resources which are likely to be required should be identified and addressed prior to improvement activity occurring, to prevent barriers from developing further down the line when a significant amount of work has potentially already been done.

Policy

Trust policies and SOPs are, by their nature, introduced as a top-down initiative to address patient safety concerns. Although it is important to have policies, guidelines and protocols in place, it is important that these are (1) well designed to suit the clinical working environment, (2) tested before being rolled out and (3) updated regularly. Ideally, there should be close liaison between teams doing improvement work

and those in charge of updating and revising policies, so that the two can easily be harmonised in real time. Currently many protocols are in place in trusts that contradict each other or are completely impractical in the clinical context. Guidance that is so extensive and detailed that no staff member could reasonably be expected to read it during working hours becomes effectively useless, and such unrealistic policies and protocols can result in staff being blamed when things go wrong. This is not only unjust, but also frequently short circuits attempts to perform a deeper systems analysis and identify the true causes of the incident in question, thereby preventing effective measures being taken to prevent a repeat. This paradoxical effect of over-elaborate policy in increasing risk is well recognised among academics studying patient safety. All guidelines, policies, SOPs and protocols could be designed in such a way that they support clinical processes and decision-making rather than hindering them. A well-written, clear operational summary that can be read in less than 5 minutes is an essential component of any effective policy.

Lessons: national training bodies

Frontline staff training

Improvement work should be a collaborative effort between frontline staff and experts in the field. To ensure that this collaboration is fruitful, health-care staff should still receive basic training in the field before or during the improvement projects. This would allow staff to understand the approaches being used and help them to identify more specifically the issues existing within their system. It would also give staff more competence in this field and contribute to their feeling of ownership of the QI work. The current state of health-care education does not incorporate this training. Without it staff are frustrated by not having the tools to make change. Furthermore, in the absence of these skills, bringing in outside consultants to 'do it for them' will alienate the staff and may hinder their involvement in these projects. It is also important that the basic evidence base for current theory on patient safety is presented to staff during their basic training, as much of it is antithetical to the cultural influences that they receive during this training from more experienced staff. Such acculturation is known to be very powerful in forming professional attitudes and it is therefore very important that the opportunity is taken to present to new staff the reasons for regarding error as normal, and harm to patients as a signal that the system, not an individual, needs correction. Much has been written about the evils of the 'shame and blame' approach to investigating safety incidents, but the attitudes that shape this approach are in fact deeply ingrained in the cultural ethos of many experienced NHS staff themselves because of the nature of the training they received early in their careers. Any long-term project to improve safety in the NHS must address this and ensure that the messages sent out by training to the young are more balanced and emphasise appropriately the importance of understanding and perfecting systems of work alongside encouraging appropriate professional behaviour.

Frequent changes in staffing make translation of this training across the organisation difficult. It is nearly impossible to ensure training initiatives reach all providers because clinicians tend to rotate between roles and trusts regularly, with a significant proportion of health-care delivery also depending on locum staff. This problem of the constantly shifting workforce provides an argument for a universal basic training programme to be administered on a regular basis to all NHS staff. However, there are significant risks in such an approach in terms of building resistance and tokenism, as has been noted in existing pan-NHS mandatory training. Therefore, any future training programme should be carefully designed and tested to avoid the defects noted in previous training attempts.

Lessons: national focus

We can do better

In order to create a culture of 'we can always do better', health-care staff need to be able to acknowledge failures and report incidents. In many sectors of the economy, failure, and the use of it as a tool for learning, is widely accepted. However, this has not yet been achieved fully in health care, probably

because of the strong cultural messages received early in clinical staff training, which emphasises the need for exceptional dedication to achieving perfection and the corollary implied message that any failure is shameful. This is particularly difficult for frontline staff as they are genuinely dealing with patient lives and this redoubles the stress associated with apparent failure, reinforcing a culture of blame, in which staff feel disempowered to make change and do not communicate failings for fear of reprimand. In a complex environment in which a certain level of error is inevitable, denial of its existence and personalisation of blame when it is undeniable can potentially lead to inappropriate scapegoating of peripheral group members such as locum staff; hypocrisy and its converse; and inappropriate self-blame by individuals, with concomitant risks to mental health. During this programme we observed this cultural refusal to accept failure resulting in serious criticism of junior medical staff by their seniors, for conducting a PDCA cycle that failed to achieve the desired objective. The idea that this should be treated as a learning experience and used to produce a success in a future cycle was foreign to the consultants concerned, who blamed their registrars for not 'getting it right first time'. The communication of failures and incidents is necessary for (1) providing an overview of the current issues within a system, (2) intervening correctly to address issues and (3) learning from failures. This needs to be one of the messages transmitted by national bodies to their members and trainees in order to improve the general safety culture in the NHS.

The problem with targets

Targets are being set locally and nationally in attempts to address safety issues and take action against poor patient safety practice. However, many of these are set reactively or out of fear and do not address the underlying issues within the system. A typical example of this would be increasing threats of disciplinary action to reduce the number of 'never events' (serious incidents that are wholly preventable as guidance or safety recommendations that provide strong systemic protective barriers are available at a national level and should have been implemented by all health-care providers) in theatre, in order to meet targets. Such activity shifts the priority from one of improving patient safety to one of hitting targets, ignoring the fact that 'never events' in themselves constitute a minuscule proportion of total patient harm, and are best regarded as an index of system reliability, rather than an issue that requires focused attention on the specific event type most recently experienced. Although targets often provide a measure of performance, many are arbitrary and are not indicators of safety. Such targets can force staff to develop workarounds and can result in decreases in quality and safety. Furthermore, when targets are reached they can create a false sense of security as they are not indicative of the standard of the processes that underpin clinical work. As targets often do not take into account the whole context, they can also be misrepresentative in terms of patient safety. We therefore recommend careful analysis of the likely consequences of targets and campaigns through prior ergonomic analysis and pilot testing before widespread adoption.

Measurement

Currently there is a need for adequate and accurate measures of quality and safety. At present there is no organisation-wide measurement of wide-scale improvement. The need for better, more easily collected and more reliable measures of safety in its various aspects was discussed extensively by Charles Vincent and colleagues in a recent monograph for the Health Foundation.⁸⁸ If we are to expect improvement to be aimed at the bigger picture, we need tools and measurement systems that are reliable and representative of the impact of improvement efforts at an organisation-wide level.

Chapter 14 Overall conclusions

This final section inevitably recapitulates some of the later chapters of our report, as it is intended as an accessible summary of all of our conclusions. We have chosen to structure it by examining, first, what we have concluded about the original study questions; second, what we learned during the process of implementing the study; third, what solutions to the problems of improving patient safety appear rational and viable, based on what we have learned and, finally, what the key research questions are for the immediate future of this field.

Our hypothesis

Our results largely support our hypothesis that two conceptually different approaches to intervention, TT and systems improvement, when integrated, are more effective than either approach alone. This finding also supports the theoretical model we proposed before the start of the study, although the programme has not dealt with the postulated role of technology intervention. The qualitative studies provided valuable explanatory underpinning for the finding that the integrated interventions worked better. This proved more complex than expected. Integrated interventions provided an understanding of patient safety principles and motivation to change, alongside practical information on how to conduct a successful change programme, whereas the single-dimension interventions provided either one or other of these. However, it was also true that the integrated intervention programmes led to a much more powerful 'pull' from motivated clinical staff to demand additional support from the expert team for more ambitious and helpful forms of intervention than the single interventions. None of the interventions resulted in dramatic clinical outcome changes; questions therefore remain on whether or not the 'dose' of intervention was adequate. Our belief is that it was, in fact, quite substantial when the degree of 'hands on' practical support provided after the training days is considered. The affordability of a massively better-resourced intervention on a large scale appears to us doubtful.

Why change is hard

Our experience of implementing the programme led us to invest much more effort into a qualitative analysis of mechanism and explanation than we had originally intended. The difficulties of achieving change in current NHS environments are complex. The following is a brief summary of the most important conclusions we made about the barriers to safety and to change in surgical environments.

There is an intimate interconnection between organisational structures and cultures. There is much emphasis at present on enhancing the culture of the NHS in favour of both patient safety and compassion. We certainly experienced some of the adverse effects of current NHS culture in terms of learned helplessness and disengagement, fear, shame and blame and a tendency to give low priority to initiatives to achieve change. We found that time and space for improvement work were non-existent in normal hospital work and were very difficult to create, even with great efforts involving well-disposed staff. However, we would argue that these cultural outcomes are strongly influenced, indeed in many ways caused, by the structural features of NHS, and indeed general, hospital organisations. Some of the key features of this that cause problems are:

- the diffusion of responsibility caused by professional 'silos', which interfere with clarity of line management responsibility, with sharing of authority between professional managers without the power to direct change and high-status clinicians, often with a weak understanding of the practicalities and ethics of management
- interdepartmental barriers and interprofessional tensions that become major barriers to change because management mechanisms to overcome them are weak or non-existent

- incentives for middle management based almost entirely on volume of activity
- lack of performance feedback to allow staff to understand what they are achieving
- a system based on individual performance management as a key tool for maintaining standards, rather than objective analysis and correction of systems problems
- a culture of individual blame based on the unrealistic expectations of an outdated model of professionalism among clinical staff.

Taken together, we feel that these features of the way work is organised in hospitals explain many of the less desirable aspects of NHS hospital culture. A direct attack on culture may therefore be unproductive without addressing these structural features as well.

What can be done now?

One possible implication of this analysis is that major advances in patient safety might require radical changes to the way in which hospitals are organised and run. This may not be possible in the short term. There is a striking degree of homogeneity in the way these institutions work across different cultures internationally, suggesting that strong forces are at work to promote convergent systems evolution. Assuming that a radical transformative solution is not feasible, what tactics do our evidence and experiences lead us to recommend when setting out to improve the quality and safety of surgical (or other clinical) services?

- Change programmes could combine TT to enhance understanding of safety and teamwork principles with systems improvement training to empower staff to implement change.
- A carefully planned implementation and engagement programme is an essential and integral part of any improvement programme. This would include a very strong element of initial communication and engagement at all levels of the organisation, maximal active engagement with and staff ownership of change, and provision of both ongoing expert advice and assistance with communications and measurement.
- Measurement of baseline status data should be organised in advance and undertaken in parallel with initial engagement. Care in selecting appropriate measures and in ensuring that the resources to collect them are available are very important.
- Provision of time and space for staff involvement is very important, and should if possible be agreed with management before any engagement or measurement activity begins.
- Senior management support to eliminate departmental, professional silo or individual agendas from halting change is necessary.
- Outside the context of research protocols, fixed timetables are unwise; significant sustainable change in less than 6 months is unusual.
- Maintenance of measurement and the establishment of a culture of review enhances sustainability.

What are the next questions?

Some of the research questions that arose or were addressed during this programme remain unanswered. The success of the integrated interventions raises others. The research agenda for patient safety is almost limitless, but we will here confine ourselves to questions that have been specifically raised by our own work on this programme. We suggest that the most important of these are detailed below.

- Can the integrated intervention approach be implemented successfully on a larger scale using less resource per staff member involved? This is clearly necessary given the cost implications of applying our methods to large NHS institutions. Specific strategies for delivering the training and supporting the resultant improvement projects need to be developed and tested to confirm whether or not they are able to deliver similar results to the small-scale intervention.

- Can the degree and nature of ‘institutional resistance’ be predicted and circumvented by careful prior analysis of organisational structure and culture before any intervention work is undertaken? A reliable way of evaluating structure and culture that allowed practical decisions to be made about whether or not and how to modify the tactics for a planned intervention would be of great value.
- Can implementation strategies based on engagement and incentivisation of staff be empirically designed and optimised? This might flow from a successful programme of analysis.
- Does intervention that systematically analyses and corrects potential safety hazards related to the use of technology (including IT and EPR) further enhance the effect of an integrated system–culture approach to intervention?

Chapter 15 Knowledge translation

Introduction

The knowledge which has been accumulated throughout the S3 study has been translated throughout a wide variety of outputs. These are presented in the following sections.

Publications

The S3 study work has resulted in a number of papers including methodological, interventional and qualitative. Some of these papers are currently published, whereas others are currently going through the publication process as outlined in *Table 45*.

Conferences

The work from the S3 study programme has been presented at various conferences both nationally and internationally. The teams have presented to audiences of a variety of backgrounds including HFs, surgical, medical and management. These conferences are listed in *Table 46*.

Work in our trust

The S3 study programme garnered attention from other areas in the trust. As a result of the S3 study we were asked to work on a number of other projects within the trust.

Pharmacy

The S3 study research team was approached by management of the pharmacy dispensary in the trust. The management believed that the pharmacy dispensary was experiencing a number of issues including waste, delays, low staff morale and poor working patterns. Key performance indicators were not being met and the dispensary was also breaching its turnaround times. This was also complicated by the introduction of a dispensary robot without considering the implications on the wider working environment. As the dispensary is a key element of the hospital system and impacts significantly on other hospital processes, the fact that it was performing suboptimally was highlighted by dispensary management as warranting urgent assessment and intervention. Three members of the S3 study research team (LM, LF and JH) allocated some time to conduct observations and collect data. Management and dispensary staff on the floor were involved in this process. A log of issues based on evidence collected was generated. Based on this, the research team devised a list of possible solutions. This was fed back to management and frontline staff for them to then make decisions as to what to implement.

System for Electronic Notification and Documentation

The expertise of the S3 study team has also been translated among IT-based projects. Significant input has been provided by a member of the research team (LM) in the design and development of the System for Electronic Notification and Documentation (SEND). This project has been focused on development of an IT solution to calculating track and trigger (T&T) scores. Research has demonstrated that physiological deteriorations in patients are often unrecognised and suggest that errors are made in calculating T&T scores. This can result in potential delays in the appropriate escalation of patient care. SEND is a tablet-based technology that allows for the documentation and evaluation of patients' vital signs. It will replace traditional

TABLE 45 The S3 study publications to date

Type	Title	Area of focus
Method papers	Capturing intraoperative process deviations using a direct observational approach: the glitch method ²²	Glitch
	Observing and categorising process deviations in orthopaedic surgery ⁸⁹	Glitch
	Oxford NOTECHS II: a modified theatre team non-technical skills scoring system ²¹	NOTECHS
	Evaluation of surgical team performance in elective operative theatres ⁹⁰	NOTECHS, glitch
	Compliance and use of the World Health Organization checklist in UK operating theatres ²³	WHO checklist
	The development of process maps in the training of surgical and human factors observers in orthopaedic surgery ³⁶	HF training
Intervention papers	The effect of teamwork training on team performance and clinical outcome in elective orthopaedic surgery: a controlled interrupted time series study ⁹¹	TT intervention
	Effectiveness of a facilitated introduction of a standard operating procedure into routine processes in the operating theatre: a controlled interrupted time series ⁹²	SOP intervention
	'Lean' participative process improvement: outcomes and obstacles in trauma orthopaedics	Lean intervention
	A combined teamwork training and work standardisation intervention in operating theatres: controlled interrupted time series study ⁹³	SOP plus TT intervention
	Quality improvement in surgery combining lean improvement methods with teamwork training: a controlled before–after study ⁹⁴	Lean plus TT intervention
	Safer Delivery of Surgical Services: a combined approach to improvement in surgery (pending)	Lean, SOP and TT
	Combining systems and teamwork approaches to enhance the effectiveness of safety improvement interventions in surgery: the Safer Delivery of Surgical Services (S3) Programme ⁹⁵	Pooled analysis
Qualitative papers	The Safer Delivery of Surgical Services Programme (S3): explaining its differential effectiveness and exploring implications for improving quality in complex systems ²⁴	Qualitative analysis of S3
	Obstacles to improvement: learning from a multimethod study in UK hospitals (awaiting publication)	Qualitative analysis of S3
Additional papers	Human factors and ergonomics in surgical safety: that was then, this is now ⁹⁶	HF in surgery
	Interventions employed to improve intrahospital handover: a systematic review ⁹⁷	Handover
	Simulation provides a window on the quality and safety of the system ⁹⁸	Simulation
	Description of an evaluation and improvement of surgical mortality and morbidity meetings (pending)	MMM
	Improving the reliability and safety of the ward round process (pending)	Ward rounds
	Intentional rounding: a staff led quality improvement intervention in the prevention of patient falls ⁷⁵	Patient falls

TABLE 46 Safer Delivery of Surgical Services conferences

Conference	Presentation title	Presenter
Behavioural Science Applied to Surgery, 2010	Does teamwork training work?	Peter McCulloch
Second Open Seminar Clinical Human Factors Group, 2011	The journey to embed human factors	Ken Catchpole
Association of Surgeons in Training Conference, 2011	The development of process maps in orthopaedic surgery for the training of surgical and human factors observers	Eleanor Robertson
Association of Surgeons in Training Conference, 2011	Evaluation of surgical team performance in elective operative theatres	Mohammed Hadi
Human Factors and Ergonomics Society 55th Annual Meeting, 2011	Observing and categorizing process deviations in orthopaedic surgery	Lauren Morgan
Presentation to Future of Work Seminar Series, Green Templeton College, 2011	Obstacles to improvement in surgical teamwork	Steve New
Fourth Open Seminar Clinical Human Factors Group, 2012	Standardisation – a research perspective	Lauren Morgan
Second World Congress in Clinical Safety, 2013	Perceptions of a frontline health-care team	Lorna Flynn
Fifth Open Seminar Clinical Human Factors Group, 2013	Human factors approach to interface design	Lauren Morgan
Balancing Creativity and Evidence for Patient Safety, 2013	Translating patient safety concepts into practice with discussion cards	Rachel Kwon
Behavioural Science Applied to Surgery, 2013	Successful introduction of an improvement intervention, by a frontline health-care team	Lauren Morgan and Francesca Stedman
Delivering Safer Care and Quality, Innovation, Productivity, Prevention (QIPP), 2013	Managing complex clinical information sharing in a systematic way and managing change	Lauren Morgan
Ergonomics and Human Factors, 2013	Developing a human factors curriculum for frontline staff training in the NHS	Lauren Morgan
Ergonomics and Human Factors, 2013	Developing a human factors curriculum for frontline staff training in the NHS	Lauren Morgan
Faculty of Medical Leadership and Management, 2013	A morning meeting that matters	Francesca Stedman
Patient and Healthcare Provider Safety Symposium, 2013	Human factors for surgical safety: culture versus system	Peter McCulloch
Royal Academy of Engineering Forum, 2013	How can high-level evidence be established for the safety and efficacy of medical devices and systems?	Peter McCulloch
Seventh Open Seminar Clinical Human Factors Group, 2014	Human factors – application beyond the frontline	Lauren Morgan
Association of Surgeons in Training Conference, 2014	Human Factors and CORESS	Peter McCulloch
BMJ International Forum, 2014	Safer surgical services: are systems and culture interventions synergistic?	Peter McCulloch
Ergonomics and Human Factors, 2014	Factors impacting staff led quality improvement in health care	Lorna Flynn
Ergonomics and Human Factors, 2014	Equipment in operating theatres – the weakest link?	Lauren Morgan
Human Factors in Complex Systems, 2014	How much can be achieved in a day of human factors analysis?	Lorna Flynn
Human Factors in Complex Systems, 2014	Integrating human factors in the design of clinical systems	Lauren Morgan

CORESS, Confidential Reporting System for Surgery.

paper charts and manual calculation of trigger scores, helping to alert frontline medical staff to early patient deterioration quickly and reliably.

Oxford Acute Referral System

The Oxford Acute Referral System is a continuation of one of the staff-led QI projects aimed at addressing issues with the process in which patients are referred in from other district general hospitals or general practitioners. This project was led by a clinician who had originally been involved in the S3 study intervention (see *Chapter 9*) and one member of the S3 study research team (LM). As the solution was an IT-based solution, this project went beyond the time frames of the S3 study. It garnered some support and funding from the trust and its development is currently in progress.

Staff-implemented feedback survey

Staff from the final intervention site (three interventions integrated) devised a feedback mechanism for the ward after the S3 study. They created a survey for trainees to complete in terms of their experiences within the department. The survey was designed using the SEIPS model¹⁷ as a basis to prompt trainees to best describe what elements of the system are working effectively and also the areas that are not doing so well. The aim of this was to create a feedback mechanism from which the department could learn and improve on. This survey can be found in *Appendix 7*.

Local and national dissemination

In addition to presenting at conferences, the research from the S3 study has been disseminated through various other channels, both locally and nationally:

- presented at Theatre Users Group
- WHO data collection and audit tool in use in trust
- participant on trust serious untoward incident investigations
- external investigators for other trust incidents (including participating in developing their new SOPs for incident investigation)
- knowledge from the S3 study is being applied and represented at a national level at Confidential Reporting System for Surgery (CORESS).

Teaching

Knowledge from the S3 study programme has also been disseminated throughout various lectures and teaching sessions, as seen in *Table 47*.

Future developments

The S3 project has provided valuable lessons and knowledge for work in the field of quality and safety in health care. The next step is to upscale this to a wider-scale project, extending this knowledge to other areas, trusts, departments and teams within the health-care sector and also to specialties other than surgery. Based on the outputs and lessons from the S3 study, the team have formed a Patient Safety Academy (PSA).

The Patient Safety Academy

The PSA is a new body, working with the Oxford Academic Health Science Network. The PSA will utilise knowledge gained during the S3 study programme to prioritise training and support for NHS staff. The PSA intends to use an integrated approach to improvement, utilising strategies and support identified as effective during the S3 study. Its objective is to provide education, training and support to existing and

TABLE 47 Safer Delivery of Surgical Services teaching

Course/body	Lecture	Lecturer(s)
CORESS Training/Royal College of Surgeons, London	Human factors and systems analysis	Peter McCulloch
Joint Committee on Surgical Training	Human factors in health care	Lauren Morgan
Royal College of Surgeons Dublin	Human factors lecture	Lauren Morgan and Eleanor Robertson
BA Ergonomics & Human Factors/Loughborough University	Systems lecture	Lauren Morgan
Towards Safer Gynaecological Laparoscopic Surgery Course/University of Nottingham	Design for safety	Lauren Morgan
MSc Surgical Science and Practice/University of Oxford	Quality improvement science and systems analysis	Lauren Morgan
MSc Surgical Science and Practice/University of Oxford	Human factors, teamwork and communication	Lauren Morgan
MSc in Surgical Science and Practice/ Patient Safety Academy, and Leadership and Quality Improvement	Human factors in system change	Lauren Morgan
Medicine/University of Oxford	Patient safety lecture	Lauren Morgan
Central Lecture/University of Oxford	Human factors in projects	Lauren Morgan
Saïd Business School/University of Oxford	Findings and experiences from the S3 study module, especially relating to the role of participative and experimental process improvement, have been used extensively in the teaching of technology and operations management to MBA students	Steve New
Saïd Business School/University of Oxford	Findings and experiences from the S3 study module used on executive education programmes for major companies, including Clifford Chance, Royal Mail, Tesco and Daiichi Sankyo	Steve New
Saïd Business School/University of Oxford	Findings and experiences from the S3 study module used on the Oxford advanced management and leadership programme, the Oxford Diploma in Strategy and Innovation, and the new Oxford University doctoral programme on cyber security	Steve New

BA, Bachelor of Arts; MBA, Master of Business Administration; MSc, Master of Science.

new clinical workforces in the NHS and through this to improve quality of care. The PSA will be developed from expertise acquired from the S3 study and previous work in the field of quality and safety in the acute and academic sectors and absorb and integrate all willing actors and workstreams in this field within the geographical region into one coherent group. In so doing, the PSA will translate the knowledge it has acquired, providing leadership for the region in this field.

The PSA aims to translate the knowledge acquired from the S3 study programme to four separate workstreams across the Oxford Academic Health Science Network region: (1) trust boards and senior management, (2) acute surgery, (3) mental health and (4) primary care. The PSA hope to provide training and support to enable these groups in identifying key safety priorities and developing joint improvement programmes/projects across the region.

Summary

In summary, the knowledge acquired from the S3 study has been translated through various channels including conference presentations, publications, teaching and training, and also through representation among other groups. It has been applied within other projects (e.g. pharmacy, Oxford Acute Referral System), some led by S3 study researchers and others led by staff who were involved in the S3 project. The final route of translation for this work will be through its use and application as part of the PSA, in which it will be utilised in the provision of training and support for various groups of NHS staff in improving quality and safety.

Chapter 16 Recommendations for future research

This programme of research has led to the development of a new system for safety and quality intervention that we have based on the findings during the programme. Having confirmed that integration of TT to improve staff safety culture with QI training to empower them to make necessary change in their environment was more effective than either initiative on its own in improving safety-relevant work processes, we have developed an integrated QI intervention that incorporates this learning. However, it is clear that many important questions about how to optimise organisational safety in surgery and in health care, in general, remain unanswered.

Models for upscaling the integrated quality intervention programme

This programme has demonstrated that some interventions can improve the function of operating theatre teams in ways that we would expect to enhance patient safety and the reliability with which effective treatment can be delivered. It has evaluated the different properties of interventions directed at improving the working culture and relationships between team members, and interventions designed to improve systems and processes of work. It has evaluated the effects of combining these approaches. All of this has been done at the small scale of an individual operating theatre or two. Following on from these studies, we have conducted an improvement intervention in a single-specialty system, addressing the problems identified not just in the operating theatre but also throughout the patient journey. Our results point to the effectiveness of integrating culture and systems interventions, and to the conclusion that this integrated approach is applicable throughout the patient journey. However, the experience of conducting these studies has also thrown up important questions that need to be answered if the learning from the primary experiments is to be applied on a large scale within the UK and internationally, and has exposed some important shortcomings in our approach to intervention that we have begun to correct, but which still require further work. The model we used in these studies is not realistic as a solution to the safety problems of surgery in the NHS or indeed any large organisation. The input in terms of resource and personnel to achieve change in a small clinical area was prohibitively high for any large-scale implementation. Even the final piece of work in neurosurgery could only be described as being an example extending the approach to a medium-sized working unit, which is a division within a large hospital. Each study relied on the involvement of a pair of researchers to collect the key data, and on a team of between two and six people to deliver the intervention to a group of staff numbering between 20 and 200. This represents a very significant input of trained workforce and would clearly be infeasible if applied to an entire hospital, or even to the NHS as a whole. Although the proof of principle has been important, it is clear that the next step needs to be demonstration that the same improvements can be achieved with a more economic approach. The first area for future research we would therefore recommend is the development of pilot studies of new delivery methods to scale up the integrated culture and systems intervention we have developed. There is a real danger of diluting the effect of the intervention to such an extent that it no longer makes a useful change. Studies will therefore need to be carefully designed to evaluate not only clinical effectiveness but also mechanism, so that any failures can be analysed. Our current plan for this is based on the implementation strategy we developed during the programme and involves identifying and selecting small teams of key leaders within a clinical area, and providing them with the training and support to carry out specific improvement projects within their clinical area. This will involve intensive training for a few days, web and printed knowledge resources, and regular support from experts in ergonomic, psychology and systems improvement by telephone or e-mail. Other models are of course possible, but the nature of the problem makes it infeasible to run direct comparisons of multiple models. We therefore recommend that groups define precisely the model they use to impart information at a large scale, to whom they impart it, and what support they supply for how long, so that some conclusions can be drawn about what approaches appear compatible with success and what resources they require.

Once a study of this type has demonstrated feasibility, a generalised version of the intervention strategy based on our findings could be tested on a suitable target problem, in a multicentre cluster randomised trial, perhaps with a stepped-wedge design to ensure engagement from each hospital involved. This would be the logical conclusion of any line of research aimed at defining generally applicable principles for improving quality and safety in health-care organisations.

Addition of a technology arm

This programme of work was based on the 3D model of safety¹⁸ that postulates three dimensions of risk and safety, the third being technology. We have not attempted to evaluate the importance of defects in technology as a contribution to safety risks in these studies, but it is clearly a factor of great importance. Equally, technology is often considered as the first line of defence against human error in health care, as in other settings. Current analyses of the positive and negative contributions of technology to patient safety tend to adopt a narrow focus on a single area, whereas proposals for technology-based solutions are generally based on a specific conception of how a computer-based template for organising hospital or clinic work should be constructed. A broad and comprehensive analysis of the contribution of technology to patient safety in NHS surgery both positive and negative is needed as an essential step towards developing a unified model for safety intervention. This will require a systematic literature review and prospective observational and analytical studies to determine how (and how often) medical devices, IT systems and instruments are key factors in causing or preventing adverse events. EPRs and associated control systems for diagnostics, prescription and recording are currently undergoing a period of rapid development, which will inevitably result in a fully integrated computer-driven management system for hospitals. Research on the benefits and problems of this is already ongoing, but to achieve maximum benefit, studies on the role of technology in health-care safety need to be still more widely drawn, to encompass the physical equipment and instruments used not only in the operating theatre but also in pre- and post-operative care. Cross-sectional and cohort studies of incidents and adverse effects may allow a picture to emerge of the role of technology in both causing and solving patient safety problems and how it interacts with the staff culture and systems of work. An intervention study to comprehensively risk analyse and then modify or replace key technologies, perhaps on the basis of a prior failure mode effects analysis, would require the development and validation of measures and methods for evaluating the impact of technology on safety. Such an intervention would be a major undertaking that would likely require the full co-operation of an entire hospital trust, but could provide important information on how technology rationalisation could contribute to the development of a truly high-reliability clinical organisation. This would in effect represent a single-dimension intervention focused only on the technology dimension. The logical conclusion of this line of work would be to conduct an integrated intervention study in a suitable surgical environment in which culture, system and technology were all optimised by an extended version of the integrated quality intervention programme, perhaps comparing this intervention directly or indirectly with the two-dimensional 'systems plus culture' model we have developed in this programme.

Analysis of organisational barriers to system change

Experientially, the most striking feature of conducting this study was the difficulty encountered in implementing the intended training and improvement projects. In every study we encountered multiple barriers to progress inherent in the very systems and culture we were attempting to improve. There are clearly a number of important questions here which need to be addressed. The organisational structure of NHS hospitals strikes most outside observers as bizarrely complicated and lines of accountability, in particular, seem hopelessly unclear. It is important to understand why this structure has evolved and persisted, despite the obvious disadvantages. The assumption that a clearer and simpler structure with stronger accountability would be more effective and safer for patients may need to be tested empirically at some stage, but this would clearly involve a very large-scale experiment (essentially a whole trust, but with

implications that would certainly attract comment and concern from national professional bodies, unions and others). There is a large literature in the business community dealing with the theory of organisational change and an important component of any investigation of organisational barriers should be an attempt to understand these through this lens. We recommend observational and analytical studies of the relationships between culture, organisational structure and resistance to change in NHS hospitals, in collaboration with experts in change management. Any attempt to develop a large pragmatic trial of an intervention approach will need to be designed using an approach grounded in an appropriately adapted theory of organisational change.

If we accept for the moment that major changes to organisational structures are unlikely to be possible, qualitative research into the attitudes to safety interventions and system improvement expressed by different staff groups may be valuable in understanding how these might be changed or circumvented when they appear unhelpful. Such research may require the use of sophisticated theory-based questionnaires to differentiate between surface attitudes and motivations, which are often moulded by the expectations of the organisation and the professional peer group, and deeper, sometimes subconscious drivers which show themselves in decision-making choices rather than public utterances. Using previous research on organisational culture and organisational change theory, it may be possible to produce a predictive model that will help to identify how a given organisation will react to attempts at systems change, and test this against the outcome of real-life improvement projects.

Action research on implementation strategies

Defining and understanding barriers to change in NHS organisations is important, but its importance is ultimately determined by whether or not it is useful in finding ways of circumventing or eliminating such barriers. This kind of investigation of implementation strategies is very challenging, as any intervention will need to be directed at large segments of organisations or at whole-hospital trusts. Providing control groups or standardising the complex intervention strategy both appear to be infeasible and this in turn leads to doubt over whether or not quasi-experimental intervention studies will be possible at all, or whether or not we will need to rely on observational studies, uncontrolled QI projects or action research formats. Studies using suitable formats should be conducted to evaluate, with as much validity as possible, theory-based implementation strategies that overlay the simple provision of training and support with a programme of action to maximise the chances of effective uptake.

Experimental compared with observational studies

It may prove useful to conduct analyses of case studies or pooled analyses of recent trust attempts to achieve organisational change to promote safety, using the existing tools for evaluation of culture and system robustness. A standardised taxonomy and language will be needed to ensure consistency in reporting what occurred and in describing the organisational environment in a consistent fashion. Collection of data from a series of well-characterised episodes may allow the use of regression to identify the importance of specific aspects of the culture in facilitating or obstructing success. The same approach may allow evaluation of the contribution of specific components of our current implementation strategy to success in different settings. The model of theory-based prediction and evaluation used in economics and in much research into organisational change in business may be applied to safety and QI in health care at a large scale.

Acknowledgements

The authors would like to thank the management, surgical, anaesthetic and theatre staff at the trusts that took part in this study for their co-operation and forbearance: Oxford University Hospitals NHS Trust, University Hospitals Coventry and Warwickshire NHS Trust and Kettering General Hospital NHS Foundation Trust. We also wish to thank all the patients, families and carers involved in this research.

Contributions of authors

Peter McCulloch conceived and designed the study programme, led the implementation of the programme at a strategic level and led the writing strategy.

Lauren Morgan focused initially on development and validation of measurement tools, working with Dr Ken Catchpole. On his departure she adopted the role of team leader for implementation of the intervention studies. She assisted with analysis of the data and led writing for a number of the papers on measurement and on individual studies.

Lorna Flynn joined the team in the middle of the programme and was responsible for the qualitative study of mechanism, supervised by Graham Martin. She devised, analysed and conducted the structured interviews that formed the basis for this analysis, and collaborated with Professor Martin to write the paper describing it and parts of this report.

Oliver Rivero-Arias was the Health Economics expert on the team. A significant part of the work he intended to carry out was disrupted by problems with data protection, but he provided very valuable help with data analysis on clinical outcomes, for which he devised the analysis scheme, and on writing.

Graham Martin supervised Ms Flynn in the conduct and analysis of the qualitative studies, and contributed to the writing of the reports and papers flowing from this aspect of the programme.

Gary Collins was the statistician to the programme and devised the analysis scheme that permitted pooled analysis of the data from the parallel intervention studies.

Steve New provided expertise in lean and QI techniques, management change theory and led the lean intervention project. He contributed to the writing of the overall report and the accompanying papers.

We also would like to thank the collaborators, research staff and support staff employed on this research programme.

We would like to thank the collaborators (non-authors) who guided and supported us as the research programme progressed:

- Alastair Gray, Health Economics Research Centre, University of Oxford
- Alison Kitson, Green Templeton College, University of Oxford
- Crispin Jenkinson, Department of Public Health, University of Oxford
- Damian Griffin, Health Sciences, University of Warwick
- Doug Altman, Centre for Statistics in Medicine, University of Oxford
- Elaine Strachan-Hall, Oxford University Hospitals NHS Foundation Trust
- Karen Barker, Oxford University Hospitals NHS Trust
- Ken Catchpole, Cedars-Sinai Medical Centre
- Ravi Dravid, Kettering General Hospital NHS Foundation Trust
- Renee Lyons, School of Public Health, University of Toronto

- Tony Berendt, Oxford University Hospitals NHS Trust
- Richard King, University Hospitals Coventry and Warwickshire NHS Trust
- Carole Johnston, Oxford University Hospitals NHS Trust
- Mary Quirke, Oxford University Hospitals NHS Trust
- Nick de Pennington, Oxford University Hospitals NHS Trust
- James Livermore, Oxford University Hospitals NHS Trust
- Jon Westbrook, Oxford University Hospitals NHS Trust
- John Chatwin, Oxford University Hospitals NHS Trust.

We would like to thank all the researchers (non-authors) who were involved in the research programme: Mo Hadi, Sharon Pickering, Laura Bleakley, Eleanor Robertson, Francesca Stedman, Rachel Kwon, Julia Matthews, Jacob Haddad, Ken Catchpole and Christopher Pennell.

We would like to acknowledge and recognise the administrative support provided throughout the programme: Sarah Hills, Sam French and Beth Bosiak. We would also like to recognise the departmental and research support within the University of Oxford and the host sites.

We would like to thank all the Scientific Advisory Committee members, past and present, for their advice and support during the research programme:

- Beverly Norris, National Patient Safety Agency
- Charles Vincent, Imperial College, London
- David Bowen, Patient Safety Representative
- Janet Anderson, King's College, London
- Joan Russell, National Patient Safety Agency
- Patrick Waterson, Loughborough University
- James Reason (past), University of Manchester
- Melinda Lyons (past), National Patient Safety Agency.

Finally, we would like to thank the NIHR for funding this research programme.

Data sharing statement

Requests for access to data should be addressed to the corresponding author.

References

1. Pormann PE, Savage-Smith E. *Medieval Islamic Medicine*. Edinburgh: Edinburgh University Press; 2007.
2. Levey M. Medical ethics of medieval Islam with special reference to Al-Ruhāwī's 'Practical Ethics of the Physician'. *Trans Am Philos Soc* 1967;**57**:1–100. <http://dx.doi.org/10.2307/1006137>
3. York EP, Brown T. The price they paid. *J Community Hosp Intern Med Perspect* 2015;**5**:26436. <http://dx.doi.org/10.3402/jchimp.v5.26436>
4. Illich I. *Medical Nemesis: The Expropriation of Health*. London: Calder and Boyars; 1975.
5. Cochrane AL. *Effectiveness and Efficiency: Random Reflections on Health Services*. London: The Nuffield Provincial Hospitals Trust; 1972.
6. Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, *et al*. The nature of adverse events in hospitalized patients. *N Engl J Med* 1991;**324**:377–84. <http://dx.doi.org/10.1056/NEJM199102073240605>
7. Kohn LT, Corrigan J, Donaldson MS. *To Err is Human: Building a Safer Health System*. Washington, DC: National Academy Press; 2000.
8. Steel K, Gertman PM, Crescenzi C, Anderson J. Iatrogenic illness on a general medical service at a university hospital. *N Engl J Med* 1981;**304**:638–42. <http://dx.doi.org/10.1056/NEJM198103123041104>
9. Reason J. Human error: models and management. *West J Med* 2000;**172**:393–6. <http://dx.doi.org/10.1136/ewj.172.6.393>
10. Rochlin GI, La Porte TR, Roberts KH. The self-designing high-reliability organization: aircraft carrier flight operations at sea. *Naval War College Review* 1987;**40**:76–90.
11. World Health Organization. *WHO Patient Safety Research: Better Knowledge for Safer Care 2009*. Geneva: WHO; 2009.
12. Helmreich RL. On error management: lessons from aviation. *BMJ* 2000;**320**:781–5. <http://dx.doi.org/10.1136/bmj.320.7237.781>
13. Vincent C, Taylor-Adams S, Stanhope N. Framework for analysing risk and safety in clinical medicine. *BMJ* 1998;**316**:1154–7. <http://dx.doi.org/10.1136/bmj.316.7138.1154>
14. Donaldson L. An organisation with a memory. *Clin Med* 2002;**2**:452–7. <http://dx.doi.org/10.7861/clinmedicine.2-5-452>
15. Amalberti R, Auroy Y, Berwick D, Barach P. Five system barriers to achieving ultrasafe health care. *Ann Intern Med* 2005;**142**:756–64. <http://dx.doi.org/10.7326/0003-4819-142-9-200505030-00012>
16. Vincent C, Taylor-Adams S, Chapman EJ, Hewett D, Prior S, Strange P, *et al*. How to investigate and analyse clinical incidents: clinical risk unit and association of litigation and risk management protocol. *BMJ* 2000;**320**:777–81. <http://dx.doi.org/10.1136/bmj.320.7237.777>
17. Carayon P, Schoofs Hundt A, Karsh BT, Gurses AP, Alvarado CJ, Smith M, *et al*. Work system design for patient safety: the SEIPS model. *Qual Saf Health Care* 2006;**15**(Suppl. 1):50–8. <http://dx.doi.org/10.1136/qshc.2005.015842>
18. McCulloch P, Catchpole K. A three-dimensional model of error and safety in surgical health care microsystems. Rationale, development and initial testing. *BMC Surg* 2011;**11**:23. <http://dx.doi.org/10.1186/1471-2482-11-23>

19. Levy MM, Dellinger RP, Townsend SR, Linde-Zwirble WT, Marshall JC, Bion J, *et al.* The Surviving Sepsis Campaign: results of an international guideline-based performance improvement program targeting severe sepsis. *Intensive Care Med* 2010;**36**:222–31. <http://dx.doi.org/10.1007/s00134-009-1738-3>
20. 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. <http://dx.doi.org/10.1136/bmj.d199>
21. Robertson ER, Hadi M, Morgan LJ, Pickering SP, Collins G, New S, *et al.* Oxford NOTECHS II: a modified theatre team non-technical skills scoring system. *PLOS ONE* 2014;**9**:e90320. <http://dx.doi.org/10.1371/journal.pone.0090320>
22. Morgan L, Robertson E, Hadi M, Catchpole K, Pickering S, New S, *et al.* Capturing intraoperative process deviations using a direct observational approach: the glitch method. *BMJ Open* 2013;**3**:e003519. <http://dx.doi.org/10.1136/bmjopen-2013-003519>
23. Pickering SP, Robertson ER, Griffin D, Hadi M, Morgan LJ, Catchpole KC, *et al.* Compliance and use of the World Health Organization checklist in UK operating theatres. *Br J Surg* 2013;**100**:1664–70. <http://dx.doi.org/10.1002/bjs.9305>
24. Flynn LC, McCulloch PG, Morgan LJ, Robertson ER, New SJ, Stedman FE, *et al.* The Safer Delivery of Surgical Services Programme (S3): explaining its differential effectiveness and exploring implications for improving quality in complex systems. *Ann Surg* 2016;**264**:997–1003. <http://dx.doi.org/10.1097/SLA.0000000000001583>
25. Carthey J. The role of structured observational research in health care. *Qual Saf Health Care* 2003;**12**(Suppl. 2):ii13–6. http://dx.doi.org/10.1136/qhc.12.suppl_2.ii13
26. Mishra A, Catchpole K, McCulloch P. The Oxford NOTECHS System: reliability and validity of a tool for measuring teamwork behaviour in the operating theatre. *Qual Saf Health Care* 2009;**18**:104–8. <http://dx.doi.org/10.1136/qshc.2007.024760>
27. Christian CK, Gustafson ML, Roth EM, Sheridan TB, Gandhi TK, Dwyer K, *et al.* A prospective study of patient safety in the operating room. *Surgery* 2006;**139**:159–73. <http://dx.doi.org/10.1016/j.surg.2005.07.037>
28. Lingard L, Reznick R, Espin S, Regehr G, DeVito I. Team communications in the operating room: talk patterns, sites of tension, and implications for novices. *Acad Med* 2002;**77**:232–7. <http://dx.doi.org/10.1097/00001888-200203000-00013>
29. Flin R, Martin L, Goeters K-M. Development of the NOTECHS (Non-Technical Skills) System for Assessing Pilots' CRM Skills. In Harris D, Muir HC, editors. *Contemporary Issues in Human Factors and Aviation Safety*. Aldershot: Ashgate; 2005. pp. 133–51.
30. Hull L, Arora S, Kassab E, Kneebone R, Sevdalis N. Observational Teamwork Assessment for Surgery: content validation and tool refinement. *J Am Coll Surg* 2011;**212**:234–43. <http://dx.doi.org/10.1016/j.jamcollsurg.2010.11.001>
31. Niitsu H, Hirabayashi N, Yoshimitsu M, Mimura T, Taomoto J, Sugiyama Y, *et al.* Using the Objective Structured Assessment of Technical Skills (OSATS) global rating scale to evaluate the skills of surgical trainees in the operating room. *Surg Today* 2013;**43**:271–5. <http://dx.doi.org/10.1007/s00595-012-0313-7>
32. Kemper PF, van Noord I, de Bruijne M, Knol DL, Wagner C, van Dyck C. Development and reliability of the explicit professional oral communication observation tool to quantify the use of non-technical skills in healthcare. *BMJ Qual Saf* 2013;**22**:586–95. <http://dx.doi.org/10.1136/bmjqs-2012-001451>

33. Flin R, Patey R. Non-technical skills for anaesthetists: developing and applying ANTS. *Best Pract Res Clin Anaesthesiol* 2011;**25**:215–27. <http://dx.doi.org/10.1016/j.bpa.2011.02.005>
34. Yule S, Flin R, Maran N, Rowley D, Youngson G, Paterson-Brown S. Surgeons' non-technical skills in the operating room: reliability testing of the NOTSS behavior rating system. *World J Surg* 2008;**32**:548–56. <http://dx.doi.org/10.1007/s00268-007-9320-z>
35. Mitchell L, Flin R, Yule S, Mitchell J, Coutts K, Youngson G. Evaluation of the SPLINTS system for scrub practitioners' non-technical skills. *Proc Hum Fact Ergon Soc Annu Meet* 2011;**55**:690–4. <http://dx.doi.org/10.1177/1071181311551143>
36. Robertson E, Hadi M, Morgan L, Pickering S, Catchpole K, Griffin D, *et al.* The development of process maps in the training of surgical and human factors observers in orthopaedic surgery. *Int J Surg* 2011;**9**:549. <http://dx.doi.org/10.1016/j.ijsu.2011.07.266>
37. Catchpole KR, Giddings AE, Wilkinson M, Hirst G, Dale T, de Leval MR. Improving patient safety by identifying latent failures in successful operations. *Surgery* 2007;**142**:102–10. <http://dx.doi.org/10.1016/j.surg.2007.01.033>
38. Parker SE, Laviana AA, Wadhwa RK, Wiegmann DA, Sundt TM, III. Development and evaluation of an observational tool for assessing surgical flow disruptions and their impact on surgical performance. *World J Surg* 2010;**34**:353–61. <http://dx.doi.org/10.1007/s00268-009-0312-z>
39. Lingard L, Regehr G, Orser B, Reznick R, Baker GR, Doran D, *et al.* Evaluation of a preoperative checklist and team briefing among surgeons, nurses, and anesthesiologists to reduce failures in communication. *Arch Surg* 2008;**143**:12–7, discussion 8. <http://dx.doi.org/10.1001/archsurg.2007.21>
40. World Health Organization. *World Alliance for Patient Safety – Safe Surgery Saves Lives*. Geneva: WHO; 2008 (updated 1 February 2013).
41. Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat A-HS, 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. <http://dx.doi.org/10.1056/NEJMsa0810119>
42. Gawande AA. *The Checklist Manifesto*. New York, NY: Profile Books; 2010.
43. Semel ME, Resch S, Haynes AB, Funk LM, Bader A, Berry WR, *et al.* Adopting a surgical safety checklist could save money and improve the quality of care in U.S. hospitals. *Health Aff (Millwood)* 2010;**29**:1593–9. <http://dx.doi.org/10.1377/hlthaff.2009.0709>
44. Harries RL, Young N, Skehan N, Woodward A. Surgical safety checklists: 'who's doing it?'. *Online J Clin Audit* 2011;**3**.
45. Bosk CL, Dixon-Woods M, Goeschel CA, Pronovost PJ. Reality check for checklists. *Lancet* 2009;**374**:444–5. [http://dx.doi.org/10.1016/S0140-6736\(09\)61440-9](http://dx.doi.org/10.1016/S0140-6736(09)61440-9)
46. Cullati S, Le Du S, Rae AC, Micallef M, Khabiri E, Ourahmoune A, *et al.* Is the Surgical Safety Checklist successfully conducted? An observational study of social interactions in the operating rooms of a tertiary hospital. *BMJ Qual Saf* 2013;**22**:639–46. <http://dx.doi.org/10.1136/bmjqs-2012-001634>
47. King N, Horrocks C. *Interviews in Qualitative Research*. Los Angeles, CA: Sage; 2010.
48. Maykut P, Morehouse R. *Beginning Qualitative Research: A Philosophic and Practical Guide*. London: Falmer; 1994.
49. Glaser BG. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York, NY: Aldine de Gruyter; 1967.

50. Spear SJ, Bowen HK. Decoding the DNA of the Toyota production system. *Harv Bus Rev* 1999;**77**:96–106.
51. Deming WE. *Out of the Crisis: Quality, Productivity and Competitive Position*. Cambridge, MA: Cambridge University Press; 1986.
52. Pronovost PJ, Berenholtz SM, Goeschel C, Thom I, Watson SR, Holzmueller CG, et al. Improving patient safety in intensive care units in Michigan. *J Crit Care* 2008;**23**:207–21. <http://dx.doi.org/10.1016/j.jcrc.2007.09.002>
53. McCulloch P, Rathbone J, Catchpole K. Interventions to improve teamwork and communications among healthcare staff. *Br J Surg* 2011;**98**:469–79. <http://dx.doi.org/10.1002/bjs.7434>
54. Bozic KJ, Maselli J, Pekow PS, Lindenauer PK, Vail TP, Auerbach AD. The influence of procedure volumes and standardization of care on quality and efficiency in total joint replacement surgery. *J Bone Joint Surg Am* 2010;**92**:2643–52. <http://dx.doi.org/10.2106/JBJS.I.01477>
55. Kortgen A, Niederprum P, Bauer M. Implementation of an evidence-based ‘standard operating procedure’ and outcome in septic shock. *Crit Care Med* 2006;**34**:943–9. <http://dx.doi.org/10.1097/01.CCM.0000206112.32673.D4>
56. 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. <http://dx.doi.org/10.1136/bmjqs-2011-000760>
57. Aveling EL, Martin G, Armstrong N, Banerjee J, Dixon-Woods M. Quality improvement through clinical communities: eight lessons for practice. *J Health Organ Manag* 2012;**26**:158–74. <http://dx.doi.org/10.1108/14777261211230754>
58. de Leval MR, Carthey J, Wright DJ, Farewell VT, Reason JT. Human factors and cardiac surgery: a multicenter study. *J Thorac Cardiovasc Surg* 2000;**119**:661–72. [http://dx.doi.org/10.1016/S0022-5223\(00\)70006-7](http://dx.doi.org/10.1016/S0022-5223(00)70006-7)
59. Lingard L, Espin S, Whyte S, Regehr G, Baker GR, Reznick R, et al. Communication failures in the operating room: an observational classification of recurrent types and effects. *Qual Saf Health Care* 2004;**13**:330–4. <http://dx.doi.org/10.1136/qshc.2003.008425>
60. Wilson G. Implementation of releasing time to care – the productive ward. *J Nurs Manag* 2009;**17**:647–54. <http://dx.doi.org/10.1111/j.1365-2834.2009.01026.x>
61. Bloodworth K. The productive ward and the productive operating theatre. *J Perioper Pract* 2011;**21**:97.
62. Boaden R, Harvey G, Moxham C, Proudlove N. *Quality Improvement: Theory and Practice in Healthcare*. Coventry: NHS Institute for Innovation and Improvement; 2008.
63. Radnor ZJ, Holweg M, Waring J. Lean in healthcare: the unfilled promise? *Soc Sci Med* 2012;**74**:364–71. <http://dx.doi.org/10.1016/j.socscimed.2011.02.011>
64. Mazzocato P, Holden RJ, Brommels M, Aronsson H, Backman U, Elg M, et al. How does lean work in emergency care? A case study of a lean-inspired intervention at the Astrid Lindgren Children’s Hospital, Stockholm, Sweden. *BMC Health Serv Res* 2012;**12**:28. <http://dx.doi.org/10.1186/1472-6963-12-28>
65. Joosten T, Bongers I, Janssen R. Application of lean thinking to health care: issues and observations. *Int J Qual Health Care* 2009;**21**:341–7. <http://dx.doi.org/10.1093/intqhc/mzp036>
66. Brandao de Souza L. Trends and approaches in lean healthcare. *Leadersh Health Serv* 2009;**22**:121–39. <http://dx.doi.org/10.1108/17511870910953788>
67. Liker JK, Franz JK. *The Toyota Way to Continuous Improvement: Linking Strategy and Operational Excellence to Achieve Superior Performance*. New York, NY: McGraw-Hill; 2011.


68. McCulloch P, Kreckler S, New S, Sheena Y, Handa A, Catchpole K. Effect of a 'lean' intervention to improve safety processes and outcomes on a surgical emergency unit. *BMJ* 2010;**341**:c5469. <http://dx.doi.org/10.1136/bmj.c5469>
69. Kreckler S, Morgan RD, Catchpole K, New S, Handa A, Collins G, *et al.* Effective prevention of thromboembolic complications in emergency surgery patients using a quality improvement approach. *BMJ Qual Saf* 2013;**22**:916–22. <http://dx.doi.org/10.1136/bmjqs-2013-001855>
70. Smith ML, Wilkerson T, Grzybicki DM, Raab SS. The effect of a Lean quality improvement implementation program on surgical pathology specimen accessioning and gross preparation error frequency. *Am J Clin Pathol* 2012;**138**:367–73. <http://dx.doi.org/10.1309/AJCP3YXID2UHZPHT>
71. Catchpole KR, de Leval MR, McEwan A, Pigott N, Elliott MJ, McQuillan A, *et al.* Patient handover from surgery to intensive care: using Formula 1 pit-stop and aviation models to improve safety and quality. *Paediatr Anaesth* 2007;**17**:470–8. <http://dx.doi.org/10.1111/j.1460-9592.2006.02239.x>
72. van Klei WA, Hoff RG, van Aarnhem EEHL, Simmermacher RKJ, Regli LPE, Kappen TH, *et al.* Effects of the introduction of the WHO 'Surgical Safety Checklist' on in-hospital mortality: a cohort study. *Ann Surg* 2012;**255**:44–9. <http://dx.doi.org/10.1097/SLA.0b013e31823779ae>
73. 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. <http://dx.doi.org/10.1136/bmjqs-2012-001325>
74. McCulloch P, Mishra A, Handa A, Dale T, Hirst G, Catchpole K. The effects of aviation-style non-technical skills training on technical performance and outcome in the operating theatre. *Qual Saf Health Care* 2009;**18**:109–15. <http://dx.doi.org/10.1136/qshc.2008.032045>
75. Morgan L, Flynn L, Robertson E, New S, Forde-Johnston C, McCulloch P. Intentional rounding: a staff led quality improvement intervention in the prevention of patient falls. *J Clin Nurs* 2016; in press. <http://dx.doi.org/10.1111/jocn.13401>
76. Amin Y, Grewcock D, Andrews S, Halligan A. Why patients need leaders: introducing a ward safety checklist. *J R Soc Med* 2012;**105**:377–83. <http://dx.doi.org/10.1258/jrsm.2012.120098>
77. Rogers E. *Diffusion of Innovations*. 5th edn. New York, NY: Free Press; 2003.
78. Cialdini R, Goldstein N. The science and practice of persuasion. *Cornell Hosp Q* 2002;**43**:40–50. <http://dx.doi.org/10.1177/001088040204300204>
79. de Vries EN, Hollmann MW, Smorenburg SM, Gouma DJ, Boermeester MA. Development and validation of the SURgical PATient Safety System (SURPASS) checklist. *Qual Saf Health Care* 2009;**18**:121–6. <http://dx.doi.org/10.1136/qshc.2008.027524>
80. Appelbaum E. *Manufacturing Advantage: Why High-Performance Work Systems Pay Off*. Ithaca, NY: Cornell University Press; 2000.
81. National Health Service Litigation Authority. *The National Health Service Litigation Authority*. 2014. URL: www.nhs.uk (accessed 4 June 2015).
82. Curtis L. *Unit Costs of Health and Social Care 2012*. Canterbury: PSSRU, University of Kent; 2012.
83. Department of Health (DH). *Agenda for Change Pay Bands and Spine Points from 1 April 2016 (England)*. 2016. URL: www.nhsemployers.org/~media/Employers/Documents/Pay%20and%20reward/AfC%20pay%20bands%20from%201%20April%202016_FINAL.pdf (accessed 26 September 2016).
84. Urbach DR. Closing in on surgical practice variations. *Ann Surg* 2014;**259**:628–9. <http://dx.doi.org/10.1097/SLA.0000000000000629>

85. Aveling EL, McCulloch P, Dixon-Woods M. A qualitative study comparing experiences of the surgical safety checklist in hospitals in high-income and low-income countries. *BMJ Open* 2013;**3**:e003039. <http://dx.doi.org/10.1136/bmjopen-2013-003039>
86. Urbach DR, Govindarajan A, Saskin R, Wilton AS, Baxter NN. Introduction of surgical safety checklists in Ontario, Canada. *N Engl J Med* 2014;**370**:1029–38. <http://dx.doi.org/10.1056/NEJMsa1308261>
87. Hill AB. The environment and disease: association or causation? *Proc R Soc Med* 1965;**58**:295–300. <http://dx.doi.org/10.1177/0141076814562718>
88. Vincent C, Burnett S, Carthey J. Safety measurement and monitoring in healthcare: a framework to guide clinical teams and healthcare organisations in maintaining safety. *BMJ Qual Saf* 2014;**23**:670–7. <http://dx.doi.org/10.1136/bmjqs-2013-002757>
89. Morgan L, Pickering S. Observing and categorising process deviations in orthopaedic surgery. *Proc Hum Factors Ergon Soc Annu Meet* 2011;**55**:685–9. <http://dx.doi.org/10.1177/1071181311551142>
90. Hadi M, Robertson E, Morgan L. Evaluation of surgical team performance in elective operative theatres. *Int J Surg* 2011;**9**:538. <http://dx.doi.org/10.1016/j.ijisu.2011.07.216>
91. Morgan L, Hadi M, Pickering S, Robertson E, Griffin D, Collins G, *et al*. The effect of teamwork training on team performance and clinical outcome in elective orthopaedic surgery: a controlled interrupted time series study. *BMJ Open* 2015;**5**:e006216. <http://dx.doi.org/10.1136/bmjopen-2014-006216>
92. Morgan L, New S, Robertson E, Collins G, Rivero-Arias O, Catchpole K, *et al*. Effectiveness of facilitated introduction of a standard operating procedure into routine processes in the operating theatre: a controlled interrupted time series. *BMJ Qual Saf* 2015;**24**:120–7. <http://dx.doi.org/10.1136/bmjqs-2014-003158>
93. Morgan L, Pickering SP, Hadi M, Robertson E, New S, Griffin D, *et al*. A combined teamwork training and work standardisation intervention in operating theatres: controlled interrupted time series study. *BMJ Qual Saf* 2015;**24**:111–19. <http://dx.doi.org/10.1136/bmjqs-2014-003204>
94. Robertson E, Morgan L, New S, Pickering S, Hadi M, Collins G, *et al*. Quality improvement in surgery combining Lean improvement methods with teamwork training: a controlled before–after study. *PLOS ONE* 2015;**10**:e0138490. <http://dx.doi.org/10.1371/journal.pone.0138490>
95. McCulloch P, Morgan L, New S, Catchpole K, Robertson E, Hadi M, *et al*. Combining systems and teamwork approaches to enhance the effectiveness of safety improvement interventions in surgery: the Safer Delivery of Surgical Services (S3) Programme [published online ahead of print December 22, 2015]. *Ann Surg* 2015.
96. Catchpole K. Human Factors and Ergonomics in Surgical Safety: That Was Then, This Is Now. In Albolino S, editor. *Healthcare Ergonomics and Patient Safety (HEPS 2011)*. Oviedo: Taylor & Francis/CRC Press; 2011.
97. Robertson ER, Morgan L, Bird S, Catchpole K, McCulloch P. Interventions employed to improve intrahospital handover: a systematic review. *BMJ Qual Saf* 2014;**23**:600–7. <http://dx.doi.org/10.1136/bmjqs-2013-002309>
98. Catchpole K, Hadi M. Simulation provides a window on the quality and safety of the system. *Resuscitation* 2011;**82**:1375–6. <http://dx.doi.org/10.1016/j.resuscitation.2011.07.042>

Appendix 1 Intentional rounding area sheet

Nurse CSW

Intentional Rounding

Oxford University Hospitals 

4 P's completed – positioning/personal needs/pain/ placement Patient not here Date: ___/___/___

	PURPLE A	12	13	21.1	21.2	21.3	21.4	22	14.1	14.2	14.3	14.4	15	23	
AM	7														
	8														
	9														
	10														
	11														
PM	12														
	1														
	2														
	3														
	4														
	5														
	6														
	7														
	8														
	9														
10															
AM	11														
	12														
	1														
	2														
	3														

Appendix 2 Intentional rounding individual patient sheet

Rm/bed no: _____

Affix pt sticker here before filing

Please stamp the corresponding box with your role

Date:					
7am	8am	9am	10am	11am	12pm
1pm	2pm	3pm	4pm	5pm	6pm
7pm	8pm	9pm	10pm	11pm	12am
1am	2am	3am	4am	5am	6am
Date:					
7am	8am	9am	10am	11am	12pm
1pm	2pm	3pm	4pm	5pm	6pm
7pm	8pm	9pm	10pm	11pm	12am
1am	2am	3am	4am	5am	6am

For this trial, please put completed forms in the box at the nursing station

Appendix 3 Ward round data collection pro forma

Observer:		Sheet: _/_	Date: / /20	Start: _:_	End: _:_
Bed number					
Pt initials					
People	SpR/Fellow	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan
	SHO/FY	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan
	Nursing	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan
	Other:_____	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan
	Patient	<input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan	<input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> plan
Charts	Medical notes	<input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> written	<input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> written	<input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> written	<input type="checkbox"/> present <input type="checkbox"/> input <input type="checkbox"/> written
	Obs chart	<input type="checkbox"/> present <input type="checkbox"/> review	<input type="checkbox"/> present <input type="checkbox"/> review	<input type="checkbox"/> present <input type="checkbox"/> review	<input type="checkbox"/> present <input type="checkbox"/> review
	Drug chart	<input type="checkbox"/> present <input type="checkbox"/> review	<input type="checkbox"/> present <input type="checkbox"/> review	<input type="checkbox"/> present <input type="checkbox"/> review	<input type="checkbox"/> present <input type="checkbox"/> review
Drug R/v	Analgesia r/v	Y / N	Y / N	Y / N	Y / N
	VTE considered?	Y / N	Y / N	Y / N	Y / N
	Abx considered?	Y / N	Y / N	Y / N	Y / N

Examination	Neuro	<input type="checkbox"/> GCS <input type="checkbox"/> LOC <input type="checkbox"/> motor <input type="checkbox"/> sensation <input type="checkbox"/> reflexes <input type="checkbox"/> CN <input type="checkbox"/> eyes <input type="checkbox"/> Cerebellum	<input type="checkbox"/> GCS <input type="checkbox"/> LOC <input type="checkbox"/> motor <input type="checkbox"/> sensation <input type="checkbox"/> reflexes <input type="checkbox"/> CN <input type="checkbox"/> eyes <input type="checkbox"/> Cerebellum	<input type="checkbox"/> GCS <input type="checkbox"/> LOC <input type="checkbox"/> motor <input type="checkbox"/> sensation <input type="checkbox"/> reflexes <input type="checkbox"/> CN <input type="checkbox"/> eyes <input type="checkbox"/> Cerebellum	<input type="checkbox"/> GCS <input type="checkbox"/> LOC <input type="checkbox"/> motor <input type="checkbox"/> sensation <input type="checkbox"/> reflexes <input type="checkbox"/> CN <input type="checkbox"/> eyes <input type="checkbox"/> Cerebellum
	Wound				
	Other system				
Plan	Lab tests	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt
	Radiology	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt
	Referrals	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt
	D/c planning	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt
	Drains	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt
	Tubes/lines	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt
	Catheters	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt
	Mobilisation	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt	<input type="checkbox"/> Request <input type="checkbox"/> Rpt
Glitches	Concurrent	<input type="checkbox"/> Conversations <input type="checkbox"/> Tasks	<input type="checkbox"/> Conversations <input type="checkbox"/> Tasks	<input type="checkbox"/> Conversations <input type="checkbox"/> Tasks	<input type="checkbox"/> Conversations <input type="checkbox"/> Tasks
	Other	<input type="checkbox"/> Ext. distraction <input type="checkbox"/> Multi-tasking	<input type="checkbox"/> Ext. distraction <input type="checkbox"/> Multi-tasking	<input type="checkbox"/> Ext. distraction <input type="checkbox"/> Multi-tasking	<input type="checkbox"/> Ext. distraction <input type="checkbox"/> Multi-tasking

	Searching for	<input type="checkbox"/> Notes <input type="checkbox"/> Staff	<input type="checkbox"/> Notes <input type="checkbox"/> Staff	<input type="checkbox"/> Notes <input type="checkbox"/> Staff	<input type="checkbox"/> Notes <input type="checkbox"/> Staff
--	---------------	------------------------------------------------------------------	------------------------------------------------------------------	------------------------------------------------------------------	------------------------------------------------------------------

Appendix 4 Subarachnoid haemorrhage care project case note template

Pt ID	Referrer	Neurological r/v	Bleed date	Glasgow Coma Scale score bleed	Referral deficit	A&E date	A&E time	Referral date	Referred to	Admit date	Admit time	Glasgow Coma Scale score admit	Admit location
-------	----------	------------------	------------	--------------------------------	------------------	----------	----------	---------------	-------------	------------	------------	--------------------------------	----------------

A&E, accident and emergency department; admit, admission; Pt ID, patient identification number; r/v, review.

CTA date	CTA time	DSA date	DSA time	Aneurysm treatment type	Aneurysm treatment date	ITU admit?	Daily r/v? Number of days	Surg cons r/v? Number of days?	IR cons r/v? Number of days?
----------	----------	----------	----------	-------------------------	-------------------------	------------	---------------------------	--------------------------------	------------------------------

admit, admission; cons, consultant; CTA, computed topography angiography DSA, digital subtraction angiogram; IR, intentional rounding; r/v, review; surg, surgeon.

Daily blood (for 14 days), number of days?	Minimum 4-hourly observations pre operation (y/n)	Minimum 4-hourly observations post-operation days 1-2 (y/n)	4-hourly observation post-operation days 3-5 (y/n)	4-hourly observations post-operation days 5-10 (y/n)	4-hourly observation post-operation days > 10 (y/n)
--------------------------------------------	---------------------------------------------------	-------------------------------------------------------------	----------------------------------------------------	------------------------------------------------------	-----------------------------------------------------

n, no; y, yes.

(60 mg 4-hourly, or 30 mg 2-hourly for 21 days) (y/n)	3 l/day (number of days)	Length of stay	Number of days admission (physiotherapy)	Discharge date	Discharge Glasgow Coma Scale score	Discharge deficit	Any other complications?
-------------------------------------------------------	--------------------------	----------------	------------------------------------------	----------------	------------------------------------	-------------------	--------------------------

Appendix 5 Subarachnoid haemorrhage care project process map

Care of the SAH patient	(1) SAH patient comes to JR Hospital	(1a) Patient bleeds and <ul style="list-style-type: none"> presents to DGH, presents to JR A&E, or already in-house at JR for other reason 	Diagnosis unknown (bed mgr accepts pt with incomplete information)
		(1b) Phone call to JR neurosurgery reg	Diagnosis equivocal (SAH is suspected and patient is transferred but later found not to be SAH)
		(1c) Neurosurgery (reg/consultant) +/- neuroradiology (reg/consultant) review of CT, decision to transfer <ul style="list-style-type: none"> Bed manager notified Floor coordinator → staff nurse notified Ward clerk manager → ward clerk notified Neurosurgery reg notifies junior dr 	No weekend (Saturday) coverage of ward clerks
		(1d) Clerk obtains previous clinical notes and assembles chart	
	(2) Initial clerking, resuscitation	(2a) Pt arrives on ward or ITU	
		(2b) Pt seen by nurse , who then bleeps junior dr and completes admission assessment	
		(2c) Pt seen by junior doctor then neurosurgery reg	
	(3) Treatment, initial recovery period	(3a) Pt taken to radiology for CT angiogram +/- coiling (or, to theatre for craniotomy w/clipping – rare)	Follow up instructions not given or unclear
		(3b) Post procedure instructions and follow up from neuroradiologist	
		(3c) Pt taken to theatre recovery area	
		(3d) Pt returns to ward or admitted to ITU	
	(4) Monitoring for delayed cerebral ischaemia	(4a) Clinical review by doctor , nurse , neurorad , with input from OT , physio , pharmacist , +/- ophthalmology , etc as required	
		(4b) Nimodipine, obs, 3L fluid/day	
		(4c) +/- CT scan if suspected hydrocephalus, +/- EVD,...	
		(4d) Transfer to ITU if clinical deterioration and/or transfer to ward from ITU when clinically stable	
	(5) Discharge and follow	(5a) Pt stable, plan for discharge	
		(5b) Junior dr writes eIDD, TTO, nurse prints eIDD and	

up	completes discharge assessment	
	(5c) Clerk arranges for follow up with neurosurgery and/or neuroradiology arranges for follow up imaging	
	(5d) mRS at discharge documented by junior dr	
	(5e) Clerk breaks down and files chart	
	(5f) Pt returns for follow up: neurosurgery outpatient clinic and/or radiology	

People:	junior doctor	Nurse	neurosurgery reg or consultant	Clerk	neuroradiologist	other
---------	---------------	-------	--------------------------------	-------	------------------	-------

Appendix 6 Subarachnoid haemorrhage pro forma

JOHN RADCLIFFE HOSPITAL		NAME
NEUROSURGERY DAILY REVIEW		DATE OF BIRTH (Patient sticker here)
Consultant:		HOSPITAL NO.
Diagnosis:		
DATE AND TIME	CLINICAL NOTES (Each entry must be signed)	
WR Lead:		
Observations	Investigations / Results / Procedures	
Stable & Apyrexial <input type="checkbox"/>	Bloods:	<u>Procedure(s) and date(s)</u>
Temp	CT:	
HR	CSF:	
BP		
Sats		
RR		
Symptoms, Examination, Assessment		
GCS:		
O/E:		
Plan		
.....		
Thromboprophylaxis	Review <input type="checkbox"/>	Plan:
.....		
Steroids	Review <input type="checkbox"/>	Plan:
.....		
Antibiotics	Review <input type="checkbox"/>	Plan:
.....		
Therapy Review required?		
<input type="checkbox"/> PT	<input type="checkbox"/> OT	<input type="checkbox"/> Pain Team
<input type="checkbox"/> SALT	<input type="checkbox"/> Dietician	
.....		
Discharge Plan		
Planned Discharge Date:	EIDD/TTO Done <input type="checkbox"/>	
To location:	Referral done <input type="checkbox"/>	Date:
SHO:	Signature:	Bleep:
V3 25/02/2013		

DATE AND TIME	CLINICAL NOTES (Each entry must be signed)
	<input type="checkbox"/> Consultant Review

ATTENTION:
Please strike through ALL blank areas above when starting a new sheet.

V2 25/02/2019

Appendix 7 Senior house officer leavers survey (August 2014)

Q1 The Department meets the Trust's values of:

Excellence	<i>Strongly disagree</i>	<i>Disagree</i>	<i>Neither agree nor disagree</i>	<i>Agree</i>	<i>Strongly agree</i>
Compassion	<i>Strongly disagree</i>	<i>Disagree</i>	<i>Neither agree nor disagree</i>	<i>Agree</i>	<i>Strongly agree</i>
Respect	<i>Strongly disagree</i>	<i>Disagree</i>	<i>Neither agree nor disagree</i>	<i>Agree</i>	<i>Strongly agree</i>
Learning	<i>Strongly disagree</i>	<i>Disagree</i>	<i>Neither agree nor disagree</i>	<i>Agree</i>	<i>Strongly agree</i>
Delivery	<i>Strongly disagree</i>	<i>Disagree</i>	<i>Neither agree nor disagree</i>	<i>Agree</i>	<i>Strongly agree</i>
Improvement	<i>Strongly disagree</i>	<i>Disagree</i>	<i>Neither agree nor disagree</i>	<i>Agree</i>	<i>Strongly agree</i>

Q2 In what three ways is patient care good in the Department?

Q3 Do you have any suggestions how this good care could be enhanced?

Q4 What three things concern you most about patient care in the Department?

Q5 Do you have any suggestions how these concerns could be addressed?

Q6 Please evaluate the following statements with regard to their impact on patient care in the Department

Members or the Department work together as a team	<i>Strongly disagree</i>	<i>Disagree</i>	<i>Neither agree nor disagree</i>	<i>Agree</i>	<i>Strongly agree</i>
The work processes in the Department enhance care	<i>Strongly disagree</i>	<i>Disagree</i>	<i>Neither agree nor disagree</i>	<i>Agree</i>	<i>Strongly agree</i>
The equipment and technology available enhance care	<i>Strongly disagree</i>	<i>Disagree</i>	<i>Neither agree nor disagree</i>	<i>Agree</i>	<i>Strongly agree</i>
The working environment enhances care	<i>Strongly disagree</i>	<i>Disagree</i>	<i>Neither agree nor disagree</i>	<i>Agree</i>	<i>Strongly agree</i>

Q7 How likely is it that you would recommend the Department's clinical care to a friend or colleague (0 – not at all likely, 10 – extremely likely)

Q8 How likely is it that you would recommend an SHO job in the Department to a friend or colleague? (0 – not at all likely, 10 – extremely likely)

Q9 How would you improve the SHO job in the Department?

Q10 Please add any further comments about your experience and / or suggestions for improvement

A decorative graphic consisting of numerous thin, parallel green lines that curve from the left side of the page towards the right, creating a sense of movement and depth.

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

Published by the NIHR Journals Library