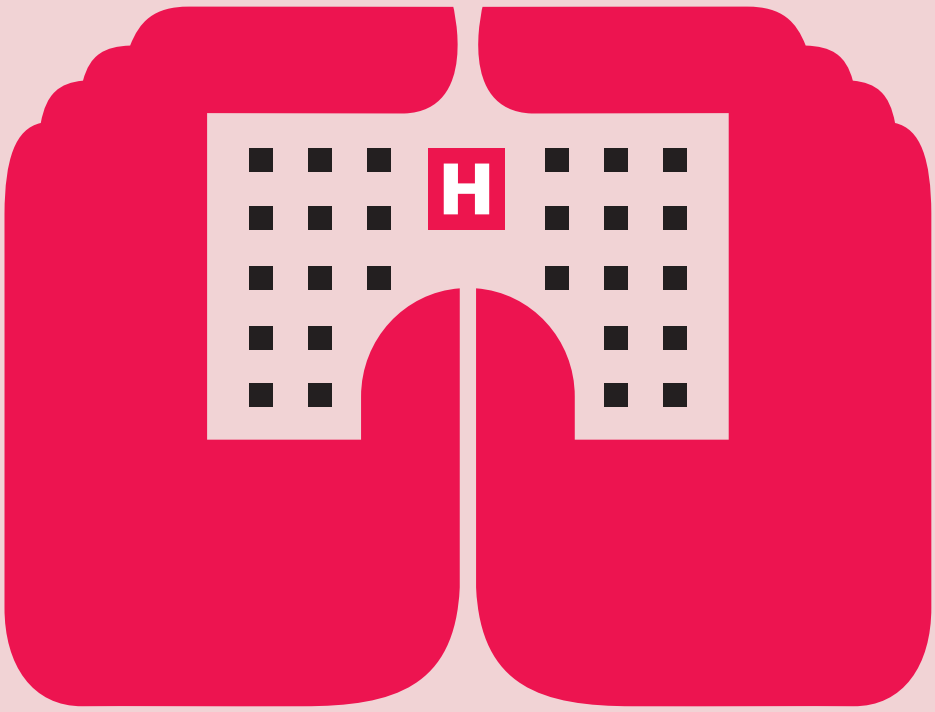
Evidence:

Safer Patients Initiative

phase one

Mixed-method evaluation of a large-scale organisational intervention to improve patient safety in four UK hospitals

February 2011



The Health Foundation
Inspiring Improvement

Identify Innovate Demonstrate Encourage

Evidence: Safer Patients Initiative phase one is published by the Health Foundation
90 Long Acre, London WC2E 9RA

ISBN 978-1-906461-21-8

Contributors

Nick Barber, professor of the practice of pharmacy, Department of Practice and Policy, School of Pharmacy, University of London.

Amirta Benning, programme manager, School of Health and Population Science, University of Birmingham.

Martin Carmalt, consultant physician, Royal Orthopaedic Hospital NHS Foundation Trust, Birmingham.

Sopna Choudhury, research associate, School of Health and Population Science, University of Birmingham.

Jeremy Dawson, research fellow, Work & Organisational Psychology Group, Aston Business School, Aston University.

Mary Dixon-Woods, professor of medical sociology, Department of Health Sciences, University of Leicester.

Bryony Dean Franklin, professor of medication safety and director, Centre for Medication Safety and Service Quality, Department of Practice and Policy, School of Pharmacy, University of London; Imperial College Healthcare NHS Trust, London.

Maisoon Ghaleb, lecturer in pharmacy practice/patient safety, Department of Practice and Policy, School of Pharmacy, University of London; School of Pharmacy, University of Hertfordshire.

Alan Girling, senior research fellow, School of Health and Population Science, University of Birmingham. Anu Suokas, PhD student, Department of Health Sciences, University of Leicester.

Karla Hemming, senior research fellow, School of Health and Population Science, University of Birmingham.

Richard Lilford, professor of clinical epidemiology, School of Health and Population Science, University of Birmingham.

Thirumalai Naicker, honorary research associate, School of Health and Population Science, University of Birmingham.

Ugochi Nwulu, research associate, School of Health and Population Science, University of Birmingham.

Gavin Rudge, data scientist, School of Health and Population Science, University of Birmingham.

Foreword

The Health Foundation is an independent charity that aims to improve the quality of healthcare across the UK. We are here to inspire and create the space for people, teams, organisations and systems to make lasting improvements to health services. Working at every level of the system, we aim to develop the technical skills, leadership, capacity and knowledge, and build the will for change, to secure lasting improvements to healthcare.

In 2004 we launched the first phase of the Safer Patients Initiative (SPI) programme. This large-scale, two-year initiative was the first major programme addressing patient safety in the UK. It had an ambitious overall aim of halving the number of adverse events within four participating hospitals (then a further 20) across the UK. The hospitals worked with safety experts from the US-based Institute for Healthcare Improvement to test ways of improving patient safety on an organisation-wide basis and develop their expertise. The SPI focused on improving the reliability of specific frontline care processes within four clinical areas and ensuring safety was a strategic priority by involving the chief executives and senior executive teams.

We appointed a consortium led by the University of Birmingham to conduct an evaluation of phase one of the programme. This evaluation sought to assess the wider organisational impact and therefore looked beyond the four pilot areas of the clinical interventions. It measured the average effect of the programme across a range of practices, based on the starting assumption that SPI would transform organisation-wide approaches to patient safety.

The evaluation reported that senior stakeholders were enthusiastic and knowledgeable about the programme and shared an understanding of its underlying theory. Through SPI, participating hospitals developed organisational learning about how to implement patient safety efforts in the future. There was, however, only modest penetration at ward level. Quality of monitoring of sick patients (which was an important SPI aim) improved both

in control hospitals and in hospitals participating in SPI, but control hospitals did not improve as much as SPI hospitals. A small improvement was found in staff attitudes to organisational climate in intervention hospitals. On a range of other measures and outcomes related to patient safety, an additive effect attributable to SPI was not detected.

The evaluators propose four potential reasons why the study did not detect improved practice:

- Improvements may have occurred at a level that eluded statistical detection.
- The study looked for organisational change and was not designed to detect changes in areas outside of acute medicine.
- The level of intervention may have been insufficient to create the anticipated changes.
- Improvements may surface in the longer term.

The achievements within the timeframe of the SPI were less than we had envisaged. However, the approaches championed in the programme continue to shape the work of many hospitals and health systems across the UK and continue to show results within individual organisations.

The external evaluation does not negate the positive results that individual hospitals reported. Rather, it challenges the wider aspiration that the intervention would lead to an organisation-wide effect. Arguably an organisational impact was not observed because the number of people exposed to the intervention was limited and the work largely remained a project rather than embedded in mainstream structures. The ambition of transforming organisations within an 18-24 month period was clearly considerable, and there were large gaps in data collection, and management and skills to overcome.

Improving patient safety is a significant challenge. The SPI began a long journey of building skills and capability in the participating organisations to measure and improve the effectiveness of care.

The degree of evaluation dedicated to this programme provides us with a unique opportunity to identify key lessons for improving patient safety and the challenges for achieving organisation-wide transformation. These include clarity about the theories of change underlying the programme; recognition of the scale of resource and organisational support required to make safety efforts work; and

improving understanding of how middle managers can be better supported and clinical ownership secured.

We have also gained important insights into the nature of evaluating complex, organisational interventions, including: the importance of aligning the evaluation design with an explicit programme theory; ensuring sensitivity to complex variation; the ability to capture the interactions between programme interventions and the local context and the importance of taking measurements at the level of the system where the intervention is taking place and where the impact can be expected to occur.

Dr Dale Webb
Director of Evaluation & Strategy
The Health Foundation

Contents

Figures and tables	viii
Abbreviations	x
Executive summary	xi
CHAPTER 1 Introduction	1
1.1 Policy background	1
1.2 The intervention	1
1.3 Selection of participating sites	4
CHAPTER 2 Methods	6
2.1 Framework for the evaluation	6
2.2 Sub-study 1: Strategic stakeholder interviews	10
2.3 Sub-study 2: Staff surveys	10
2.4 Sub-study 3: Ethnographic study	12
2.5 Sub-study 4: Error rates/quality of care among acute medical patients	13
2.6 Sub-study 5: Outcomes	21
CHAPTER 3 Results	23
3.1 Sub-study 1: Strategic stakeholder interviews	23
3.2 Sub-study 2: Staff survey	23
3.3 Sub-study 3: Ethnographic study	27
3.4 Sub-study 4: Error rates/quality of care among acute medical patients	36
3.5 Sub-study 5: Outcomes	53
CHAPTER 4 Discussion	59
4.1 Main findings	59
4.2 Strengths and weaknesses	61
4.3 Explaining the results	62
4.4 Interpretation	65
4.5 Next steps	66

REFERENCES 68

NOTES FROM THE AUTHORS 74

APPENDICES **76**

Appendix 1: Staff survey 76

Appendix 2: Stages of the drug use process 78

Appendix 3: Patient survey 81

Figures and tables

Figure 2.1:	Causal chain linking SPI to outcomes	7
Figure 2.2:	Staggering the data collection allowed use of the pre-SPI2 intervention data collection as control data for the evaluation of the first phase of SPI	9
Figure 3.1:	Correlation between knowledge and enthusiasm for SPI among strategic hospital stakeholders	24
Figure 3.2:	Prescription error rates and date of review	37
Table 1.1:	Key generic and specific elements of the SPI	3
Table 1.2:	Hospitals that participated in the SPI1	5
Table 2.1:	Summary of sub-studies comprising the evaluation of SPI1	8
Table 2.2:	Staff survey variables deemed relevant to the SPI	11
Table 2.3:	Detectable effect sizes, at 5% significance and 80% power for a sample of 800 case notes split equally between epochs	14
Table 2.4:	Vital signs that should be recorded	16
Table 2.5:	Appropriate clinical response for abnormal observations	17
Table 2.6:	Classification of errors and adverse events	20
Table 2.7:	Patient survey questions deemed relevant to SPI	22
Table 3.1:	Staff survey scores in SPI1 hospitals at the two periods	25
Table 3.2:	Staff survey scores in control and SPI1 hospitals at the two periods	26
Table 3.3a:	Medical history taking (% of patients asked)	38
Table 3.3b:	Medical history taking – differences between control and intervention hospitals, changes over time and the effect of SPI	38
Table 3.4a:	Vital signs – percentage compliance with standards	39
Table 3.4b:	Vital signs – differences between control and intervention hospitals, changes over time and the effect of SPI	40
Table 3.5a:	Appropriate clinical response – compliance with standards	42
Table 3.5b:	Appropriate clinical response – differences between control and intervention hospitals, changes over time and the effect of SPI	43
Table 3.6a:	Features of good care for specific classes of patient	44

Table 3.6b:	Steroids and antibiotics – differences between control and intervention hospitals, changes over time and the effect of SPI	44
Table 3.7a:	Prescribing errors	45
Table 3.7b:	Prescribing Errors – differences between control and intervention hospitals, changes over time and the effect of SPI	46
Table 3.7c:	Anti-coagulant prescribing errors	46
Table 3.7d:	Anti-coagulant prescribing errors – analysis	46
Table 3.8:	Examples of prescribing errors relating to each stage of the drug use process found in this study	48
Table 3.9a:	Reconciliation errors at admission	48
Table 3.9b:	Reconciliation errors at admission – differences between control and intervention, changes over time and the effect of SPI	49
Table 3.10:	Holistic review changes in ratings and numbers of adverse events and errors between control and SPI1 hospitals	51
Table 3.11:	Rates (per 100 patients) of errors identified by broad category of error – holistic review	52
Table 3.12:	Rates per 100 patients of adverse events identified by broad category of adverse events	54
Table 3.13:	Preventable adverse events identified as being strongly or certainly preventable out of the 1260 case notes reviewed in the holistic review	55
Table 3.14a:	Mortality rates	56
Table 3.14b:	Analysis of mortality rates	56
Table 3.15:	Patient survey scores in SPI1 hospitals at the two periods	58
Table 3.16:	Patient survey scores in control and SPI1 hospitals at the two periods	58

Abbreviations

A&E	Accident and emergency
BNF	British National Formulary
BTS	British Thoracic Society
CI	Confidence intervals
COPD	Chronic obstructive pulmonary disease
CURB	Confusion, urea, respiratory rate, blood pressure
EWSS	Early warning score systems
FMEA	Failure mode effects analysis
ICC	Intra-class correlation coefficients
ICU	Intensive care unit
IHI	Institute for Healthcare Improvement
IOM	Institute of Medicine
OR	Odds ratios
PDSA	Plan, do, study, act
QALY	Quality of life adjusted year
QUERI	Quality nhancement Research Initiative
SBAR	Situation background assessment recommendation
SE	Standard errors
SPI	Safer Patients Initiative
SPI1	Pilot phase hospitals of the Safer Patients Initiative
SPI2	Second phase hospitals of the Safer Patients Initiative
U&E	Urea and electrolytes

Executive summary

Background

Improving patient safety is one of the most challenging and pressing issues facing practitioners, managers and policy makers, but to date there is little evidence of how best to do it. Progress in improving safety has been frustratingly slow and there is keen interest in finding ways to make sustained change.

In the last five years, attention has turned to approaches that involve interventions at an organisational level and include participative elements and the active involvement of clinicians and managers. The approaches are both top down – engaging with senior managers and other strategic stakeholders – and bottom up, seeking to capitalise on the skills, knowledge and interventions of clinicians and managers. An example is the well-publicised Saving 100,000 Lives campaign promoted by the Institute for Healthcare Improvement (IHI), a US-based not-for-profit organisation dedicated to improving healthcare.

The Safer Patient Initiative

In 2005, the Health Foundation developed the IHI methodology for use in the NHS and launched the Safer Patients Initiative (SPI). This is the evaluation report of the first phase of that programme, SPI1, which began in January 2005 and provided direct support to four NHS hospitals as they implemented an organisation-wide patient safety programme for 18 months.

The stated aim was to make a 50% reduction in the total number of adverse events at the four trial hospitals. Each of the hospitals was carefully selected through a competitive bidding process. The Health Foundation gave £775,000 to each of the four hospitals to secure the services of IHI and to provide the capacity for change. The evaluation looked for organisational level change, using interviews to evaluate changes in culture and knowledge, and case note review to investigate changes in adverse events. This included

how practitioners responded to acutely ill and deteriorating patients and how well they recorded vital signs. The results of this quantitative analysis were compared with 18 control hospitals.

Results

Overall, there was a marked and significant improvement in the response to acutely ill patients during the study, including the recording of vital signs. However, this was replicated at the control hospitals and cannot be attributed to the SPI. The evaluation found no reduction in errors or in adverse events in the patient group examined – patients with acute respiratory disease.

On the qualitative side, the interviews with senior stakeholders found that they were generally enthusiastic and knowledgeable about the SPI and shared an understanding of the programme and its underlying theory of change. Ward staff, on the other hand, tended to know little about the SPI procedures, practices and principles, or viewed them as top down rather than something they had been involved in developing. There was little evidence of a shared sense of ownership and some evidence of a sense of elitism that had grown up around those who had taken part in the initiative. The SPI had little measurable impact on ward level staff, leading to the conclusion that its impact at ward level was, at best, modest.

Conclusions

For those who were hoping for dramatic improvements in safety and a reduction in adverse events, this evaluation will be disappointing. The results will not be a surprise to others who believe that achieving quality improvements and reducing error rates through management initiatives is difficult. Either way, the results warrant further discussion, not least as there are several reasons that the study may have failed to identify improved practice.

For a start, improvements may have occurred on a scale that eluded statistical detection. The English threshold for judging the cost effectiveness of a clinical intervention is about £30,000 per Quality Adjusted Life Year (QALY). To be cost effective, the SPI would need to save fewer than seven lives with a mean duration of five years to justify the investment of £775,000 per hospital. In a large-scale study such as this in a hospital where hundreds of deaths take place each year, a signal of this magnitude would be lost in the noise.

The study looked for organisational level change and was not designed to detect changes in areas outside acute medicine. For example, intensive care units (ICU) and surgical departments were not included in this evaluation.

The qualitative elements of the study raise some questions about the design and implementation of the SPI. The initiative secured buy-in from senior managers but not from ward level staff. There was some evidence from the interviews that ward staff viewed the scale of the challenge as daunting and that the resource implications and degree of cultural change required was underestimated. It may simply take a long time for programmes such as the SPI to achieve this effect. It may take more resources: £775,000 spent over 18 months in hospitals with annual budgets of £150m – £300m might simply be too small a dose, especially when little of that money reached the sharp end of practice.

Hospitals did report effects from SPI participation, including heightened managerial awareness or commitment (or both) to safety, and organisational learning about how to implement safety improvement in future. Participating in the SPI may secure greater long-term commitment to quality and safety, and improvements made in participating hospitals will either surface at a later date or prove more sustainable than the improvements seen in control hospitals. This hypothesis can only be tested with further data collection.

Next steps

The results presented here should neither be a cause for despair nor an excuse to search for positive results to prove the value of the SPI as it moves forward. The challenge is to build on the observations in this study. The staff we interviewed had their own ideas about how to do this, suggesting more support for middle managers, engaging clinical leaders at an earlier stage and encouraging clinical ownership as well as reducing the number of areas to be tackled. Senior stakeholders recognised that initiatives requiring more paperwork were unlikely to be popular at ward level and acknowledged that organisational change may require large-scale resourcing and structural support.

It is important to remember that improving patient safety is hard and that achieving change is likely to be a marathon rather than a sprint. Far from abandoning the topic of patient safety improvement or decrying the SPI, the results point to promising and reasonable hypotheses about how to introduce a more holistic approach to safety.

Chapter 1

Introduction

1.1 Policy background

How best to secure improvements in patient safety is one of the most pressing and challenging questions facing practitioners, managers and policy-makers.

Increasing effort, attention and resource have been focused on patient safety since the publication of two key reports in 2000 (the UK chief medical officer's *Organisation with a Memory*,[1] and the US Institute of Medicine's (IOM) *To Err is Human*).[2] However, a review published five years after the IOM report found that progress had been frustratingly slow and that there was little evidence of systematic improvements in safety.[3]

There is now keen interest in finding ways to make more sustained progress. Patient safety initiatives working at an organisational level that include participative principles, such as the involvement of workers in risk management, may provide the greatest hope of improving patient safety. In the UK, the Health Foundation's Safer Patients Initiative (SPI) (see table 1.1) is an important example of such an approach.

1.2 The intervention

The first phase of the SPI programme (SPI1) began in January 2005 and was mentored by the Institute for Healthcare Improvement (IHI), a US-based not-for-profit organisation dedicated to improvement in healthcare, over an 18-month period. Hospitals were expected to embed and spread learning following the IHI mentoring.

The Health Foundation provided funding of £775,000 per hospital to secure the services of the IHI and to provide the capacity for change in the individual hospitals. SPI had a number of features similar to the well publicised US Saving 100,000 Lives campaign,[6;7] and set out to 'transform organisational approaches to delivering

safe care'.[8] The IHI aimed to penetrate deeply into organisations, changing not only specific processes and standards, but also the attitudes, motives and behaviour of staff and how they understand the nature of their work.

The SPI programme is complex, consisting of many components, both specific and generic. The SPI focused on improving specific front-line care processes within designated clinical areas (general wards, critical care, perioperative care and medicines management).

Specific interventions targeted problems identified by the IHI, including medication error and identifying and responding to patient deterioration. The IHI worked to ensure 'that safety was a strategic priority through involvement with chief executives and senior executive teams'.[8]

The programme aimed to secure commitments to safety, culture and behaviour across hospitals and to improve performance in relation to patient safety.

Generic interventions – those not specific to any particular clinical problem – included training on how to conduct a structured process to identify problems and then to develop and evaluate customised solutions using the plan, do, study, act (PDSA) technique.[9;10] This technique is based on quality improvement methodology with a long provenance going back to Deming.[11] Staff at hospitals were also asked to participate in patient safety culture surveys.

As part of the programme, the IHI led selected hospital staff in four learning sessions. Here, teams of 15–20 people (known as change agents) from each hospital had time away from their normal duties to learn the principles and methods of safety science.

The times between meetings were described as action periods and staff were encouraged to work towards safety goals. Hospitals were asked to remain in contact with the IHI throughout these periods and to team up with other intervention hospitals via site visits, conference calls and a microsite.

This paper focuses on the first phase of the Safer Patients Initiative (SPI1) involving four UK NHS hospitals.

The Health Foundation has subsequently rolled out the SPI methodology to a second phase of hospitals (SPI2), and the evaluation of this is reported in *Evidence: Safer Patients Initiative phase two*.[12]

Table 1.1: Key generic and specific elements of the SPI

General aim: to avoid unnecessary harm, pain or suffering as a result of error in medical interventions	
Aim	**Method**
Generic improvement in the system to reduce adverse events whatever their cause	
Building a culture of safety and good leadership Training to enable organisations to identify problems and develop and evaluate methods to reduce risk Fostering an understanding of the principles of safe practice	a. Collaborative residential learning sessions with IHI faculty b. Web-based learning and site visits from IHI c. Leadership projected in part by management walk rounds d. Know how for PDSA cycles e. Electronic information sharing facility – for example to share results of PDSA cycles f. Participation in safety culture surveys using the Sexton tool
Specific interventions	
1. Identifying and responding to deteriorating patients. To reduce: • Need for 'crash calls' • Avoidable mortality	a. Review of 50 deaths b. Tools for monitoring patients' condition and for triggering action. These tools include a pro forma to record vital signs and other salient information (EWSS) c. Promote the use of risk (severity) scores d. Establishing a rapid response team
2. Reducing medication error	a. Medication safety assessment by involving staff in FMEA – educating staff to identify and remedy weak links in medication practice from prescribing to administration and monitoring b. Tool to reduce adverse events to anti-coagulant therapy c. Education to improve medicine reconciliation on admission
3. Communication between staff to reduce adverse events/mortality	a. SBAR tool to ensure that information is communicated in a structured way b. Safety briefings – briefings at shift changes to ensure staff are aware of relevant information for patients
4. Infection control, including methicillin-resistant Staphylococcus aureus (MRSA)	a. Perioperative antibiotics to reduce surgical site infection b. Catheter insertion and maintenance drill to prevent central-line infections in intensive care c. Following the tenets of ventilator guidelines (bundles) to reduce ventilator-acquired pneumonia, venous thromboembolism and stress ulcers in intensive care units d. Improve hand hygiene, for example, by means of prominently displayed posters

1.3 Selection of participating sites

The four SPI hospitals were selected following a competitive bidding process to demonstrate that they would be receptive to the intervention. A review panel with an international perspective as well as safety, clinical, and organisational expertise, was convened by the Health Foundation to select the sites.

The panel used a three-stage selection process. The first involved analysing all applications against explicitly agreed criteria, including:

- leadership commitment
- capacity and capability
- openness, transparency and communication
- exemplar status.

A shortlist of eight organisations was entered into the second stage, which involved site visits to explore:

- information outlined in the hospitals' applications.
- feasibility and sustainability of the programme within the specific hospital's context.

The third stage was a selection panel meeting to consider the bids against these criteria:

- capacity and capability
- leadership commitment
- patient involvement
- openness and transparency
- willingness and capacity to be an exemplar for others
- sustainability and believability.

An explicit assessment of current safety work was not a criterion for selection at any stage of the process. The four participating hospitals are described in table 1.2.

The Health Foundation commissioned and funded an independent evaluation of the programme, the results of which are reported here.

Table 1.2: Hospitals that participated in the SPI1

Hospital no.	Rural/ urban	Bed no.	Teaching status	A&E	ICU	Consultant (specialists) FTE
Hospital 1	Urban	625	Associate teaching hospital	Yes	Yes	112
Hospital 2	Rural	750	No	Yes	Yes	120
Hospital 3	Urban	903	Principal teaching hospital	Yes	Yes	242
Hospital 4	Rural	280	No	Yes	No	36

These hospitals are all part of the NHS and have no private beds. Figures provided as of October 2004.
A&E = accident and emergency department; ICU = intensive care unit; FTE = full time equivalent.

Chapter 2

Methods

2.1 Framework for the evaluation

The evaluation was based on a systems-wide approach that has been described in detail elsewhere, and used a mixed-method design.[13–17] In this approach, the system is conceptualised as the setting in which care is delivered. Building on Donabedian's causal chain,[18] five levels can be distinguished (figure 2.1):

- structure (for example, size of hospital and types of services provided – see table 1.2)
- management processes (for example, leadership style, management walk rounds)
- intervening (mediating) variables (for example, culture, morale, sickness absence rates) that connect management to clinical process
- clinical processes (error rates/compliance with evidence based care)
- outcomes (adverse events, mortality, patient satisfaction).

The data collection and analysis was organised around this conceptual model, using a series of five linked sub-studies:

- Management process were studied using qualitative interviews with strategic stakeholders (sub-study 1).
- Intervening variables were assessed using a quantitative staff survey (sub-study 2) and a qualitative study using ethnographic methods on hospital wards (sub-study 3).
- Clinical processes were studied by case note review (sub-study 4)
- Outcomes were studied by case note review (mortality and adverse events) and quantitative patient questionnaires (sub-study 5).

The qualitative studies (sub-studies 1 and 3) were conducted in the four SPI1 hospitals only. All other sub-studies were controlled before and after studies in 18 control and four intervention hospitals. Using before and after observations across control and SPI1 sites enables comparison of rates of change across control and SPI1 hospitals – an approach sometimes referred to as the 'difference in difference' method.[14] The sub-studies are summarised in table 2.1.

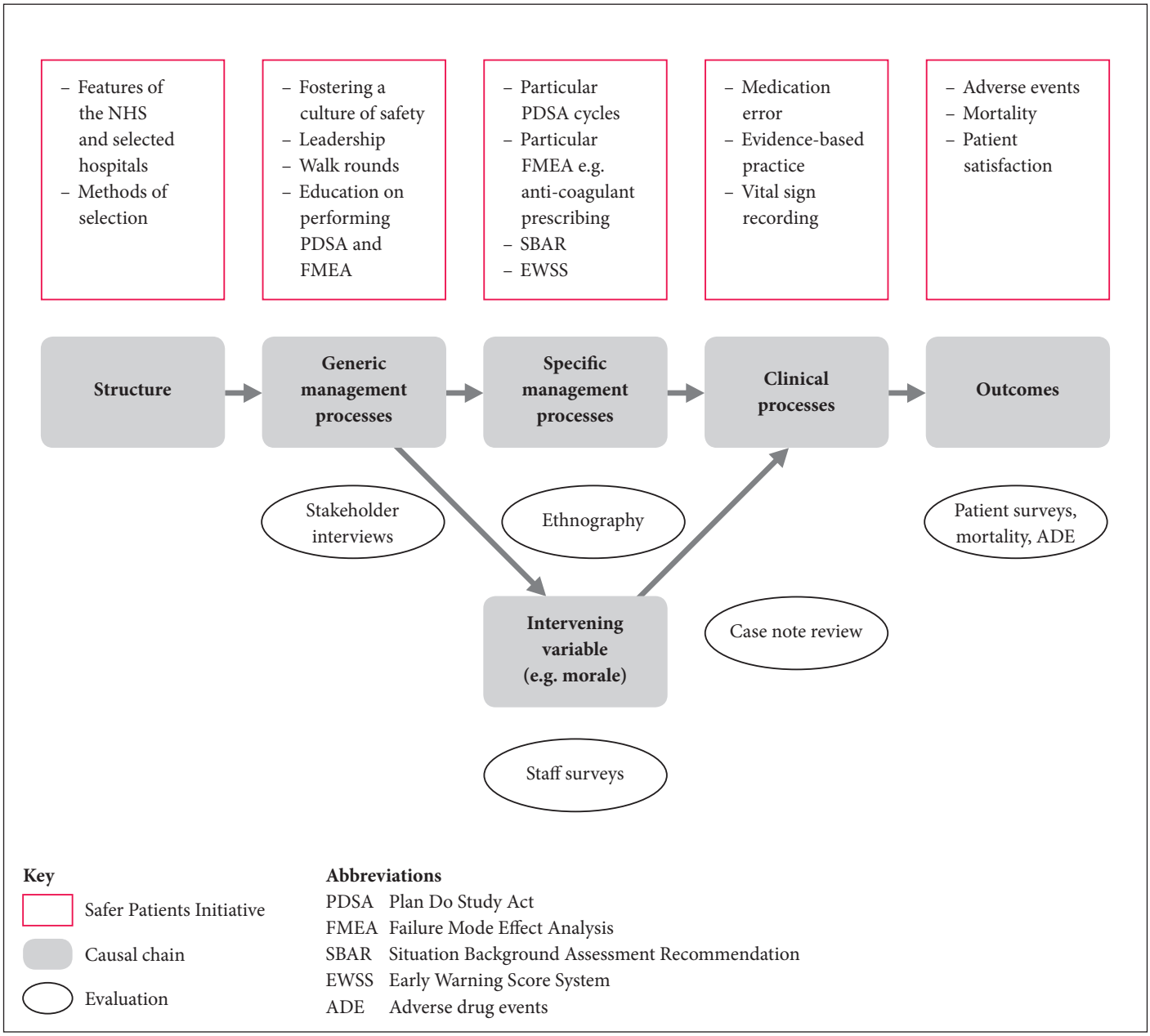


Figure 2.1: Causal chain linking SPI to outcomes. The five evaluation sub-studies were made at points across the chain to provide information on context, fidelity and effectiveness of SPI

Selecting the control hospitals capitalised on the evaluation of the second phase of the intervention (SPI2). The SPI2 intervention was scheduled to start after the completion of the SPI1 intervention phase.

For this reason, it was possible to use both control and intervention hospitals from SPI2 as controls for SPI1. This was achieved by choosing two separate pre-intervention epochs for SPI2. Thus, nine of the control hospitals for SPI1 were destined to be SPI2 intervention sites (selected by the Health Foundation) and nine were SPI2 matched control sites. SPI2 controls were selected using the following criteria:

Table 2.1: Summary of sub-studies comprising the evaluation of SPI1

Sub-study	Purpose	Location	Data collection	Analysis
1) Interviews with strategic stakeholders	Study impact at a senior management level. Arguably a necessary, if not sufficient, condition for effectiveness	SPI1 hospitals and strategic commissioners	Semi-structured interviews	Constant comparative method
2) Staff survey	Measure effects of SPI1 on staff morale, culture and opinion	Control and SPI1 hospitals	Validated structured questionnaire	• Comparison of control versus SPI1 hospitals: a. At baseline b. Over time i.e. difference in difference • Comparisons within control and SPI1 cohorts
3) Ethnography study at ward level	Discover the impact of SPI1 on 'hearts and minds' of those delivering care	Medical wards treating patients with acute respiratory disease in SPI1 hospitals	Ethnographic study consisting of three rounds of visits, including observations, interviews and focus groups	Constant comparative method
4) Clinical process measure	Measure effects of SPI on the quality of care being delivered using independent case note reviews (both explicit and holistic)	Medical wards treating patients with acute respiratory disease in control and SPI1 hospitals	Before and after intervention phase of SPI1	• Comparison of control versus SPI1 hospitals: a. At baseline b. Over time i.e. difference in difference • Comparisons within control and SPI1 cohorts • Measurement of reliability and learning/fatigue effects
5) Outcomes	Measure effects of SPI on • adverse events and mortality, • patient satisfaction	Medical wards treating patients with acute respiratory disease in control and SPI1 hospitals	Before and after study using • Case notes • Validated structured questionnaire	• Comparison of control versus SPI1 hospitals: a. At baseline b. Over time i.e. difference in difference • Comparisons within control and SPI1 cohorts • Measurement of reliability

- Only non-specialist acute trusts in England were considered.
- Pairs of trusts should have a similar directorate structure (using data from the NHS national staff survey).
- Pairs of trusts should have the same foundation or non-foundation status.
- Pairs of trusts should be similarly located in either urban or rural settings.
- Once these criteria were satisfied, the trust with the most similar size (usually within 1,000 staff) to the SPI2 trust was selected as the control trust.
- If a trust had more than one hospital, quantitative data collection was focused on the largest hospital with an ICU.

The method by which SPI2 hospitals could serve as controls for SPI1 is explained in figure 2.2.

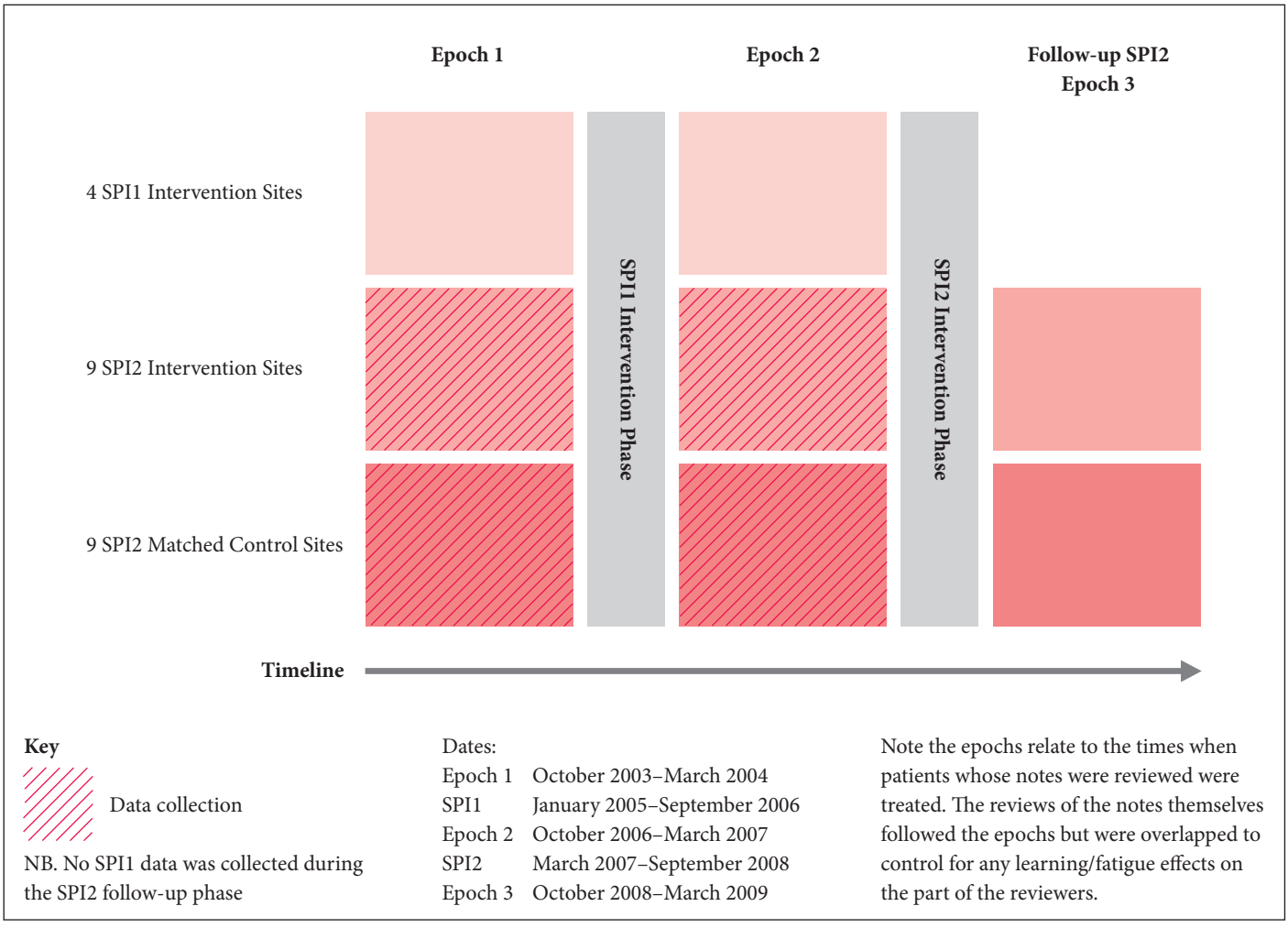


Figure 2.2: Staggering the data collection allowed use of the pre-SPI2 intervention data collection as control data for the evaluation of the first phase of SPI. In this way SPI1 and SPI2 have been "zipped together"

2.2 Sub-study 1: Strategic stakeholder interviews

The potential importance of the blunt end (senior management) in influencing the sharp end (where practitioners care for patients) is well recognised.[19]

This evaluation investigated how strategic stakeholders (people in senior positions) understood and responded to the SPI1 as a programme of change. We conducted semi-structured telephone interviews with 60 strategic-level hospital stakeholders and five external stakeholders involved in commissioning, designing and introducing the initiative.

The responses were tape-recorded and transcribed verbatim and subjected to two separate analyses.

Firstly, we analysed 65 transcripts based on the constant comparative method to generate thematic categories into which all participants' accounts could be categorised.[20]

Secondly, we quantified the responses given by hospital stakeholders. Hospital stakeholder transcripts were read by three independent reviewers who scored each individual's interview (on a one to 10 scale) for level of knowledge and level of enthusiasm.

Inter-observer agreement was tested using the intra-cluster correlation coefficient (ICC) statistic and Pearson's correlation coefficient to measure association between enthusiasm and knowledge.

2.3 Sub-study 2: Staff surveys

Strategic support for any programme may not reflect views at the sharp end of practice. We aimed to assess intervening variables such as staff morale, attitudes and various factors relevant to culture that might be affected by the SPI.

All hospitals in England participate in the National Staff Survey, a yearly survey run by the Healthcare Commission (now the Care Quality Commission). Arrangements were made to conduct the same survey, using the same survey methods and questionnaire in the three non-English hospitals in SPI1.

It was not possible for the surveys to take place at exactly the same time in all hospitals. The first round of the survey was undertaken

in the English hospital in Autumn 2004 and in the non-English hospitals in Spring 2005 (three months after the intervention had started). The second round of the survey was conducted in Autumn 2006 for the English hospital and Spring of 2007 for the non-English hospitals, but this time it was not possible to include one of the SPI1 hospitals because it was undergoing a merger and a period of major reorganisation. Data were therefore available for three of the four hospitals in the second round.

Questionnaires were sent to all staff in the four SPI1 hospitals. In the 18 control hospitals, a simple random sample of 850 staff was used instead, as this is the standard methodology employed by the Care Quality Commission. A sample size of 850 is such that an average 60% response rate (around 500 responses per site) would yield 95% confidence intervals of no greater than 10% for all scores within a single organisation. The detail of the survey methods is not repeated here but is available from www.nhsstaffsurveys.com

Table 2.2: Staff survey variables deemed relevant to the SPI

1.	Well structured appraisals[21;22]
2.	Working in well-structured teams[23]
3.	Witnessing potentially harmful errors or near misses in previous month
4.	Suffering work-related injury
5.	Suffering work-related stress
6.	Experiencing physical violence from patients/relatives
7.	Intention to leave
8.	Job satisfaction
9.	Quality of work-life balance
10.	Support from supervisors
11.	Organisational climate[24]

Around 28 survey scores are regularly reported by the Care Quality Commission (although the precise number has varied from year to year according to the precise content of the questionnaires). Eleven of these (table 2.2) were identified at the start of the evaluation as being of likely relevance to the SPI programme. This was either because they reflect safety issues directly, or because they relate to working practices known from research to be linked to safety and health outcomes.

Details of these questions and how they are calculated can be found in appendix 1.[21;22]

Differences between the control and SPI1 hospitals, in terms of changes between the two survey periods, were tested using a generalised linear mixed model with SPI1/control and survey period as fixed factors (with interaction), and hospital as a random factor.

To control for known differences between groups of staff, the following background factors were included as covariates in the models:

- age
- sex
- ethnic background (white or other)
- occupational group (nursing/midwifery, medical/dental, allied health professional/scientific & technical, admin/clerical, general management, maintenance/ancillary, or other)
- length of service, and management status (line manager or not).

A statistical correlation for multiple observations was not applied but the confidence intervals were set at 0.99 ($p<0.01$).

2.4 Sub-study 3: Ethnographic study

An ethnographic study was conducted at ward level in each of the four hospitals. Three rounds of data collection were undertaken.

1. A week-long visit to each of the wards involving approximately 150 hours of observations and 47 interviews with different types of ward staff, focusing on general issues relating to patient safety and the SPI. These visits were conducted between April and September 2006 as the SPI was being rolled out.
2. A week-long second visit to each ward, involving approximately 150 hours of observations and 41 interviews, this time with a particular (although not exclusive) focus on observing the patients' condition and responding to abnormalities, thus allowing insight into the early warning and rapid response systems used to detect and support deteriorating patients. These visits were conducted from April to June 2007, during the embedding and spread phase of SPI.
3. A third visit involving three focus groups at each site (one at study ward level, one involving people with patient safety/SPI responsibilities and one at strategic level). In these focus groups, preliminary findings from the first two visits were fed back and reflections were sought on the SPI and on the way forward for patient safety. These visits were carried out from May to July 2008, towards the end of the formal completion of the SPI programme in the SPI1 hospitals (September 2008).

Data analysis was based on the constant comparative method.[20] For the interviews, initial open codes were revised, expanded and collapsed as the analysis progressed and then organised into categories in a coding scheme, through which data was processed. This was facilitated by the use of NVivo software. For focus groups and field notes, simple coding procedures were used to categorise the data. Categories were inspected to build a theoretically-informed interpretation. In order to ensure anonymity, extracts from the data have not been labelled by site.

2.5 Sub-study 4: Error rates/quality of care

Case note selection criteria

Resources would not allow all relevant clinical topics in table 1.1 to be examined so it was necessary to be selective, focusing on patients over the age of 65 with acute respiratory disease. This cohort of patients was selected because:

- Improving recognition and response to acute deterioration in a patient's condition was a specific SPI target (topic 1 in table 1.1), and patients admitted with acute respiratory disease are at high-risk of such deterioration.[25;26]
- A single set of case notes could be used to assess end points targeted by several SPI interventions (see table 1.1), for example early warning score systems (EWSS) in managing the deteriorating patient, the use of failure mode effects analysis FMEA in targeting prescribing errors (particularly high-risk medicines such as anticoagulants), and the deployment of PDSA cycles to reduce reconciliation errors where previously prescribed medicines are inadvertently discontinued on admission
- There is a high incidence of co-morbidities in people aged over 65, making this a high-risk and hence a potentially error-rich population in which an effective intervention might yield detectable improvements
- There was evidence that monitoring and medication practice was sub-optimal in NHS hospitals, thus providing sufficient headroom for improvements to be detected with samples of affordable size[27;28]
- It was important to focus on an area where all four intervention hospitals would be implementing a specific SPI.

Table 2.3: Detectable effect sizes, at 5% significance and 80% power for a sample of 800 case notes split equally between epochs

For example, if baseline compliance with a standard was 50% then an improvement to 65% or a deterioration to 35% would be detectable. The assumed analysis adjusts for unexplained variation between hospitals.

Baseline Proportion	Modified proportions detectable with 80% power	
0.05	0.14	0.00
0.10	0.21	0.02
0.15	0.27	0.05
0.20	0.34	0.09
0.25	0.39	0.13
0.30	0.45	0.17
0.35	0.50	0.21
0.40	0.56	0.25
0.45	0.61	0.30
0.50	0.65	0.35
0.55	0.70	0.39
0.60	0.75	0.44
0.65	0.79	0.50
0.70	0.83	0.55
0.75	0.87	0.61
0.80	0.91	0.66
0.85	0.95	0.73
0.90	0.98	0.79
0.95	1.00	0.86

Case note assembly (and statistical power calculation)

We collected case notes from both the four intervention and 18 control hospitals from two time periods (epoch 1 and epoch 2) that preceded and followed the SPI1 intervention period (see figure 2.2).

The number of SPI1 hospitals was fixed and we spread control observations across a greater number of hospitals to provide a more robust sample. We aimed to analyse 100 case notes from each SPI1 hospital per epoch (800 in total) and 15 from each control hospital per epoch (540 in total). This would give 80% power to detect effects summarised in table 2.3. For example, for a standard (such as six-hourly measurement of respiratory rate) with a baseline

compliance of 70%, the study is powered to detect an SPI associated improvement to 83% compliance, or a deterioration to 55% at p=0.05, two-tailed.

These calculations are appropriate for analysis in binary data where each patient is associated with a single opportunity for error. However, the power available to analyse prescribing errors will tend to be considerably greater than that in table 2.3 since the typical patient is associated with more than one medication order and thus has several opportunities for error. That said, some actions, such as use of blood culture in people with evidence that they may have blood stream infection, were contingent (did not apply to the whole sample) and less power would be available in such cases.

Epoch 1 ran from October 2003 to September 2004 in the SPI1 hospitals and from October 2006 to September 2007 in epoch 2, thereby largely controlling for any seasonal effects (due, for example, to staff changeovers at particular times of the year). As fewer patients were needed for each time period in control hospitals, the epoch only extended from October to March of the corresponding years.

Patients over 65 years of age and admitted with acute respiratory disease, primarily community-acquired pneumonia, exacerbation of chronic obstructive pulmonary disease (COPD) or acute asthma were included in the study (for rationale see case note selection criteria p12). The case notes from the first eight or nine patients who fulfilled the eligibility criteria were selected from each SPI1 hospital, in each month from each epoch. In the control hospitals, the first two or three of such cases were selected.

For each case note, the admission of interest was photocopied and anonymised (with respect to the patient's name, hospital name and year of admission) by medical-record clerks in each hospital. Photocopied notes were dispatched to Birmingham before being sent to reviewers. In Birmingham, anonymisation was quality assured, the notes were digitised, and the year of admission was removed so that reviewers would be blinded to the epoch from which the case notes originated.

We audited the quality of anonymisation by asking the reviewer in the explicit review to note if the hospital of origin, the year of origin and the patient name had been identified.

Box 2.1: Components of an ideal respiratory history

- Duration of presenting symptoms
- Normal (pre-morbid) exercise tolerance
- Presence/absence of shortness of breath
- Presence/absence of orthopnoea
- Presence/absence of cough
- Whether or not cough was productive (if present)
- Smoking history taken
- Presence/absence haemoptysis
- Whether or not chest pain was present
- Occupation/previous occupation
- Pet ownership

Explicit case note review

We developed a set of explicit criteria to define medical care for respiratory patients from the British Thoracic Society (BTS) guidelines,[29;30] the British National Formulary (BNF [versions 53, 54 and 56 – the editions that covered the study period][31–33]) and expert opinion (consultant respiratory physicians from a teaching and a general hospital – see acknowledgements). The areas of review and source of guidelines were:

– Quality of medical history-taking. Eleven items (box 2.1) were identified, using expert opinion, as constituting the ideal history for a patient admitted with acute respiratory disease.
– Proportion of routine investigations (urea and electrolytes, chest x-ray and full blood count) ordered within six hours of a patient's admission (expert opinion – see above).
– Observations and signs of patient deterioration. The completeness with which patients' vital signs were recorded (table 2.4) was evaluated on admission and then for the first and subsequent six hour time periods (BTS). Vital sign data that *were* recorded in the case notes constituted the numerator, while all vital signs that *should* have been recorded constituted the denominator .
– Appropriate clinical response for abnormal vital signs was measured (table 2.5) (BTS).

Table 2.4: Vital signs that should be recorded

	Admission	6 and 12 hours later
Temperature	✓	✓
Respiratory rate	✓	✓
Cyanosis/oxygen saturation	✓	–
Presence of confusion/mental state (new onset)	✓	–
Pulse	✓	✓
Blood pressure	✓	–
Oxygen saturation	–	✓

– Investigating features of good care for specific classes of patients by:
  - calculating the confusion, urea, respiratory rate, blood pressure (CURB) (see box 2.2) score to determine the severity of community acquired pneumonia and hence appropriate antibiotic selection (BTS, BNF)
  - use of intravenous steroids for patients with acute exacerbations of asthma and COPD (BTS)
  - measurement of peak flow in asthma patients (expert opinion)
  - to exclude hypercapnia in COPD patients, by performing arterial blood gases, before prescribing/administering oxygen (BTS).
– Rates of prescribing errors. The following definition was used:

  'A clinically meaningful prescribing error occurs when, as a result of a prescribing decision or prescription writing process, there is an unintentional significant (1) reduction in the probability of treatment being timely and effective or (2) increase in the risk of harm when compared with generally accepted practice'.[34]

Errors were identified using a previously developed pro forma[35] and categorised according to stage of the drug use process (appendix 2).[36] SPI had identified reductions in the number of adverse effects related to anticoagulant therapy as a key aim (see 2.6 sub-study: outcomes p20) so prescribing error in this area was investigated as a sub-category (as listed in section 2.8 of the BNF). Finally, medicines reconciliation on admission was also a target of the SPI (table 1.1). Therefore, we examined failures to continue to prescribe medicines on the transition from primary to secondary care where no explanation for this was recorded in the notes.

Table 2.5: Appropriate clinical response for abnormal observations

Abnormal vital sign	Appropriate clinical response
Oxygen saturation <90, at any time	Full blood gases within 2 hours Given oxygen if not on oxygen Doctor called or transferred to ICU if on oxygen
Blood pressure systolic <90	At least next 6 hours, hourly observations Blood culture
Sputum present	Sputum culture
Respiratory rate >20 at any time after admission	Given oxygen (if not on oxygen) Doctor called (if on oxygen)
Temperature over 38 C – any episode	Blood culture
Failure to improve within 48 hours or subsequent deterioration	Review by consultant Repeat chest x-ray White cell counted/repeated Appropriate addition of further antibiotics

Box 2.2: Assessment of severity of community acquired pneumonia using the CURB score

CURB score

Confusion: new mental confusion (an abbreviated mental test score of eight or less)
Urea: raised >7 mmol/l
Respiratory rate: raised > 30/min
Blood pressure: low blood pressure (systolic blood pressure <90 mm Hg , diastolic blood pressure < 60 mm Hg)

Interpretation of CURB score

- Patients who have two or more core adverse prognostic features are at high risk of death and should be managed as having *s*evere pneumonia
- Patients who display one core adverse prognostic feature are at increased risk of death. The decision to treat such patients as having severe or non-severe pneumonia is a matter of clinical judgement, preferably from an experienced clinician. This decision can be assisted by considering pre-existing and additional adverse prognostic features.

Influence on antibiotic therapy

Non-severe community-acquired pneumonia
Most patients can be adequately treated with oral antibiotics. Combined oral therapy with amoxicillin and a macrolide (erythromycin or clarithromycin) is preferred for patients who require hospital admission for clinical reasons. When oral treatment is contraindicated, recommended parenteral choices include intravenous ampicillin or benzylpenicillin, together with erythromycin or clarithromycin.

Severe community acquired pneumonia
Patients with severe pneumonia should be treated immediately after diagnosis with parenteral antibiotics. An intravenous combination of a broad spectrum b-lactamase stable antibiotic such as co-amoxiclav or a second generation (e.g. cefuroxime) or third generation (e.g. cefotaxime or ceftriaxone) cephalosporin together with a macrolide (e.g. clarithromycin or erythromycin) is preferred.

All case notes were reviewed by a qualified pharmacist (Maisoon Ghaleb) from November 2006 to August 2009. Ideally reviews would be conducted in a random sequence once all records had been collected. This was not possible because of the time taken to collect the case notes and the reporting requirements of the evaluation. Instead, the case notes were scrambled to ensure that the notes were not reviewed entirely in series and in particular so that the same hospitals and epochs were not examined in series. This was done so that any learning or fatigue effects on the part of the reviewer could be detected and could be adjusted for in the analysis.

Inter-observer agreement on prescribing error was evaluated by assigning every 10th case note to two observers (Maisoon Ghaleb and Bryony Dean Franklin) who were both qualified pharmacists, and who assessed cases in batches blinded to each others' assessments, but compared and discussed results after each batch.

Generalised linear mixed models were used to analyse the effect of

the SPI intervention. Fixed effects were included:

a) For differences in pre-intervention levels between control and SPI1 hospitals (baseline comparisons).

b) The temporal change experienced in the control hospitals between the pre-intervention period (epoch 1) and the post-intervention period (epoch 3).

c) The effect of SPI, interpreted as the difference between the temporal changes pre/post intervention experienced in the control and SPI1 hospitals.

Adjustment for the patient-level covariates, age and sex was included in all analyses. Cubic polynomials in the time of review were used to adjust for learning/fatigue effects in the review process and were included in all analyses save that for mortality. Binary observations were modelled using mixed effects logistic regressions with a random component for variation between hospitals.

Medication errors (per recorded prescription) were analysed with population-averaged negative binomial models with grouping by hospital, fitted using generalised estimating equations.

Where the data were insufficient to support a full analysis as described here, the hospital effects were excluded from the model leading to logistic regression analyses (for binary data) and negative binomial regression models (for prescribing errors.)

The calculations were performed in STATA 11.0. Statistical significance is claimed for p-values less than 0.01, and 99% confidence intervals are used throughout. Levels of inter-agreement were tested using the Kappa statistic.

Box 2.3: Definitions of error and adverse events

Error:

Undesirable event in healthcare management which could have lead to harm, or did so, but which did not impact on duration of admission or lead to disability at discharge.

A failure to complete a planned action as it was intended or to adopt an incorrect plan.

Adverse Event

Unintended injury or complication.

Prolonged admission, disability at discharge or death.

Caused by healthcare management rather than the disease process.

Poor outcomes, some of which are the result of preventable actions or poor plans.

Holistic case note review

In addition to the explicit review, each case note was evaluated holistically (implicit review) by a specialist in general medicine (Martin Carmalt). Martin Carmalt has considerable experience in case note review and has investigated hospitals who were outliers on hospital mortality statistics.[37] To measure inter-observer reliability, a subset of case notes (n=91) was independently re-evaluated by an experienced trainee in respiratory medicine (Thirumalai Naicker,). Using expert clinical judgement, an overall quality score was assigned, graded on a scale from one (unsatisfactory, an error had occurred) to 10 (very best care). A specific score for each of three stages of care – admission, management and pre-discharge – was also allocated on a scale
from one (unsatisfactory) to six (excellent care).Reviewers classified errors and adverse events using standard definitions found in box 2.3.[38–42]

The number of errors and adverse events (of all types, not just those relating to medication) were recorded for each patient. It was possible for a patient to have more than one error or adverse event.

The results are presented as average numbers of errors or adverse events per 100 patients. Average ratings and average numbers of adverse events and errors were calculated for both control and intervention groups, for both epoch 1 and epoch 2 (with standard errors).

Table 2.6: Classification of errors and adverse events

Category	Nature of the problem
Diagnosis/assessment admission error	• failure to diagnose promptly/correctly • failure to assess patient's overall condition adequately (including comorbidities)
Hospital-acquired infection	• hospital-acquired infection
Technical/management	• technical problem relating to a procedure • problem in management/monitoring (including nursing and other professional care)
Medication/maintenance/test results	• failure to give correct medicaton or monitor its effects • failure to maintain correct hydration/electrolytes • failure to follow up abnormal test
Clinical reasoning	• obvious failure of clinical reasoning
Discharge information	• information needed by GP not transferred at discharge for whatever reason

Note that a particular error/event could be assigned to more than one category. For example, a test result showing severe hyperthyroidism was ignored and this error could be classified under 'Medication/Maintenance/Test results' and 'Discharge information'.

Adverse events and errors were further classified by broad categories (table 2.6), and adverse events were also categorised into four levels of preventability.

Generalised linear mixed models with random effects for each hospital were used to estimate the difference in changes (although in one instance the random effect model did not converge so a fixed model was used instead). No formal adjustments were made for multiple comparisons, although 99% confidence intervals are quoted in all cases. Inter-observer reliability was assessed by ICC for the four scores.

For the adverse events and errors, inter-observer reliability was assessed comparing errors and adverse events identified by both reviewers, using the Kappa statistic.

2.6 Sub-study 5: Outcomes

Adverse events

The SPI aimed to make a 50% reduction[43;44] in the total number of adverse events. The incidence of patient harm caused by medication was measured as part of the explicit review.

The holistic review also measured adverse events both overall and by degree of preventability (see holistic case note review p20); and results are given as total adverse events per 100 patients.

Mortality

It was not possible to measure overall adjusted (or crude) hospital mortality rates for all four SPI1 hospitals because of difficulty accessing this information in the non-English countries. Moreover, mortality is an insensitive marker of quality[45] and a change in this outcome might not be expected in a study with only four intervention hospitals.[46]

We did compare mortality rates across epochs among patients whose case notes were selected for review. This was because it was feasible and, arguably, a higher signal to noise ratio would be expected among this group, which not only benefits from specific SPI interventions but also has high mortality.

Patient satisfaction

Since quality of care and avoidance of adverse events are important to patients, improvements in practice might plausibly affect patients' views of their care. Their views were assessed by means of a survey.

All English hospitals participate in an annual patient survey. In our study, the same questionnaire was also used in the non-English hospitals, using the same methods as those used in the Care Quality Commission's National NHS Acute Inpatient Survey in England. The detail of this methodology is available from www.nhssurveys.com

The dates of the surveys were aligned with those of the staff surveys, and the same control hospitals were selected. Methods similar to those for the staff survey were used in the analysis, except that only organisational level data were available for control hospitals. An organisational level analysis was conducted using a two-way analysis of variance (ANOVA) (the factors being SPI1 versus control hospital, and survey one versus survey two).

Five scores (table 2.7) were identified for analysis: three overall satisfaction scores and two related to cleanliness. The details of these scores can be found in appendix 3. Organisation-level scores in each arm of the study were formed by averaging all respondents' scores within each hospital.

Table 2.7: Patient survey questions deemed relevant to SPI

1. Overall, how would you rate the care you received?
2. How would you rate how well the doctors and nurses worked together?
3. Overall, did you feel you were treated with respect and dignity while you were in the hospital?
4. In your opinion, how clean was the hospital room or ward that you were in?
5. How clean were the toilets and bathrooms that you used in hospital?

Chapter 3

Results

3.1 Sub-study 1: Strategic stakeholder interviews

Strategic stakeholders generally saw the aims of the SPI as legitimate and sound. They accepted that there was a need to control risk and that patient safety was an important priority for hospitals. Only seven of the 60 hospital stakeholders were unable to describe the SPI accurately or in detail. The majority of accounts from hospital stakeholders appeared to demonstrate a shared understanding of the main elements of the programme. Most, for example, gave a reasonable account of the PDSA cycle.

There was considerable enthusiasm for the programme. However, there were also concerns about the ambitious reach of the programme, whether resources would be equal to the demands, and whether resistance might be encountered at the sharp end. Further details are reported elsewhere.[9]

The quantitative analysis corroborates the qualitative analysis: 73% scored above five on the knowledge scale and 83% scored above five on the enthusiasm scale (figure 3.1).

The correlation between knowledge and enthusiasm varied depending on the rater (the ICC's between enthusiasm and knowledge for the three raters were 0.61, 0.69 and 0.91.

The correlation between raters was medium to high (the ICC's between the three pairs of raters for knowledge were 0.54, 0.56 and 0.63; and 0.58, 0.70 and 0.71 for enthusiasm).

3.2 Sub-study 2: Staff surveys

For the first survey, the mean response rate was 45% (7,826 of 17,507 returned) in the four SPI1 hospitals. For the second, it was just 35% (4,191/11,922) across the three participating hospitals (see table 3.2). However, there was a national decrease in response rates among acute hospitals in England over the same period (from 57% to 52%).[47;48]

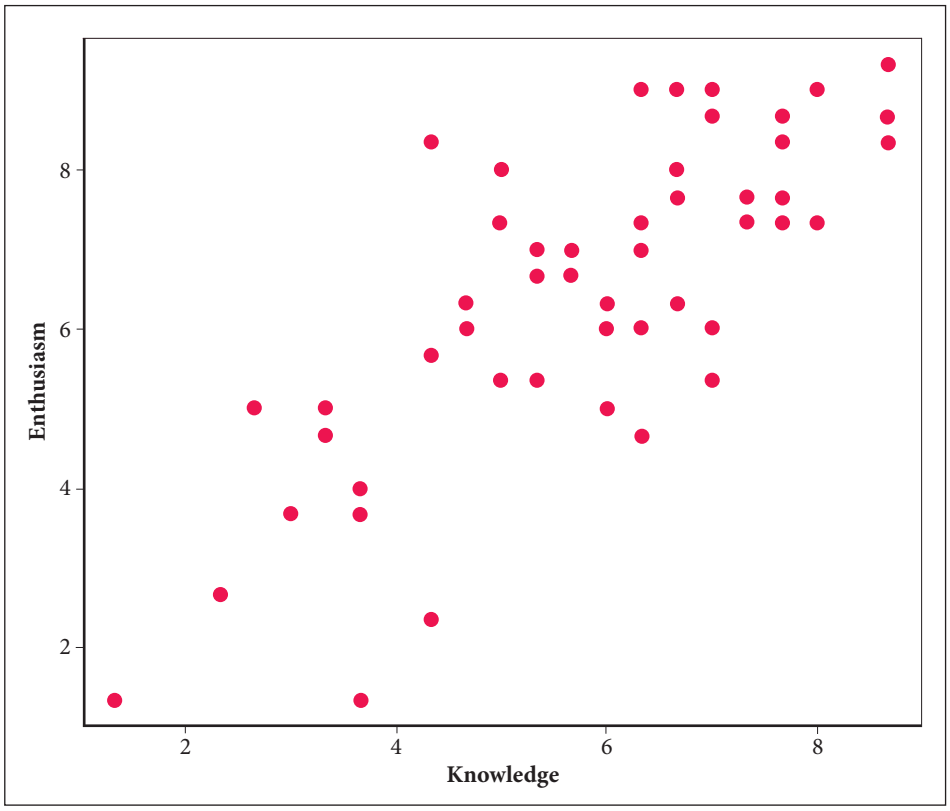


Figure 3.1: Correlation between knowledge and enthusiasm for SPI among strategic hospital stakeholders (some dots represent more than one interviewee)

In the 18 control hospitals, the response rate for the first survey varied from 38% to 71%, with an average of 57%, and for the second survey varied from 26% to 62% with an average of 52%. These figures are similar to contemporaneous national response rates. Response rates were lower in the non-English hospitals, possibly due to the lack of profile for the survey outside England.

There was no significant difference in response rates between control and SPI1 hospitals at baseline. Table 3.1 shows the values of the 11 survey scores in each of the four SPI1 hospitals for the two surveys, along with details of response rates. Table 3.2 shows the changes in both control and SPI1 hospitals on each of the 11 scores identified, along with the differences between the groups in these changes and associated 99% confidence intervals.

Comparison with control hospitals is important because national changes in the NHS over this period resulted in generally more negative scores at the second survey than at the first.[50] At baseline, the percentage of staff reporting well-structured appraisals within the previous 12 months was significantly lower in SPI1 hospitals than control hospitals. 'Job satisfaction' and 'support from supervisors' were also significantly lower in SPI1 hospitals than control hospitals ($p<0.01$). None of the other baseline differences was statistically significant.

Table 3.1: Staff survey scores in SPI1 hospitals at the two periods*

	Hospital 1				Hospital 2				Hospital 3				Hospital 4**	
	N	Survey 1 score (SE)	N	Survey 2 score (SE)	N	Survey 1 score (SE)	N	Survey 2 score (SE)	N	Survey 1 score (SE)	N	Survey 2 score (SE)	N	Survey score (SE)
% staff having well structured appraisals within previous 12 months	1094	41 (1)	1079	33 (1)	2458	39 (1)	1045	28 (1)	2101	27 (1)	1869	23 (1)	458	28 (2)
% staff working in well structured teams	1091	37 (1)	1073	38 (1)	2450	41 (1)	1043	37 (1)	2090	34 (1)	1845	33 (1)	457	42 (2)
% staff witnessing potentially harmful errors or near misses in previous month	1119	54 (1)	1122	45 (1)	2517	44 (1)	1074	47 (2)	2138	50 (1)	1937	42 (1)	468	46 (2)
% staff suffering work-related injury in previous 12 months	1129	22 (1)	1082	19 (1)	2514	20 (1)	1045	20 (1)	2139	22 (1)	1896	17 (1)	464	19 (2)
% staff suffering work-related stress in previous 12 months	1147	36 (1)	1104	32 (1)	2541	32 (1)	1072	30 (1)	2178	35 (1)	1918	29 (1)	474	39 (2)
% staff experiencing physical violence from patients/relatives in previous 12 months	1139	14 (1)	1112	9 (1)	2531	13 (1)	1073	11 (1)	2157	17 (1)	1932	13 (1)	463	18 (2)
Intention to leave[49]	1144	3.41 (0.03)	1118	3.39 (0.03)	2508	3.50 (0.02)	1074	3.33 (0.03)	2158	2.99 (0.02)	1919	3.03 (0.02)	472	3.33 (0.05)
Staff job satisfaction[49]	1155	3.46 (0.02)	1120	3.46 (0.02)	2550	3.56 (0.01)	1079	3.35 (0.02)	2185	3.25 (0.02)	1923	3.26 (0.02)	476	3.40 (0.03)
Quality of work-life balance[49]	1152	2.64 (0.03)	1116	2.60 (0.03)	2529	2.50 (0.02)	1074	2.75 (0.03)	2173	2.73 (0.02)	1916	2.66 (0.02)	476	2.63 (0.05)
Support from supervisors[49]	1146	3.40 (0.03)	1120	3.48 (0.03)	2536	3.50 (0.02)	1072	3.36 (0.03)	2179	3.18 (0.02)	1917	3.21 (0.02)	473	3.33 (0.04)
Organisational climate[49]	1137	3.32 (0.02)	1098	3.20 (0.02)	2516	3.19 (0.01)	1070	2.90 (0.02)	2155	2.82 (0.02)	1902	2.79 (0.02)	459	3.08 (0.03)
Response rate	–	43%	–	40%	–	39%	–	39%	–	39%	–	31%	–	48%

* The first six of these scores were percentages, simply reflecting the percentage of respondents who answered 'yes' to a single question or a set of questions. The other five are on a scale of one to five, and are based on the mean of between three and six questions, each of which was scored between one and five for each respondent. For four of these five scores (quality of work-life balance, staff job satisfaction, support from supervisors and organisational climate), the higher the score the better, although for intention to leave, lower scores are better.

** Due to reorganisation this hospital only participated in one survey.

Standard Error (SE)

Table 3.2: Staff survey scores in control and SPI1 hospitals at the two periods*

	Control hospitals					SPI1 hospitals					Range at baseline	Difference in change (99% CI)	p
	N	Survey 1 score (SE)	N	Survey 2 score (SE)	Absolute % change	N	Survey 1 score (SE)	N	Survey 2 score (SE)	Absolute % change			
% staff having well structured appraisals within previous 12 months	8046	39 (1)	7260	28 (1)	-10	6111	34 (1)	3993	27 (1)	-7	27–46	3 (-2, 8)	0.095
% staff working in well structured teams	8052	40 (1)	7279	37 (1)	-3	6088	38 (1)	3961	35 (1)	-2	34–52	1 (-4, 6)	0.510
% staff witnessing potentially harmful errors or near misses in previous month	8236	47 (1)	7520	39 (1)	-8	6242	48 (1)	4133	44 (1)	-4	41–56	5 (-2, 11)	0.068
% staff suffering work related injury in previous 12 months	8286	22 (0)	7372	19 (0)	-3	6246	21 (1)	4023	18 (1)	-3	18–26	0 (-4, 3)	0.854
% staff suffering work related stress in previous 12 months	8368	34 (1)	7457	33 (1)	-1	6340	34 (1)	4094	30 (1)	-4	29–39	-5 (-12, 0)	0.013
% staff experiencing physical violence from patients/ relatives in previous 12 months	8283	13 (0)	7482	11 (0)	-2	6290	15 (0)	4117	11 (0)	-3	9–18	-3 (-8, 0)	0.026
Intention to leave[49]	8263	3.36 (0.01)	7437	3.29 (0.01)	-0.08	6282	3.29 (0.01)	4111	3.21 (0.01)	-0.09	2.99–3.51	0.04 *(-0.03, 0.10)*	0. 139
Staff job satisfaction[49]	8357	3.47 (0.01)	7495	3.37 (0.01)	-0.10	6366	3.42 (0.01)	4122	3.34 (0.01)	-0.09	3.25–3.57	0.03 (-0.02, 0.08)	0.132
Quality of work-life balance[49]	8249	2.63 (0.01)	7436	2.72 (0.01)	0.10	6330	2.61 (0.01)	4106	2.67 (0.02)	0.06	2.45–2.80	-0.05 (-0.12, 0.03)	0.106
Support from supervisors[49]	8310	3.45 (0.01)	7477	3.41 (0.01)	-0.05	6334	3.36 (0.01)	4109	3.32 (0.01)	-0.04	3.18–3.55	0.02 (-0.04, 0.08)	0.358
Organisational climate[49]	8302	3.11 (0.01)	7424	2.89 (0.01)	-0.22	6267	3.08 (0.01)	4070	2.93 (0.01)	-0.15	2.82–3.32	0.08 (0.02, 0.13)	0.000*

* The first six of these scores were percentages, simply reflecting the percentage of respondents who answered 'yes' to a single question or a set of questions. The other five are on a scale of one to five, and are based on the mean of between three and six questions, each of which was scored between one and five for each respondent. For four of these five scores (quality of work-life balance, staff job satisfaction, support from supervisors and organisational climate), the higher the score the better, although for intention to leave, lower scores are better. To aid interpretation scores where a lower value is better are shown in *italics*. Range at baseline indicates the range of scores across control and SPI1 hospitals in the first survey to give some context for the level of change shown. The difference in change and corresponding confidence interval does not necessarily reflect the difference in absolute change because of the inclusion of covariates in the models tested.

Only one of the 11 scores shows a statistically significant difference ($p<0.01$) in changes between the control hospitals and SPI1 hospitals between the two surveys. Organisational climate, which refers to the extent of positive feeling within the organisation about communication, staff involvement, innovation and patient care was similar between the control and SPI1 hospitals at baseline (3.08 versus 3.11 on a scale where one is very negative and five is very positive). This score decreased by 0.22 in the control hospitals but only by 0.15 in the SPI1 hospitals ($p=0$). The effect size for this difference in change between the control and SPI1 hospitals after covariates are taken into account was modest, at 0.08 points on a five point scale where there was a range at baseline of 0.5 points between hospitals.

3.3 Sub-study 3: Ethnographic study

The first two visits provided insights into the sharp end of practice on wards, while the focus groups undertaken on the third visit also provided insights into the other layers of management and strategy in the hospitals. We were able to identify staff views on the SPI as they experienced, and what they thought would help in the spread of safer practices for the future.

Reflecting the findings of the management interviews, the focus groups across the four sites agreed that the senior people in the hospitals were committed and enthusiastic about SPI, made a significant strategic contribution, gave weight to the programme and generally set a good example for staff.

> *'If these guys [senior leaders] aren't behind it, very quickly your clinical directors and …other directors… and senior people start to fall by the wayside. So I think that's absolutely paramount, having the top guys leading the way, so I think that has been one of the big successes.'*
>
> (Focus group)

The SPI was seen as having encouraged a change from patient safety being something taken up by individual (but sometimes unaccountable) voluntary enthusiasts, towards being a more mainstream priority. The involvement of the IHI was seen as crucial in lending credibility and support to the progamme, and was valued as a source of knowledge and expertise.

> *'It's fundamentally important for people who teach you [to] have credibility, and I think IHI in the United States know their stuff and they have their way of teaching things. It's culturally a bit difficult to get into it but once you know you've got it, you are there, it's really empowering.'*
>
> (Focus group)

Despite the enthusiasm and support at a strategic level, the management layer (often ward sister or consultant) appeared less engaged in SPI than the strategic level.

> *'I think it starts from the top but I don't know if it actually gets right down and the same from the bottom up, I think we've got that middle layer that often it gets lost in sometimes.'*
>
> (Focus group)

Middle managers and front-line staff were not uninterested in, or unconcerned about, patient safety. In interviews, these staff gave detailed accounts of risks they confronted on wards, and they expressed anxiety that some of these were not managed well. The risks they described were often those being targeted by either the generic or the specific interventions of the SPI, such as, for example, communication and handovers.

> *'I think the biggest problem we have on this ward, and I think that would be anywhere where you have a big establishment with a lot of people – you've got the doctors, the physios, the OTs – it's communication. The doctors will quite often come and do their ward round and they'll go around say, "Oh Mr so and so you can go home today", but they won't tell [...] anyone else.'*
>
> (Interview with ward staff member)

Some of problems of engaging middle-managers and ward staff can be explained by the day-to-day nature of their work, which is often focused on managing complex clinical and organisational demands.

Observations suggested that the wards were often very busy and stressful places to work. Staff interviews pointed to problems of managing with limited resources, especially inadequate staffing or problems of skill mix and shortages of equipment.[51]

This meant that staff felt they were often too stretched to give priority to things other than the tasks in hand. There were suggestions that the effort to improve patient safety should focus on improving aspects of structure rather than on processes.

> *'What I want is trained nurses on the ward so I can manage the ward better, discharge many patients sooner and reduce the number of complaint letters...'*
>
> (Interview with ward staff member)

It was clear that staff were tired of implementing lots of different initiatives, and middle managers found it tough to balance a raft of competing priorities.

'There's so many things now that the nurses in the wards have to do... the same time we're bringing in the Safer Patients Initiative I'll be bringing in health and safety risk management training, I'll be bringing in fire safety training… and they all go through the senior nurse on the ward…

'So you've got things being dropped from a great height down onto the nurses so you can see how in some ways we're creating part of the confusion because we all think it's important they know that and they're doing it but if it's 20 or 25 different issues for a nurse then it's a heck of a lot to take on board.'

(Focus group)

Where ward staff did know what was going on in relation to the SPI, they were generally positive.

'The safety briefings have improved the way we work from a point of view that we have... a better system of monitoring our equipment, ... we get it fixed quicker because it's been highlighted in the morning as part of a safety brief, we come out of handover and we're all aware of particular wandering patients. [...] If there was some medication error, we're using the safety briefings to maybe say, not naming names but, you know a patient isn't receiving this, can we please make sure they do because they were complaining overnight about it. So I think it's improving the communication between the nurses.'

(Interview with ward staff)

However, the ethnographic work suggested that the impacts of the SPI at the level of medical wards were mostly difficult to detect. Apart from improved monitoring of patients using EWSS, the SPI was not routinely evident in the everyday practices of people caring for patients on the front line.

For the most part, the sharp-end staff tended either to know relatively little about SPI procedures, practices and principles, or they viewed them as handed down rather than something that they had been involved in developing.

Outside a small number of pockets of activity, there was little evidence that front-line staff perceived a sense of ownership over the initiative. There was also a perception, in interviews and the focus groups, that the SPI had allowed a small number of people to become an elite group with enhanced career prospects who then moved on, while others were excluded.

'SPI was a select group of 20 people. I think we could only bring down 20 people and you're starting off in small areas and of course the by-product of that is that you've got a small group dealing with those small areas so there is, although we may not like it, a perception in some parts of the organisation that SPI is a an elite entity.'

(Focus group)

The gap between the strategic level view and what was happening at the sharp end was evident in a number of different ways. For example, leadership walk rounds (where a senior executive visited the wards) were discussed enthusiastically in the leadership focus groups, and were seen as a highly effective way of understanding the issues that the sharp end finds difficult. They raised senior managers' understanding and awareness of life at the sharp end.

> *'If I wanted to find out what was going on [on] a ward I could do a sort of incident reporting system but I also know that that won't tell me what life is really like on that ward. So actually going on the ward and listening to staff talk specifically about harm to patients is something that I don't believe that most executives get in their normal practice. We all get trapped in offices.'*
>
> (Focus group)

However, walk rounds were only seldom mentioned by ward staff in interviews (and they were not witnessed over the 300 hours of observations – though this may simply have been an artefact of the data collection process). When discussed in one focus group, it was evident that staff at the sharp end felt that the process was disappointing and may even have undermined the SPI, because it appeared to demonstrate a failure to connect senior management with the wards in meaningful ways. Here is an anecdote from a focus group discussion:

> *Have you had any leadership walk rounds on your ward?*
>
> *A: 'Yes we had one of the guys came down with [name] not that long ago, about a month ago, he came down for a walk round.'*
>
> *What was it like?*
>
> *A: 'Well he came around and spoke to a few people and just asked about any concerns. He said he was interested to know how the nursing staff felt and he wanted to know one thing that he could take back to the rest of the board about any issues that nursing staff had. Afterwards they sent a letter to say thanks but you never hear any … well we haven't heard anything more than that so.'*
>
> *How long time ago was that?*
>
> *B: 'I think it was about [two months ago].'*
>
> *Was there an item they then took back?*
>
> *A: 'Well we had said that we were concerned about working short-staffed so often and also about the lack of opportunities for staff to do ongoing study.'*
>
> *B: 'Yes about the lack of equipment.'*

A: 'The lack of equipment and … well we said a few things and I think it was quite general across all the wards because they went along the whole floor on different days and visited all the wards along the medical floor. Everybody was sort of mentioning the same things. But we've never heard anything about change because of it, but we did have our little moan.'

Similarly, the hospital stakeholders and focus group participants agreed that there were great benefits of the PDSA approach: it developed expertise, enabled the hospitals to try out new ways of working, allowed staff to experiment, gave space and privacy for correcting mistakes, and allowed local customisation.

'It gives you permission to try new things and if it doesn't work it doesn't work. You haven't broken any rules because in hospitals we are very much bound by, is this the accepted way, is this allowed. But PDSA has given us permission to try different things even for a day, a shift, you know make changes.'

(Focus group)

Several PDSA success stories were reported in the focus groups. However, few frontline ward staff who were interviewed seemed aware of PDSA. Thus, somewhere between the blunt end and the sharp end, the model of participative engagement on which the SPI was based got lost.

There were several important influences on the extent to which SPI interventions became embedded on the wards. One was legitimacy. Sometimes staff simply did not see particular interventions as being scientifically legitimate.

'Something that appears very simple on the surface like the definition of a surgical infection caused an absolute riot.'

(Focus group)

Scientific legitimacy problems were, perhaps paradoxically, compounded by the use of PDSA cycles. Some clinical staff were reported to see the data collected during the cycles as unreliable and lacking in credibility, and providing a case for change. However, it may also be that claims of poor evidence were being used to resist change and reinforce inertia. Legitimacy problems also arose when staff did not recognise that the problem being tackled was a real one, or they did not feel that the resources required to implement the intervention was a legitimate use of staff time. There were also suggestions that legitimacy varied among different staff groups and junior doctors were often seen (by nurses in particular) as especially difficult to engage.

'I think it's very variable, I think there are some senior clinicians that are very supportive... but we still don't seem to be solving the problems with the junior doctors.'

(Focus group)

The second important influence related to how trackable any improvements were. One of the reasons why the EWSS seemed to penetrate practice was that they left a visible trace (in the form of a record of observations of vital signs), and thus promoted a sense of accountability. It was also clearly linked to a long nursing tradition of conducting observations of patients' vital signs, and was seen as addressing the legitimate and important problem of identifying and responding to patient deterioration. In contrast, ethnographic observations suggested that safety briefings were rarely used in a recognisable form on the wards, even though staff saw handovers as a risky area. Similarly, the situation, background, assessment, recommendation (SBAR) communication technique seemed to have a relatively low uptake. In the focus groups, the apparent failure of the safety briefings was attributed partly to the fact that it was difficult to demonstrate the improved practice and hence there was little incentive to comply.

'I think one of the reasons why it [safety briefings] wasn't complied to as part of SPI [was because] it wasn't something we were asked to report on.'

(Focus group)

'One of the issues for me with SBAR is that we've been so focused on measurement [but] it's one of those things that's really difficult to measure.'

(Focus group)

Observations identified several further barriers to adopting safety initiatives, including the instability of teams caused by rotating staff and frequent substitutions by agency staff. This meant that it was difficult to sustain a collective knowledge and faith in the SPI over time. For example, in a discussion of hand-washing, ward staff reported:

A: 'Nursing-wise it's quite consistent but in medical staff, the doctors.'

B: 'We're having still poor response, poor compliance.'

A: 'I think we've had a lot of bank and agency staff working and they sometimes don't seem to be as hot on it as the regular staff.'

(Focus group)

It is important to emphasise that the SPI did have positive impacts, which were clearly evidenced in some of the interviews and the focus groups. In particular, it helped to increase managerial recognition and focus on patient safety, and promoted a systematic approach to tackling patient safety problems. One of its more lasting benefits was that over time, hospitals began to recognise challenges more clearly.

> *'I mean with the [recently commenced patient safety strategy/ initiative] we've got our tasks set out for us but I suppose most of the participants we, we've got a bit of history on this, we have progressed quite substantially, okay there's a duty on us to ever do better so in terms of those interventions… … we actually have the opportunity to innovate, we've got the methodology and we should be looking to spread, and so we are looking to spread that methodology, that's the challenge for us to go forward.'*
>
> (Focus group)

A key achievement of the SPI was encouraging organisational learning about how to manage quality improvement efforts in the future.

> *'I think we have to be careful and I think we have to recognise that yes, we've done some great work but only in small areas and now this is a huge thing for us to spread this across the organisation and admit that some of the things that we did in SPI we didn't maybe do as well as we could have done.'*
>
> (Focus group)

Hospitals reported that they had begun to devise and implement strategies for future implementation of patient safety programmes. One of the major lessons learned was the scale of resource and organisational support required to make patient safety efforts work. There was a perception that hospitals had underestimated this, and that they had been, in the early stages, too ambitious and too ready to assume that something that worked in a defined clinical area (such as ICUs) would easily transfer to other environments.

> *'I wonder if we tried to do too much. [...] I know the American experience was they could choose only six and now with the latest experience they can choose from 12. I just wonder if we tried too much and weren't able to devote enough focus on every element of it.*
>
> *'I feel that we've not made the same progress… well clearly we haven't made the same progress on all 29 initiatives in all five work streams and I wonder if we spread ourselves too thinly at times.'*
>
> (Focus group)

For the future, engaging senior clinicians and encouraging local ownership was widely seen as the key to success.

> *'It's getting the leadership there you know, the clinical lead, to sign up to it and to really drive it because we don't have that leadership. I think we're at an advantage now because we have learned a lot from what we have done over the past three years. Now when we do move to spread throughout the organisation we have all that learning behind us and we're able to reflect on that and take it forward perhaps in a slightly different way and learn to engage people maybe more at the start which is something that we didn't do three years ago.*
>
> *[Name] is now going to the monthly meetings between the medical director and the clinical director and also the clinical managers so that we're starting to get that information to them and starting to ask to get ownership from them about what change needs to happen because that middle layer wasn't really working.'*
>
> (Focus group)

Strategies for the future included using reputational incentives to encourage people to cooperate:

> *'I don't think we've got to a position where the peer pressure's breaking through.'*
>
> (Focus group)

Avoiding paperwork associated with patient safety was also seen as important in securing the cooperation of front-line staff. One focus group discussed how important it is to develop more meaningful ways to measure and prompt compliance without overloading staff with audits and data collection. In their organisation, each clinical area can decide how they implement and measure safety briefings for themselves.

> *'You can have tick box and say "yes a safety briefing has taken place", but how effective is that? One of the ways that we've advised the staff to think about is if you asked a member of staff later on that day "what were the three things on the safety briefing?" they should be able to tell you.*
>
> *'It's making it doable, measurable but not more data, not more audits, it's how you capture that.'*
>
> (Focus group)

Creating new structures to support patient safety work was also seen as important in some hospitals.

'By having the small groups it was actually preventing us from spread. Each clinical group now has responsibility and each hospital has responsibility for patient safety within whichever structure they choose,whether it's clinical governance or health and safety they all have a small patient safety group within each clinical group. What we found is if you were in the general ward work stream you just focused on the general ward work but some of the critical care stuff actually applied to you and vice versa.

One of the first things we did actually was not ban the things we learned from SPI but stop calling things SPI committees… we didn't need to use it any more so we just literally changed it to Patient Safety Committee as opposed to an SPI committee but taking forward all the stuff that was learned of course, part of the culture change as well and involving everybody.

We are going to stop having a Safer Patients Initiative steering group and we're going to have a patient safety committee safety representative for each directorate who will be responsible with the clinical director and general manager for that directorate for delivering on all the workstreams, i.e. they will monitor them within their directorate and at the safety committee will just monitor certain high level...'

(Focus groups)

Organisational changes, critically, also meant embedding the work within those less engaged 'middle layers' of the hospitals. Thus, some hospitals gave departments/divisions more responsibility for implementing and monitoring patient safety.

'We will leave it to individual departments to monitor them and make correction action where necessary and we'll just monitor the outcomes, so it's going to be a different way of.. well we are hoping it will embed it by making it everyone's responsibility but at the same time and in order to introduce new work streams because we're introducing training and a few other things as well as… … and there will be improvement team support for the directorates and the safety leads in the directorates, so it's going to be a different way of looking at it in … with the aim of trying to embed it more.'

(Focus group)

3.4 Sub-study 4: Error rates/quality of care among acute medical patients

Explicit case note review

The sample

The smallest SPI1 hospital in our sample could not identify the target numbers of case notes, leading to a slight shortfall in the intended SPI1 sample size of 400 case notes in each epoch: 381 (epoch 1, before SPI1) and 380 (epoch 2, after implementation of SPI1). The corresponding numbers for control hospitals are 236 case notes in epoch 1 and 240 in epoch 2.

Effects associated with the review process

Case note reviews took place from November 2006 to August 2009. The review of SPI1 hospital case notes was done first, and was 90% complete by August 2008. By contrast, 90% of the control hospital reviews were not carried out until after September 2008 (as this was commissioned at a later date). In the intervention arm, randomisation of the order of review was only partial: on average, the review date for epoch 1 was earlier than for epoch 2 – there was considerable overlap between the dates of the two sets of reviews (see figure 3.2).

A cubic polynomial adjustment for the timing of the review was employed to minimise the potential for confounding between the SPI effect and any temporal effects associated with the date of review. These effects were found to be most significant in the analysis of prescribing errors and have been routinely incorporated into the results described in prescribing error (p41) and medicines reconciliation errors at admission (p47). Elsewhere they were found to be significant for just three items not associated with prescribing error (see observations and signs of patient deterioration (p37) and appropriate clinical response to abnormal vital signs (p41)).

Reliability and anonymisation of case notes

The comparison of prescribing error results between two observers showed substantial[52] inter-observer agreement with a Kappa value of 0.71 and 0.70 in epochs 1 and 2, respectively. Prescribing errors were used to assess reliability as it is the most difficult of the explicit review criteria to assess, being based on hundreds of potential errors in the BNF.

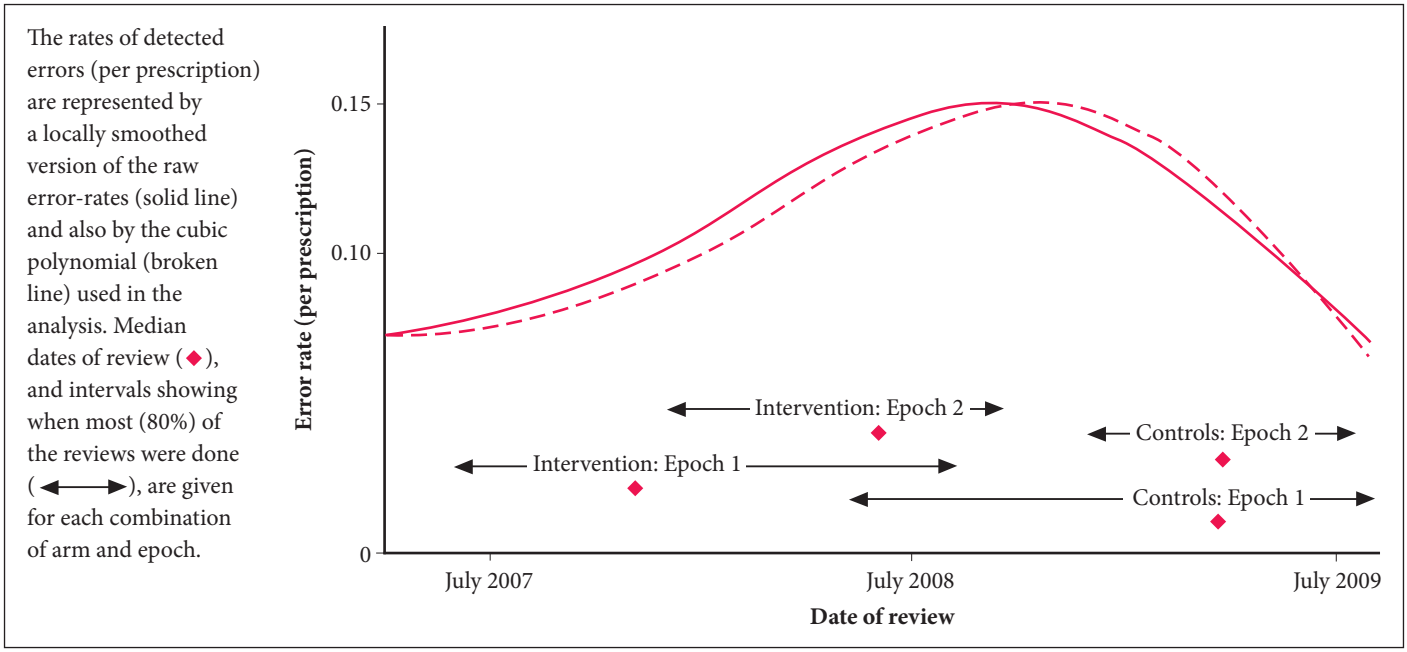


Figure 3.2: Prescription error rates and date of review

During the review the primary reviewer was able to detect the hospital of origin in 1% of cases (11/1,154), the epoch in 14% of cases (158/1,154) and the patients' name in 4% of cases (42/1,154).

Quality of medical history taking

The effect of SPI is not apparent and is not statistically significant for any item. The baseline comparisons showed no significant differences between control and SPI1 hospitals, nor is there significant evidence of temporal improvement for any item (see table 3.3b).

There was some evidence of a temporal effect in the review process (learning/fatigue effect) for item two (exercise tolerance), $p<0.001$ and for item 9 (chest pain), $p=0.002$.

Several of the questions were asked less often for older patients. Age was a significant predictor for items two, three, five, six, seven and nine ($p<0.001$ in all cases), typically reducing the odds of the question being asked by about 5% per year of age.

Observations and signs of patient deterioration

The effect of SPI is not apparent, and is not statistically significant for any item. The baseline comparisons showed no significant differences between control and SPI1 hospitals. However, compliance appears to have improved both at six hours and 12 hours, irrespective of the SPI. In fact there is significant ($p<0.01$) evidence of improvement in the control hospitals for items eight, 12, 13 and 14 (see table 3.4b).

Table 3.3a: Medical history taking (% of patients asked)

	Control hospitals				SPI1 hospitals			
	Epoch 1		Epoch 2		Epoch 1		Epoch 2	
No. of patients	**236**		**240**		**381**		**380**	
	%	**SE**	**%**	**SE**	**%**	**SE**	**%**	**SE**
1. Duration of presenting symptom	94.5	1.5	94.6	1.5	94.5	1.2	95.5	1.1
2. Normal exercise tolerance	32.6	3.1	34.9	3.1	32.5	2.4	37.1	2.5
3. Presence/absence shortness of breath	89.8	2.0	92.1	1.7	93.2	1.3	93.9	1.2
4. Presence/absence orthopnoea	28.0	2.9	28.7	2.9	24.1	2.2	20.3	2.1
5. Presence/absence cough	89.8	2.0	90.4	1.9	84.8	1.8	89.2	1.6
6. If cough, was it productive	82.6	2.5	86.3	2.2	81.6	2.0	88.2	1.7
7. Smoking history taken	75.7	2.8	80.4	2.6	80.3	2.0	82.1	2.0
8. Presence/absence of haemoptysis	23.7	2.8	25.7	2.8	26.0	2.2	27.4	2.3
9. Chest pain (of any type)	61.3	3.2	68.6	3.0	74.8	2.2	71.8	2.3
10. Occupation/previous occupation	39.7	3.2	38.1	3.1	63.5	2.5	63.9	2.5
11. Pets	2.6	1.0	3.0	1.1	1.8	0.7	1.1	0.5
% over all items	56.6		58.6		59.8		61.0	

Note that entries are percentages with binomial standard errors (SE).

Table 3.3b: Medical history taking – differences between control and intervention hospitals, changes over time and the effect of SPI

	Baseline comparisons		Changes in controls		Effect of SPI	
	OR (99% CI)	**p**	**OR (99% CI)**	**p**	**OR (99% CI)**	**p**
1. Duration of presenting symptom	2.0 (0.5, 8.5)	0.207	0.8 (0.3, 2.4)	0.607	1.6 (0.4, 6.9)	0.414
2. Normal exercise tolerance†	0.6 (0.2, 1.5)	0.158	1.2 (0.7, 2.0)	0.421	1.5 (0.7, 3.0)	0.178
3. Presence/absence shortness of breath	2.1 (0.6, 7.7)	0.149	1.3 (0.6, 3.2)	0.388	0.9 (0.3, 3.1)	0.843
4. Presence/absence orthopnoea	0.7 (0.3, 1.5)	0.230	1.0 (0.6, 1.8)	0.817	1.0 (0.5, 2.2)	0.966
5. Presence/absence cough	0.4 (0.1, 1.1)	0.018	1.2 (0.5, 2.7)	0.610	1.9 (0.7, 5.3)	0.129
6. If cough, was it productive	0.8 (0.3, 2.2)	0.533	1.5 (0.7, 3.0)	0.142	1.3 (0.5, 3.3)	0.453
7. Smoking history taken†	1.0 (0.3, 3.1)	0.963	1.6 (0.8, 2.9)	0.060	0.8 (0.4, 1.9)	0.519
8. Presence/absence of haemoptysis	1.1 (0.5, 2.4)	0.733	1.2 (0.7, 2.0)	0.505	0.9 (0.4, 1.9)	0.769
9. Chest pain (of any type)†	1.1 (0.4, 2.6)	0.872	1.5 (0.9, 2.6)	0.031	0.7 (0.3, 1.5)	0.230
10. Occupation/previous occupation†	1.6 (0.7, 3.7)	0.159	1.0 (0.6, 1.6)	0.939	1.1 (0.6, 2.2)	0.622
11. Pets	0.3 (0.03, 1.6)	0.048	1.5 (0.3, 6.9)	0.502	0.6 (0.1, 6.1)	0.571

† Denotes items with significant ($P < 0.010$) between hospital variation within the arms of the study.

Table 3.4a: Vital signs – percentage compliance with standards

	Control hospitals				SPI1 hospitals			
	Epoch 1		Epoch 2		Epoch 1		Epoch 2	
	%	SE	%	SE	%	SE	%	SE
On admission:								
1. Temperature	96.7	1.6	99.2	0.8	99.0	0.4	99.2	0.4
2. Respiratory rate	95.8	1.8	99.2	0.8	92.1	1.2	98.4	0.6
3. Cyanosis/O2 sat	98.3	1.2	98.4	1.1	98.0	0.6	99.2	0.4
4. Confusion/Mental state	53.3	4.6	71.5	4.1	65.7	2.1	66.2	2.1
5. Pulse	98.3	1.2	99.2	0.8	99.0	0.4	99.4	0.3
6. Blood pressure	98.3	1.2	99.2	0.8	99.0	0.4	99.4	0.3
At 6 hours:								
7. Temperature	61.7	4.5	69.9	4.2	72.9	2.0	84.1	1.6
8. Respiratory rate	40.8	4.5	69.1	4.2	43.8	2.2	80.3	1.8
9. Pulse	69.2	4.2	73.2	4.0	78.6	1.8	86.5	1.5
10. O2 saturation	61.7	4.5	71.5	4.1	72.9	2.0	85.5	1.6
At 12 hours:								
11. Temperature	58.3	4.5	70.7	4.1	67.9	2.1	79.4	1.8
12. Respiratory rate	35.0	4.4	69.9	4.2	38.8	2.2	75.5	1.9
13. Pulse	63.3	4.4	76.4	3.8	71.7	2.0	80.3	1.8
14. O2 saturation	54.2	4.6	75.6	3.9	62.4	2.2	79.3	1.8
Routine investigations:								
15. U & E	99.2	0.8	98.4	1.1	99.0	0.4	98.8	0.5
16. Chest X-ray	96.7	1.6	97.6	1.4	94.9	1.0	94.8	1.0
17. Full Blood Count	98.3	1.2	97.6	1.4	99.0	0.4	98.4	0.6

Entries are percentages, with binomial standard errors (SE).

Table 3.4b: Vital signs – differences between control and intervention hospitals, changes over time and the effect of SPI

	Baseline comparisons		Changes in controls		Effect of SPI	
	OR (99% CI)	p	OR (99% CI)	p	OR (99% CI)	p
On admission:						
1. Temperature*	3.4 (0.6, 19.5)	0.072	4.2 (0.2, 75.8)	0.204	0.3 (0.01, 8.8)	0.360
2. Respiratory rate*	0.5 (0.1, 1.8)	0.166	5.3 (0.3, 90.9)	0.130	1.0 (0.05, 20.1)	0.989
3. Cyanosis/O2 saturation†	0.9 (0.1, 7.6)	0.868	2.0 (0.1, 50.0)	0.578	1.6 (0.04, 60.1)	0.723
4. Confusion/mental state*	1.3 (0.5, 3.6)	0.459	2.1 (1.0, 4.5)	0.011	0.6 (0.2, 1.4)	0.106
5. Pulse*	1.7 (0.2, 14.6)	0.546	2.1 (0.1, 49.3)	0.555	0.8 (0.02, 32.5)	0.884
6. Blood pressure	1.7 (0.2, 14.6)	0.546	2.1 (0.1, 49.3)	0.555	0.8 (0.02, 32.5)	0.884
At six hours:						
7. Temperature†	1.4 (0.7, 2.9)	0.197	1.6 (0.8, 3.3)	0.095	1.2 (0.5, 2.8)	0.606
8. Respiratory rate†	1.1 (0.5, 2.5)	0.764	3.6 (1.7, 7.5)	<0.001	1.7 (0.7, 3.8)	0.124
9. Pulse†	1.2 (0.5, 2.6)	0.581	1.4 (0.6, 3.0)	0.266	1.4 (0.6, 3.5)	0.313
10. O2 saturation	1.3 (0.6,2.5)	0.391	1.7 (0.8, 3.6)	0.062	1.5 (0.6, 3.6)	0.225
At 12 hours:						
11. Temperature†	1.2 (0.6, 2.4)	0.608	2.0 (1.0, 4.2)	0.014	1.1 (0.5, 2.5)	0.846
12. Respiratory rate†	1.1 (0.5, 2.6)	0.658	5.1 (2.4, 10.9)	<0.001	1.1 (0.5, 2.5)	0.819
13. Pulse†	1.0 (0.5, 2.1)	0.943	2.3 (1.1, 5.0)	0.005	0.9 (0.4, 2.2)	0.744
14. O2 saturation†	1.1 (0.6, 2.3)	0.618	3.1 (1.5, 6.5)	<0.001	0.9 (0.4, 2.1)	0.786
Routine investigations:						
15. U & E	1.7 (0.1, 45.2)	0.667	1.3 (0.03, 52.0)	0.872	0.3 (0.003, 18.4)	0.406
16. Chest X-ray	0.7 (0.2, 3.3)	0.599	2.2 (0.2, 21.2)	0.378	0.5 (0.04, 5.4)	0.430
17. Full blood count	2.4 (0.2, 30.2)	0.363	1.4 (0.1, 20.4)	0.727	0.2 (0.01, 6.3)	0.257

* Unadjusted analyses only, owing to convergence difficulties.

† Denotes items with significant (P < 0.010) between hospital variation within the arms of the study.

Compliance at 12 hours deteriorates by about 2% per year of patient age, and this is significant (p<0.01) for items 12, 13 and 14. Item 11 showed a similar reduction, but gave a non-significant result. Item four was positively associated with age (2% greater compliance per year) with p=0.004. It appears that women are more likely than men (p=0.009) to receive a chest X-ray (item 16). The odds-ratio (OR) of 2.1 translates to a percentage increase of around 2% more women than men.

Appropriate clinical response to abnormal vital signs

The data are summarised in tables 3.5a and 3.5b. There is wide variation in the denominators (N) for these items, reflecting the conditional nature of the responses. Mixed effects analysis was attempted for each item, but there were no significant effects between arms or between epochs, and no evidence for any effects associated with the SPI. The component of variation between hospitals was negligible in most cases, and achieved a p value less than 0.1 (= 0.09 in both cases) for only two items.

For most items the data are sparse. No substantive conclusions are indicated.

Other features of good care for specific classes of patients

The effect of SPI is not apparent, and is not statistically significant for any item, as shown in tables 3.6a and 3.6b. The baseline comparisons showed no significant differences between control and SPI1 hospitals. However, compliance appears to have improved on item 5 (use of CURB score), though from a very low base.

A strong negative age-effect was apparent for item 4 yielding a reduction in odds of compliance of about 8% per year of age.

There is a strong positive reviewer learning effect ($p<0.001$) for item 2 (oxygen prescription for COPD).

Prescribing error

These results are presented in tables 3.7a and 3.7b. There are more prescriptions per patient in the SPI1 hospitals (29.3), compared to the control hospitals (24.9), and a small increase (about two per patient) across epochs in both arms. Unadjusted analysis suggests an increase in error rate associated with the SPI of marginal significance ($p=0.041$).

The rate of error detection was found to change with time in a systematic way as the (single) reviewer gained more experience with the semi-structured task of identifying medication errors from case notes.

Reviews took place from November 2006 to August 2009. The rate of detected errors of the case note review was found to improve at first, peaking at around July/August 2008 but declined thereafter.

The control hospital data was reviewed during the later period, when the detected error rate was declining. SPI1 hospitals were reviewed while it was increasing.

Table 3.5a: Appropriate clinical response – compliance with standards

	Control hospitals						SPI1 hospitals					
	Epoch 1			Epoch 2			Epoch 1			Epoch 2		
	N	%	SE	N	%	SE	N	%	SE	N	%	SE
Oxygen saturation <90, at any time												
Full blood gases within 2 hours	15	60.0	12.6	20	60.0	11.0	130	36.9	4.2	100	54.0	5.0
Given oxygen if not on oxygen	16	68.8	11.6	16	68.8	11.6	100	79.0	4.1	58	77.6	5.5
Doctor called or transferred to ITU if on oxygen	10	30.0	14.5	11	63.6	14.5	114	36.0	4.5	60	36.7	6.2
Blood pressure systolic <90												
At least next 6 hours, hourly observations	21	19.0	8.6	27	14.8	6.8	36	25.0	7.2	35	20.0	6.8
Blood culture	18	33.3	11.1	23	39.1	10.2	27	33.3	9.1	31	38.7	8.7
Sputum present												
Sputum culture	141	39.0	4.1	150	47.3	4.1	215	47.0	3.4	256	51.6	3.1
Respiratory rate >20 at any time after admission												
Given oxygen (if not on oxygen)	5	0.0	0.0	1	0.0	0.0	96	20.8	4.1	55	7.3	3.5
Doctor called (if on oxygen)	8	0.0	0.0	3	33.3	27.2	27	7.4	5.0	18	16.7	8.8
Temperature over 38 C – any episode												
If yes, blood culture	35	71.4	7.6	39	74.4	7.0	73	72.6	5.2	89	76.4	4.5
Failure to improve within 48 hours or subsequent deterioration												
Review by consultant	20	100.0	0.0	22	100.0	0.0	41	100.0	0.0	38	100.0	0.0
Repeat chest X-ray	18	100.0	0.0	17	100.0	0.0	36	69.4	7.7	30	73.3	8.1
White cell counted/repeated	19	100.0	0.0	22	100.0	0.0	41	95.1	3.4	37	97.3	2.7
Appropriate addition of further antibiotics	16	93.8	6.0	11	100.0	0.0	31	64.5	8.6	24	66.7	9.6
Follow up												
Clinical review arranged 6 weeks after discharge	112	61.6	4.6	112	65.2	4.5	264	43.9	3.1	277	45.5	3.0

Note: The columns headed N represent the opportunities for error. The opportunities vary within categories, e.g. the reviewer may judge that it would have been inappropriate to call a doctor or move a patient to ICU despite falling oxygen saturation, e.g. because death was expected.

Entries are error rates as percentages of N, with binomial standard errors.

Table 3.5b: Appropriate clinical response – differences between control and intervention hospitals, changes over time and the effect of SPI

	Comparisons at epoch 1		Change in controls		Effect of SPI	
	SPI1/control		Epoch 2/epoch 1		Ratio of temporal changes	
	OR (99% CI)	p	OR (99% CI)	p	OR (99% CI)	p
Oxygen saturation <90, at any time						
Full blood gases within 2 hours	1.0 (0.2, 5.9)	1.000	0.6 (0.1, 4.9)	0.531	2.3 (0.2, 22.6)	0.333
Given O2 if not on O2	0.9 (0.1, 8.5)	0.893	0.8 (0.1, 7.2)	0.893	1.3 (0.1, 14.3)	0.763
Doctor called or transferred to ITU if on O2	3.3 (0.3, 36.4)	0.201	1.9 (0.1, 29.3)	0.554	0.3 (0.01, 4.6)	0.223
Blood pressure systolic <90						
At least next 6 hours, hourly observations	4.1 (0.2, 80.2)	0.223	0.6 (0.1, 5.1)	0.579	0.6 (0.04, 8.8)	0.627
Blood culture	1.3 (0.1, 15.2)	0.794	1.3 (0.2, 7.5)	0.726	0.7 (0.1, 7.2)	0.651
Sputum present						
Sputum culture	1.4 (0.5, 4.4)	0.391	1.4 (0.7, 2.6)	0.221	0.9 (0.4, 2.0)	0.651
Respiratory rate >20 at any time after admission						
Given oxygen (if not on oxygen)	–	–	–	–	–	–
Doctor called (if on oxygen)	–	–	–	–	–	–
Temperature over 38 C – any episode						
If yes, blood culture	1.3 (0.2, 8.0)	0.754	1.0 (0.3, 4.2)	0.957	0.9 (0.1, 5.4)	0.865
Failure to improve by 48 hours or subsequent deterioration						
Review by consultant	–	–	–	–	–	–
Repeat chest x-ray	–	–	–	–	–	–
White cell counted/repeated	–	–	–	–	–	–
Appropriate addition of further antibiotics	1.1 (0.04, 317.5)	0.960	–	–	–	–
Follow up						
Arrange clinical review within 6 weeks	0.9 (0.3, 2.2)	0.672	1.2 (0.6, 2.6)	0.461	0.7 (0.3, 1.8)	0.388

Table 3.6a: Features of good care for specific classes of patient

	Control hospitals						SPI1 hospitals					
	Epoch 1			Epoch 2			Epoch 1			Epoch 2		
	N	%	SE	N	%	SE	N	%	SE	N	%	SE
1. Asthma or COPD: given steroids within 24 hrs	129	87.6	2.9	135	92.6	2.3	224	91.1	1.9	199	88.4	2.3
2. Peak flow record	34	79.4	6.9	29	82.8	7.0	78	82.1	4.3	37	64.9	7.8
3. Severity of pneumonia patients assessed	101	75.2	4.3	113	73.5	4.2	170	73.5	3.4	189	70.4	3.3
4. Is this based on CURB score in notes?	102	2.0	1.4	111	23.4	4.0	170	2.4	1.2	189	8.5	2.0
5. Was appropriate antibiotic treatment given?	100	93.0	2.6	110	95.5	2.0	169	94.7	1.7	189	93.1	1.8

Table 3.6b: Steroids and antibiotics – differences between control and intervention hospitals, changes over time and the effect of SPI

	Baseline comparisons		Change in controls		Effect of SPI	
	OR (99% CI)	p	OR (99% CI)	p	OR (99% CI)	p
1. Asthma or COPD given steroids within 24 hrs	1.6 (0.4, 5.9)	0.385	1.7 (0.6, 5.3)	0.219	0.7 (0.2, 2.9)	0.500
2. Peak flow record†	1.0 (0.03, 30.2)	0.974	1.1 (0.1, 11.1)	0.896	0.6 (0.04, 8.2)	0.570
3. Severity of pneumonia patients	1.7 (0.5, 5.2)	0.245	1.1 (0.4, 2.6)	0.854	0.8 (0.2, 2.4)	0.553
4. Is this based on CURB score in notes?	0.7 (0.04, 11.7)	0.753	17.0 (2.3, 125.8)	<0.001	0.3 (0.02, 3.4)	0.173
5. Was appropriate antibiotic treatment given?	1.5 (0.2, 14.8)	0.626	2.0 (0.4, 10.8)	0.303	0.5 (0.1, 4.2)	0.383

† Denotes items with significant between hospital components of variation within the arms of the study ($p<0.01$).

Table 3.7a: Prescribing errors

	Control hospitals				SPI1 hospitals			
	Epoch 1		Epoch 2		Epoch 1		Epoch 2	
No. of patients	233		239		381		378	
No. of prescriptions	5482		6207		10664		11538	
Prescriptions per patient	23.5		26.0		28.0		30.5	
Errors								
Total	596		564		1157		1530	
By type of error								
Counsel	1		0		0		2	
Monitor	0		0		1		1	
Need	56		95		114		190	
Dose	287		224		591		616	
Drug	23		13		46		55	
Formula	40		39		41		73	
Supply	189		193		364		593	
	Rate	**SE**	**Rate**	**SE**	**Rate**	**SE**	**Rate**	**SE**
Unadjusted rates								
Error rate per prescription	0.115	0.010	0.093	0.008	0.111	0.012	0.132	0.014
Rates adjusted for date of review								
Overall rate (all errors)	0.137	0.016	0.111	0.014	0.146	0.017	0.146	0.013
By type of error								
Need	0.015	0.003	0.023	0.004	0.014	0.002	0.018	0.002
Dose	0.067	0.009	0.048	0.008	0.059	0.009	0.053	0.007
Drug	0.007	0.002	0.004	0.002	0.003	0.001	0.004	0.001
Formula	0.007	0.002	0.006	0.002	0.006	0.001	0.007	0.001
Supply	0.035	0.006	0.029	0.005	0.069	0.011	0.077	0.010

Table 3.7b: Prescribing errors – differences between control and intervention hospitals, changes over time and the effect of SPI

	Baseline comparisons		Changes in controls		Effect of SPI	
	Rate ratio (99% CI)	p	Rate ratio (99% CI)	p	Rate ratio (99% CI)	p
Overall rate (all errors)	1.0 (0.6, 1.5)	0.789	0.8 (0.6, 1.1)	0.048	1.2 (0.9, 1.8)	0.138
By type of error						
Need	1.0 (0.5, 1.9)	0.879	1.5 (0.9, 2.5)	0.045	0.8 (0.4, 1.6)	0.438
Dose	0.9 (0.5, 1.5)	0.553	0.7 (0.5, 1.0)	0.011	1.2 (0.8, 1.9)	0.201
Drug	0.3 (0.1, 0.8)	0.002	0.5 (0.2, 1.5)	0.123	2.7 (0.8, 9.7)	0.041
Formula	0.8 (0.3, 2.2)	0.659	0.9 (0.5, 1.7)	0.598	1.4 (0.5, 3.4)	0.319
Supply	1.8 (1.0, 3.1)	0.012	0.8 (0.6, 1.2)	0.179	1.4 (0.9, 2.3)	0.064

Rate-ratios are estimated from a population-averaged negative binomial model.

Table 3.7c: Anti-coagulant prescribing errors

	Control hospitals				SPI1 hospitals			
	Epoch 1		Epoch 2		Epoch 1		Epoch 2	
No. of patients	52		93		167		224	
No. of prescriptions	83		132		274		362	
No. of errors	1		5		25		32	
	Rate	SE	Rate	SE	Rate	SE	Rate	SE
Unadjusted rates								
Error rate per prescription	0.011	0.012	0.038	0.018	0.088	0.024	0.089	0.022
Rates adjusted for date of review								
Overall rate (all errors)	0.020	0.023	0.070	0.050	0.169	0.055	0.094	0.025

Table 3.7d: Anti-coagulant prescribing errors – analysis

	Baseline comparisons		Changes in controls		Effect of SPI	
	Rate ratio (99% CI)	p	Rate ratio (99% CI)	p	Rate ratio (99% CI)	p
Overall rate (all errors)	8.5 (0.4, 181.1)	0.071	3.1 (0.2, 56.7)	0.317	0.2 (0.01, 3.7)	0.146

In the intervention arm, randomisation of the order of review was only partial: on average, the review date for epoch 1 was earlier than for epoch 2, though there was considerable overlap between the dates of the two sets of reviews.

Therefore there is the potential for confounding between the SPI effect and the date of review.

After adjustment for date of review (which was highly significant, $p<0.001$) there are no significant differences between arms or epochs and no effect associated with SPI.

A specific breakdown of errors relating to anti-coagulant administration was carried out because this treatment was particularly stressed by IHI (table 3.7c).

No differences were observed, but the denominators are small, especially in control hospitals. Examples of prescribing errors are given in table 3.8.

Medicines reconciliation errors at admission

The results can be found in tables 3.9a and 3.9b. Again, there is no significant evidence that the SPI has an effect. The arms of the study are very similar at baseline and there is a tendency for this type of error to increase over epochs in both control and SPI1 hospitals.

Holistic review – quality of care and errors

The sample

The number of case notes reviewed by the holistic method differs, being higher than the number reviewed by the explicit review method (see Explicit case note review p36). This is because the commissioning of this review started at a later date. In the four SPI1 hospitals, 390 and 381 case notes were holistically reviewed from epoch 1 and epoch 2 respectively (roughly equally divided between the four hospitals. For the 18 control hospitals, 243 and 246 case notes were reviewed from epoch 1 and epoch 2 (range eight to 15 cases per hospitals).

Reliability

In total, 122 case notes were reviewed by both reviewers. Measures of reliability between the two holistic reviewers were, as expected, low[53] (ICCs were 0.05 (99% CI: -0.13, 0.23) for the admission rating; 0.19 (99% CI: -0.05, 0.23) for the management rating; 0.21 (99% CI: -0.02,0.42) for the pre-discharge care rating and 0.29 (99% CI: 0.06, 0.49) for the overall care rating).

Table 3.8: Examples of prescribing errors relating to each stage of the drug use process found in this study

Category of prescribing error	Examples from case notes reviewed
Need for drug	• Rabeprazole 10mg oral once a day was taken by patient before admission but was not prescribed during admission • Patient usually takes digoxin 125mcg oral once a day, but this was not prescribed on admission
Selection of drug	• Tiotropium 18mcg inhaler once a day prescribed at the same time as Combivent (salbutamol and ipratropium) inhaler two puffs four times a day. This is drug duplication as both of these drugs have the same pharmacological action • Patient is allergic to penicillin but was given one stat dose of 500mg oral amoxicillin
Selection of dose	• Doctor prescribed Combivent (salbutamol and ipratropium) inhaler four puffs four times a day. This was a wrong dose (overdose) as the maximum should have been two puffs four times a day • Paracetamol 1g oral to be given when required prescribed without indicating the maximum daily frequency/dose
Selection of formulation	• Seretide 250 inhaler two puffs twice a day prescribed without specifying whether evohaler or accuhaler • Dipyridamole 200mg orally twice a day prescribed without indicating that modified release formulation intended
Provide information needed for supply	• Co-amoxiclav 625mg three times a day prescribed without indicating the route of administration • Clopidogrel 75mg oral once a day prescribed and given without having a signature of prescriber

Table 3.9a: Reconciliation errors at admission

	Control hospitals		SPI1 hospitals	
	Epoch 1	Epoch 2	Epoch 1	Epoch 2
No. of admissions	203	188	380	377
Admissions with reconciliation errors				
N	14	21	24	41
% (SE)	6.9 (1.8)	11.1 (2.3)	6.3 (1.2)	10.9 (1.6)
Mean no. of errors when error is present (SE)	1.6 (0.2)	2.2 (0.3)	2.3 (0.4)	2.1 (0.2)

Table 3.9b: Reconciliation errors at admission – differences between control and intervention, changes over time and the effect of SPI

	Comparisons at epoch 1		Changes in controls		Effect of SPI	
	SPI1/control		Epoch 2/epoch 1		Ratio of temporal changes	
	OR (99% CI)	p	OR (99% CI)	p	OR (99% CI)	p
Admission with reconciliation error	1.1 (0.3, 4.3)	0.839	1.5 (0.6, 4.0)	0.241	0.8 (0.3, 3.0)	0.770

Odds-ratios (OR) derive from a logistic model with random effects for hospitals, adjusted for the date of review.

The main reviewer tended to assign higher average ratings with more variability, whereas the second reviewer tended to assign lower average ratings with less variability. The inter-rater agreement measures between reviewers, for identifying patients who had experienced an error as part of their overall care, were low (Kappa: 0.15, se 0.08).

Quality of care ratings

The average scores during epoch 1 (with standard errors) for admission, management and pre-discharge ratings were 5.0 (0.05), 4.2 (0.07) and 4.3 (0.07) respectively on a scale of one (below best practise) to six (excellent care). The average score for overall care was 7.4 (0.06), on a scale of one (unsatisfactory) to 10 (very best care).

Admission, management and pre-discharge care ratings were higher in the SPI1 hospitals compared with the control hospitals, during both epoch 1 and epoch 2 (table 3.10), although not significantly so. However, the overall care rating was higher in the control hospitals during epoch 1 (although again not significantly so), but similar during epoch 2.

In addition, all ratings tended to increase in epoch 2 as compared with epoch 1. This pattern was more consistent across intervention hospitals, where not only did all rating increase, but the admission rating increased significantly between epochs (increase 0.28, p=0.001) However, differences in changes across control and SPI1 hospitals were not significant for any of the four ratings (table 3.10).

Errors

The numbers of errors per 100 patients were lower in the SPI hospitals compared to the control hospitals, for both for epoch 1 and epoch 2 (table 3.10). In the control hospitals, there was around

one error for every two patients, whereas in the SPI1 hospitals there was around one error for every three patients. The numbers of errors decreased in epoch 2 (for both the control and SPI1 hospitals), although this difference was not significant. Again, differences in changes across control and SPI1 hospitals were not significant for errors.

A total of 425 errors were identified (table 3.11). The most frequent categories of errors related to diagnosis, assessment or admission, or were errors relating to poor clinical reasoning.

Errors relating to poor clinical reasoning were more frequent in the control hospitals (in both epoch 1 and epoch 2), and although they decreased in the control hospitals in epoch 2, they increased in the SPI1 hospitals in epoch 2.

Rates of other errors also differed between control and SPI1 hospitals and between epoch 1 and epoch 2, although no differences in changes were significant.

3.5 Sub study 5: Outcomes

Adverse events

The holistic review estimated an adverse event rate of about four per 100 patients treated, which is comparable to the published literature (table 3.10).[38–41] The inter-rater agreement for identifying patients who had experienced an adverse event was low (Kappa=0.25 se 0.09).

The rate of adverse events per 100 patients was higher in the SPI1 hospitals compared with the control hospitals in epoch 1, but the reverse was the case for epoch 2 (table 3.10).

The number of adverse events per 100 patients decreased during epoch 2 in the SPI1 hospitals, while the number of adverse events increased during epoch 2 in the control hospitals (table 3.10). However, once again differences in changes were not significant. A trend in favour of the SPI1 hospitals was observed for five of the six categories of adverse events (table 3.12), but no difference in change was significant.

For approximately one quarter of the adverse events, there was strong or certain evidence that the event was preventable. At around 1.3% (16/1,260), this is a somewhat lower rate of adverse events than reported for hospital inpatients elsewhere.[41] Four patients died where there was more than 50% probability that death resulted from an adverse event.

Table 3.10: Holistic review – changes in ratings and numbers of adverse events and errors between control and SPI1 hospitals

	Control hospitals				SPI1 hospitals				Difference in change
	Epoch 1		Epoch 2		Epoch 1		Epoch 2		(99% cis)^
No. of patients		243		246		390		381	
Quality ratings	Ave	SE	Ave	SE	Ave	SE	Ave	SE	
Admission rating[1]	4.9	0.08	4.9	0.08	5.0	0.07	5.3	0.05	0.23 (-0.14, 0.60)
Management rating[1]	4.2	0.12	4.1	0.12	4.3	0.10	4.6	0.09	0.35 (-0.19, 0.90)
Pre-discharge rating[1]	4.2	0.11	4.2	0.10	4.3	0.08	4.4	0.08	0.06 (-0.42, 0.54)
Overall care rating[2]	7.6	0.09	7.6	0.09	7.4	0.08	7.6	0.07	0.27 (-0.17, 0.70)
Errors/adverse events	*Rate*	*SE*	*Rate*	*SE*	*Rate*	*SE*	*Rate*	*SE*	
No. errors per 100 patients*	44.44	3.8	42.3	3.8	29.7	2.5	24.4	2.3	-2.42 (-17.99, 13.31)
No. adverse events per 100 patients*	2.9	1.2	4.8	1.3	6.2	1.2	3.7	1.1	-3.92 (-10.39, 2.55)
% of preventable adverse events	0		30	13	28	9	36	11	-22 (-67, 30)

[1] Score scale: 1 (below best practice) to 6 (excellent care).

[2] Score scale: 1 (unsatisfactory) to 10 (very best care).

Patients could experience more than one error and more than one adverse event.

^ Difference in changes are estimated from a mixed effects model (see methods for details).

Table 3.11: Rates (per 100 patients) of errors identified by broad category of error – holistic review

*Errors can be of multiple categories.	Control hospitals				SPI hospitals				Effect of SPI
	Epoch 1		Epoch 2		Epoch 1		Epoch 2		Rate ratio (99% CI)
No. of patients	243		246		390		381		
No. of errors*	111		104		116		94		
Rate of errors	44.4 (3.8)		42.3 (3.8)		29.7 (2.5)		24.4 (2.3)		0.87 (0.52, 1.44) p=0.48
	Control hospitals				SPI hospitals				Effect of SPI
	Epoch 1		Epoch 2		Epoch 1		Epoch 2		Rate ratio (99% CI)
	Rate	*SE*	*Rate*	*SE*	*Rate*	*SE*	*Rate*	*SE*	
No. errors per 100 patients*	44.44	3.8	42.3	3.8	29.7	2.5	24.4	2.3	-2.42 (-17.99, 13.31)
Diagnosis/assessment/admission error	54.3	5.0	48.8	4.7	31.8	3.1	27.8	3.1	1.02 (0.64, 1.65) p=0.90
Hospital-acquired infection	0		0		0.26	0.26	1.05	0.52	–
Technical/management	7.4	1.7	8.9	1.8	3.3	0.9	2.4	0.8	0.60 (0.15, 2.43) p=0.35
Medication/maintenance/follow-up	23.5	3.1	16.7	2.6	18.7	2.2	15.2	2.0	1.20 (0.59, 2.42) p=0.51
Clinical reasoning	30.9	3.0	28.5	2.9	10.5	1.6	16.5	1.9	1.83 (0.92, 3.63) p=0.02
Discharge information	11.9	2.1	15.4	2.3	10.3	1.5	9.4	1.5	0.75 (0.31, 1.81) p=0.41

In two cases, the reviewer felt that the death was definitely caused by the error (untreated, documented hyperkalaemia and failure to recognise adrenal crisis) and in two further cases that it was more likely than not (wrong choice of antibiotic and insulin overdose).

Patient mortality – outcome

The analysis was adjusted for age, sex and the number of co-morbidities, though only age was significant ($p<0.001$). The odds of death increased by 8% (CI 5%–11%) per year of patient age.

The effect of SPI was not significant. Baseline comparisons showed no significant differences between control and SPI1 hospitals; neither was there significant evidence of temporal change in the control hospitals (tables 3.14a and 3.14b).

Patient survey

The response rate for the first survey was 54% (1,961 of 3,624 returned) in the four SPI1 hospitals. For the second survey it was 51% (1,720 out of 3,397). In the 18 control hospitals there was a greater drop, from 63% to 56%.

Table 3.15 shows the values of the five survey scores in each of the four SPI1 hospitals for the two surveys, along with details of response rates.

Table 3.16 shows the changes in both control and SPI1 hospitals on each of the five scores identified, along with the differences between the groups in these changes and associated 99% confidence intervals.

At baseline, there were no statistically significant differences between control and SPI1 hospitals on any of the scores. One of the survey scores showed a significantly different change between the control and SPI1 sites.

The rating of cleanliness of toilets and bathrooms decreased in the control sites, from 79 to 77 points, whereas this increased in SPI1 hospitals, from 74 to 76 points ($p=0.009$). It is noteworthy that there was apparently a baseline difference between the two groups of hospitals here, and although this difference was not statistically significant ($p=0.115$), the SPI1 hospitals were still slightly poorer than the control hospitals in the second survey despite the change.

None of the other four scores showed any significantly different changes between the two groups

Table 3.12: Rates per 100 patients of adverse events identified by broad category of adverse events

	Control hospitals				SPI1 hospitals				Difference in change
	Epoch 1		Epoch 2		Epoch 1		Epoch 2		(99% CIs)
No. of patients		243		246		390		381	
No. of adverse events*		7		11		24		14	
Category of adverse event	*Rate*	*SE*	*Rate*	*SE*	*Rate*	*SE*	*Rate*	*SE*	
Diagnosis/assessment/admission error	2.47	1.00	2.44	0.99	6.67	1.59	3.67	1.16	-2.79 (-9.83, 4.26)
Hospital-acquired infection	0.82	0.58	2.03	0.90	3.08	0.88	1.57	0.64	-2.62 (-6.74, 1.50)
Technical/management	0.41	0.41	1.22	0.70	0.51	0.36	0.52	0.37	-0.79 (-3.19, 1.60)
Medication/maintenance/follow-up	0.41	0.41	0.81	0.57	2.05	0.71	0.52	0.37	-1.97 (-5.01, 1.07)
Clinical reasoning	0.41	0.41	0		2.05	0.72	0.79	0.45	-0.87 (-3.79, 2.05)
Discharge information	0.82	0.58	0		0.26	0.26	1.04	0.52	1.63 (-0.62, 5.65)
									Rate ratio (99% CI)
Rate of adverse events	2.9	1.2	4.8	1.3	6.2	1.2	3.7	1.1	0.40 (0.09, 1.84) p=0.12

*A single adverse event may occupy more than one category.

Table 3.13: Preventable adverse events identified as being strongly* or certainly preventable out of the 1260 case notes reviewed in the holistic review

	Control hospitals	Intervention hospitals
Epoch 1	1. Given oxygen and became unrousable from CO_2 retention requiring ITU admission	1. Loss of consciousness due to hypoglycaemia caused by an overdose of insulin to control hyperkalaemia (patient died)
		2. Supra-ventricular tachycardia in patient with untreated hypokalaemia (patient died)
		3. Wrong choice of antibiotic for severe community-acquired pneumonia (patient died)
		4. Deterioration in breathlessness because nurse omitted scheduled use of nebuliser
		5. Sent home with severe uninvestigated anaemia. Symptoms likely and very high risk †
		6. Started on treatment for hypothyroidism despite equivocal test result (and in wrong dose)
		7. Bronchospasm could have been avoided or lessened had beta blocker been stopped
Epoch 2	1. Loss of consciousness due to hypoglycaemia caused by an overdose of insulin to control hyperkalaemia*	1. Collapse due to adrenal crisis because corticosteroids were not prescribed for patient with known Addison's disease (patient died)
	2. Delay in administration of vitamin K leading to haematoma	2. Failure to treat MRSA and GP not informed on discharge. No absolute evidence of harm but very high risk
	3. Breathlessness increased, requiring transfer to high dependency unit, following failure to administer prescribed antibiotics	3. Severe anaemia not investigated and GP not informed. No harm in hospital but very high risk and symptoms likely†
		4. Bronchospasm could have been avoided or lessened had beta blocker been stopped
		5. Failure to inform GP of the risk of CO_2 retention by giving patient oxygen†

*More likely than not on the balance of probabilities.

† There is no absolute evidence of harm in these cases but patients were discharged in clear danger and this influenced the reviewer.

Table 3.14a: Mortality rates

	Control hospitals		SPI1 hospitals	
	Epoch 1	Epoch 2	Epoch 1	Epoch 2
No. of patients	236	240	381	380
Deaths	27	39	63	49
% mortality (SE)	11.4 (2.1)	16.3 (2.4)	16.5 (1.9)	12.9 (1.7)
Age: mean (SD)	77.6 (7.6)	79.7 (7.7)	77.4 (7.6)	78.2 (8.0)
% female	58.5	52.1	50.4	51.8
Co-morbidities: mean	2.8	3.1	3.3	3.8

Table 3.14b: Analysis of mortality rates

	Baseline comparisons		Changes in controls		Effect of SPI	
	Odds ratio (99% CI)	p	Odds ratio (99% CI)	p	Odds ratio (99% CI)	p
Admission mortality (adjusted for age, sex, no. of co-morbidities)	1.9 (0.6, 5.6)	0.149	1.4 (0.7, 2.9)	0.274	0.5 (0.2, 1.4)	0.085

Odds-ratios (OR) derive from a logistic model with random effects for hospitals, adjusted for the date of review.

Table 3.15: Patient survey scores in SPI1 hospitals at the two periods

	Hospital 1		Hospital 2		Hospital 3		Hospital 4	
	Survey 1	Survey 2	Survey 1	Survey 2	Survey 1	Survey 2	Survey 1	Survey 2
Overall, how would you rate the care you received?	71	70	81	79	76	78	80	78
Overall, did you feel you were treated with respect and dignity while you were in the hospital?	84	82	90	89	87	84	91	88
How would you rate how well the doctors and nurses worked together?	72	70	79	78	74	75	79	77
In your opinion, how clean was the hospital room or ward that you were in?	77	76	83	82	79	78	78	80
How clean were the toilets and bathrooms that you used in hospital?	70	71	78	79	73	74	77	78
Response rate	56%	46%	71%	59%	49%	53%	49%	45%

Table 3.16: Patient survey scores in control and SPI1 hospitals at the two periods

	Control hospitals			SPI1 hospitals			Range at baseline	difference in change (99% CI)	p
	Survey 1	Survey 2	Absolute % change	Survey 1	Survey 2	Absolute % change			
Overall, how would you rate the care you received?	79	77	-2	77	76	-1	71–83	1 (-1, 3)	0.330
Overall, did you feel you were treated with respect and dignity while you were in the hospital?	89	87	-1	88	86	-2	84–91	-1 (-2, 1)	0.269
How would you rate how well the doctors and nurses worked together?	78	76	-2	76	75	-1	72–82	1 (-1, 3)	0.135
In your opinion, how clean was the hospital room or ward that you were in?	82	80	-1	79	79	0	71–89	1 (-1, 4)	0.288
How clean were the toilets and bathrooms that you used in hospital?	79	77	-2	74	76	1	69–86	3 (0, 6)	0.009

Chapter 4

Discussion

4.1 Main findings

As the stakeholder interviews and focus groups demonstrated, the SPI was greeted enthusiastically at a strategic level.

However, the ethnography suggested that front-line staff on medical wards had a vague idea of the intervention and few had direct experience of most of its components, except in the area of recognising and responding to the deteriorating patient.

A similar picture emerges from the staff survey. Control and SPI1 hospitals were mostly indistinguishable at baseline and only one of the 11 dimensions of staff satisfaction changed significantly over time (but to a small degree) – on the item relating to organisational climate.

Taken together, these findings suggest that the impact of SPI at medical ward level was at best modest.

Quantitative evaluation of response to specific SPI targets (items 1a, b, c; 2a, b and 3a in table 1.1) also yielded a null result, thereby corroborating the qualitative finding of apparently low impact of the programme on the sharp end of practice. The important SPI aim of improving response to acutely ill patients, including the quality of the recording of vital signs, improved markedly and significantly during the study period in the SPI1 hospitals, but a similar improvement was also observed in the control hospitals. This is probably due to policy shifts and other external imperatives, encouraging better detection and response to deteriorating patients. The use of the CURB score also improved markedly, but there was a trend towards this being in favour of control hospitals.

Prescribing error rates are very sensitive to the methodology[54] used to make the measurements. Inter-observer reliability was good. By overlapping observation periods between epochs, we were able to detect learning and fatigue effects and hence allow for these in the analysis. The prescribing error rates we observed were quite high when compared with the only other study using the same methods in two hospitals (7.4 and 8.6%).[35]

This study found no improvement in prescribing over epochs, and there was no difference between control and SPI1 sites, suggesting that there was no SPI effect on prescribing error rates.

Many prescribing errors are of a minor nature. The extent to which such minor errors are a surrogate for more serious errors, as implied by the Heinrich ratio, is contested.[54] Although we uncovered a high rate of prescribing error, very few of these errors resulted in adverse events for patients. With so few events due to prescribing error, this end-point cannot reliably be used to confirm or refute an SPI effect.

Errors associated with anticoagulation therapy are potentially a particular cause of concern[40] and were therefore a specific SPI target. We found no trends towards fewer errors over time in anticoagulation therapy and only one adverse event associated with this class of drug (see table 13.13).

Medicines reconciliation was another key SPI target, but there was no trend towards improvements either over time, or between the SPI hospitals and the control hospitals. We observed a number of clinical processes that were not specific SPI targets. These might have been expected to improve, if the overall goal of strengthening the system and achieving cultural and organisational realignments around safety had been achieved. Again there were no significant differences between control and SPI1 hospitals over time.

For some measures – such as use of corticosteroids in COPD and asthma – this was because practice was already good at baseline and there was little room for further improvement. However, there was also no change in the quality of medical history taking or appropriateness of antibiotic selection, even though there was room for improvement here.

Use of blood gases, when indicated, improved in SPI hospitals but this was not significantly different from control hospitals. The holistic review corroborated findings from the explicit review, showing no improvement in quality and no reduction in either errors or adverse events in medical wards treating patients with acute respiratory disease in control versus SPI1 hospitals.

A non-statistically significant drop in mortality in SPI1 hospitals was observed among patients included in the case note review while the mortality rate in control hospitals increased, even after adjusting for age differences across epochs.

It seems unlikely that failure to implement SPI caused an increase in mortality among respiratory patients in control hospitals. This suggests that the trend towards differences in changes across control and SPI1 hospitals should be interpreted with caution.

The difference in rate of change mortality was not significant after adjusting for differences in age.

4.2 Strengths and weaknesses

The main strengths of the study lie in:

- Objective measures of safety practice using observers who were entirely independent of the hospitals. Care was taken to allow for changing discrimination (learning and fatigue effects) during the course of the study.
- A number of different (qualitative and quantitative) observations were made across the hospital system to explain and contextualise the direct measurements of safe practice and allow triangulation[55] of both data collection and interpretation.
- Use of a before and after controlled design.[16] With some notable exceptions[56;57] most quality improvement reports lack contemporaneous controls. Such a design would evidently have been misleading in this case since the sharp improvement of monitoring of vital signs and use of a formal scoring system in the SPI1 hospitals could have been incorrectly attributed to the SPI.
- Measurement of learning/fatigue effects (not just inter-observer variation) for quantitative components. This showed that, as expected from a review of the literature,[53] reliability was high for explicit review of clinical process (that is, error), low for holistic review of clinical process and intermediate for holistic review of adverse events. This allows the reader to be discriminating, placing more weight where reliability is moderate or high rather than where it is low.

A limitation of this study was that non-randomised and controls were matched with SPI2 rather than first phase SPI1 intervention hospitals. Hospitals were selected for SPI1 because they were perceived to have contained positive features (see introduction).

Results might be biased because SPI hospitals have less headroom for improvement and controls have higher than average performance, particularly since half were also selected as future SPI2 intervention sites. However, the possibilities are not supported by baseline comparisons – performance data and staff and patient survey results were very similar across cases and controls. The assessment of quality of medical history taking identified two of 11 items where compliance was higher among SPI1 than control hospitals at baseline.

Completeness of vital signs recording also had a slight (but non-significant) difference in favour of SPI1 hospitals at baseline.

Error rates at baseline were lower in SPI1 hospitals on holistic review but adverse events were observed less often among controls. Results might have been biased in favour of SPI because intervention sites were selected, not chosen at random (the reverse of the possible bias mentioned above).

In addition, both control and intervention sites gave consent for the evaluation and this may have had a differentially motivating effect in intervention hospitals – an effect that could not be avoided by randomisation. These potential biases against controls would have been scientifically more worrisome had the results not been mostly null.

Overall, there do not appear to be material differences in performance between control and SPI1 hospitals at baseline. Most observations are similar and where statistically significant differences exist, these are small. The data do not support the idea that the SPI1 hospitals had such excellent practice at baseline compared to controls that they were jeopardised in the comparison.

We did not make observations of adverse events after discharge from hospital, though it is likely that some did occur beyond this stage. However, there is little evidence from this study to support the hypothesis that comparison of event rates post-discharge would have favoured either arm of the study.

4.3 Explaining the results

Science cannot prove a null result and the possibility of positive effects at some level within the targeted hospitals cannot be ruled out. The study may have failed to identify improved practice for many reasons.

Detecting improvements

Improvements may have occurred at a magnitude that eluded statistical detection. The sample was large, with over 1,200 case notes reviewed, with sufficient statistical power to detect material changes in actions that should affect all patients, such as regular monitoring of all vital signs. Power was lower for contingent actions that only applied to smaller sub-groups (that is, for patients whose condition deteriorated).

The English threshold under which an intervention is judged cost-effective is about £30,000 per QALY. The SPI would, therefore, need to save fewer than seven lives with a mean duration of five years to justify the investment of about £775,000 per hospital (ignoring discounting and assuming disability-free life).

It would not be possible to exclude an effect of this magnitude in a study of any feasible size. With many hundreds of deaths taking place in each hospital each year, the signal would be lost in the noise.

Organisational-level change

Our evaluation sought to assess organisational-level change. This was based on the prior hypothesis that the multi-component SPI intervention would affect endpoints specified in advance of the data collection.

However, improvements may have taken place in clinical areas targeted by SPI but not observed in our study (items 4a, b, c and d in table 1.1), such as ICUs and surgical departments.

These were not included in our study because of the expense of auditing the quality of intensive care and because one of the hospitals selected for the intervention did not have an ICU. It is nonetheless possible that the SPI did have stronger impacts in these more highly controlled settings.

Indeed, there were some suggestions in the focus groups that ICUs were perhaps more receptive to the programme and that it was possible to mobilise support for the programme more easily. Some of these clinical areas were included in the evaluation of SPI2.

There may have been improvements in aspects of safety targeted by individual hospitals

Care of cardiovascular disease was identified for special attention in some hospitals. It is likely that such initiatives, if vigorously pursued, would result in improvement.

However, the study was not designed simply to answer the question: 'Can a clinical practice ever be improved as a result of specific managerial intervention?' The answer to this question is clearly 'yes' – many spectacular examples, including the Michigan study of prevention of central-line infections,[58] can be found in the literature.

The question in this study concerned the average effect that may be expected among a series of practices aimed at improving patient safety, some specific and some more generic, that were specified in the study protocol in advance of the data collection.

The effect of participation in SPI

There is an argument that participation in SPI may secure greater long-term commitment to quality and safety, and improvements made in the intervention hospitals will either surface at a later date or be sustained better.

This hypothesis can only be tested with further data collection.

Design and implementation of SPI

It is possible that the design and implementation of SPI might not have been optimal.

While senior stakeholders stressed the bottom-up nature of the intervention, this was not how it was perceived by most ward staff. Although there were examples of PDSA cycles triggered by clinical staff, these were not replicated on a scale where the benefits were likely to show up in an independent quality audit, based on predefined criteria.

Despite the enthusiasm and broad understanding of the principles underlying the SPI at a strategic level, the programme and organisational theories of change may not have been sufficiently explicit, and more pre-intervention work might have identified. More precisely, how it would work and under what conditions.

It is also possible that SPI needed a longer time scale or greater intensity to achieve change and for its improvements to show up in the kinds of observations we made. There is some evidence from the qualitative work that the scale of the task was seen as daunting, and that the resource implications and degree of organisational re-gearing that was required had been underestimated.

Changes that may be relatively easy to achieve in highly contained areas (such as the ICU) may have been much more difficult to achieve in other clinical environments, and it may simply take a long time for programmes like SPI to penetrate.

Likewise, the intensity of the intervention may not have been sufficient to engender large-scale change. Spending £775,000 over 18 months in hospitals with annual budgets of £150m–£300m might simply have been too small an amount, especially when little of that money made its way to the sharp end of practice.

The techniques used may have low effectiveness in general use. For example, one element of the IHI approach, the use of FMEA, has recently been challenged.[59–60] It may also be the case that the impact of measures such as walk rounds, safety briefings, and SBAR may be too diffuse to have discernible impacts.

Multiple patient safety initiatives

Lastly, there have been many bottom-up and top-down initiatives over the course of the study.

The NHS has not stood still and English organisations have committed to adopting approaches to patient safety with many similarities to the SPI.[61;62]

For example, as we mentioned earlier, there were multiple initiatives in relation to deteriorating patients. This is reflected in the improvements we have measured across both control and SPI1 hospitals.

However, the hypothesis examined in this study is that SPI would add value to changes that were happening anyway. It is the marginal value of SPI over independent temporal change that is interesting.

The possibility of such temporal effects underscores the need for contemporaneous controls in conducting external, summative evaluations of service delivery interventions.[16] Any changes may otherwise be falsely attributed to the intervention.

4.4 Interpretation

Whether or not the findings reported here are provocative will depend on starting beliefs.

Our results will be disappointing to anyone who thought that the effects would be dramatic. The SPI was introduced as a radical initiative that would have profound effects and which would 'reduce adverse event rates in hospitals by 50%'.[43;44]

Our results suggest that much more temperate claims should be made in future. It must also add to the doubts that have already been expressed about whether the Saving 100,000 Lives campaign was responsible for (all of) the observed reduction in mortality in participating hospitals in the United States.[63] From a Bayesian perspective,[64] enthusiasts should be at least a little less confident about such an intervention demonstrating dramatic effects across hospitals over relatively short periods.

Our results will come as less of a surprise to observers who believe that it is difficult to achieve improvements in quality of care and reduce error rates through generic management initiatives, however, enthusiastically they are welcomed. Creating deep-seated, systemic cultural change through an external initiative with a modest budget over a limited time scale, may be viewed as almost quixotic by more sceptical observers.

We found that the principles and practice of SPI had limited penetration at the medical ward level. The quantitative results are consistent with this finding. More disappointingly, we failed to find an intervention effect on more specific targets, such as monitoring vital signs or medicines reconciliation.

In Bayesian terms, these data are likely to reinforce a neutral prior probability distribution.

Patient safety[65] is hard and achieving change is likely to be a marathon rather than a sprint. Any detectable effects of such interventions may take some time to surface. Their effective implementation requires:

- clarity about the theories of change underlying the programme
- recognition of the scale of resource and organisational support required to make patient safety efforts work
- improved understanding of how, in the face of daunting complexity and multiple priorities, practitioners, middle managers and organisational systems can be better supported.

4.5 Next steps

From the authors' perspective there are two dangers to be avoided. The first danger is to despair and resort to nihilism. The corresponding danger is to privilege positive results over null results.

Objective proof of subjective interpretations is even more difficult to come by in the evaluation of service delivery interventions, than in other branches of science. Yet null results remain valuable; face validity is not enough.

It is important to recognise that hospitals did report effects from SPI participation, including heightened managerial awareness of and commitment to patient safety, and organisational learning about how to implement patient safety improvement efforts in the future.

The intervention did register in the hospitals even if it did not penetrate deeply throughout. The challenge is to build on these observed effects.

The staff we interviewed theorised about the way forward. They proposed offering more support to middle managers, engaging clinical leaders at earlier stages and encouraging clinical ownership as a way of securing success in the future. They suggested reducing the number of areas to be tackled and avoiding areas where there is scientific dispute about whether something is an important problem.

It was clear that hospitals had learned that addressing issues of legitimacy was a key task. They knew that introducing initiatives that generated more paperwork would be unpopular among

stretched ward staff, and that large-scale resourcing and structural support may be needed to implement many patient safety efforts successfully.

The results of the ethnographic sub-study have started to shed light on a fundamental dilemma in many aspects of management. Managers are held accountable for the quality of services. Yet quality is more likely to improve if based on initiatives arising from staff caring for patients.

The task of managers might thus be seen as providing the conditions that might foster bottom-up change and exerting a subtle form of leadership that inspires and empowers. Or perhaps even by making front-line staff feel they were the objects, not subjects of inspiration. Although the SPI intervention clearly intended to achieve this effect, it seems that success in the round was limited.

In contrast, the US Veterans Health Administration's Quality Enhancement Research Initiative[66] is held up as an example of a successful programme that has managed to orchestrate a genuinely bottom-up process. This programme is militantly clinician-based, and built around ideas agreed by clinicians working with managers and researchers in tasks groups. Effort, focus and resources are invested in finding out where specific practice is sub-standard and then tackling the specific causes one by one.

References

1 Department of Health Expert Group (2000). *An organisation with a memory*. London: The Stationery Office.

2 Khon L, Corrigan J, Donaldson M (2000). *To err is human: building a safer health system*. National Academies Press.

3 Leape LL, Berwick DM (2005). Five years after to err is human: what have we learned? *JAMA*, vol 295, pp2,384-2,390.

4 Hutter B (2001). *Regulation and risk: occupational health and safety on the railways*. London: Oxford University Press.

5 Macrae C. Learning from patient safety incidents: Creating participative risk regulation in healthcare (2008). *Health, Risk & Society*, vol 10, pp53-67.

6 Bisognano C, McCannon J, Botwinick L (2005). A campaign for 100,000 lives.

7 Berwick DM, Calkins DR, McCannon CJ, Hackbarth AD (2006). The 100 000 lives campaign: setting a goal and a deadline for improving health care quality. *JAMA* vol 295, pp324-327.

8 Health Foundation. Description of the SPI intervention. 10-9-2009. Personal communication.

9 Dixon-Woods M, Tarrant C, Willars J, Suokas A (2010). How will it work? A qualitative study of strategic stakeholders' accounts of a patient safety initiative. *Qual Saf Health Care*, vol 19, pp74-78.

10 Varkey P, Reller MK, Resar RK (2007). Basics of quality improvement in health care. *Mayo Clinic Proceedings*, vol 82, pp735-739.

11 Deming WE (2000). *Out of the crisis*. 2nd edition ed. MIT Press.

12 Benning A, Dixon-Woods M, Nwulu U, Ghaleb M, Dawson J, Barber N et al (2011). Multiple component patient safety intervention in English hospitals: controlled evaluation of second phase. *BMJ,* vol 342:d199.

13 Brown C, Hofer T, Johal A, Thomson R, Nicholl J, Franklin BD et al (2008). An epistemology of patient safety research: a framework for study design and interpretation. Part 1. Conceptualising and developing interventions. *Qual Saf Health Care,* vol 17, pp158-162.

14 Brown C, Hofer T, Johal A, Thomson R, Nicholl J, Franklin BD et al (2008). An epistemology of patient safety research: a framework for study design and interpretation. Part 2. Study design. *Qual Saf Health Care*, vol 17, pp163-169.

15 Brown C, Hofer T, Johal A, Thomson R, Nicholl J, Franklin BD et al (2008). An epistemology of patient safety research: a framework for study design and interpretation. Part 3. End points and measurement. *Qual Saf Health Care*, vol 17, pp170-177.

16 Brown C, Hofer T, Johal A, Thomson R, Nicholl J, Franklin BD et al (2008). An epistemology of patient safety research: a framework for study design and interpretation. Part 4. One size does not fit all. *Qual Saf Health Care,* vol 17, pp178-181.

17 Brown C, Lilford R. Evaluating service delivery interventions to enhance patient safety (2008). *BMJ,* vol 37, pa2,764.

18 Donabedian A (1980). Explorations in quality assessment and monitoring, In: The definition of quality and approaches to its assessment. Ann Arbor Michigan: Health Administration Press.

19 Cook R, Woods D (1994). Operating at the sharp end: the complexity of human error. In: Bogner M, editor. Human Error in Medicine. Lillsdale NJ: Lawrence Erlbaum.

20 Glaser B, Thomson R (1967). The discovery of grounded theory: Strategies for qualitative research. Chicago: Aldine Publishing Company.

21 West MA, Guthrie JP, Dawson JF, Borrill CS, Carter MR (2006). Reducing patient mortality in hospitals: The role of human resource management. *Journal of Organizational Behavior*, vol 27, pp983-1002.

22 Guzzo RA, Jette D, Katzell RA (1985). The effects of psychologically based intervention programs on worker productivity: a meta-analysis. *Personel Psychology*, vol 38, pp275-292.

23 Borrill CS, West MA, Shapiro D, Rees A (2000). Team working and effectiveness in health. *British Journal of Healthcare Management*, vol 68, pp364-371.

24 Michie S, West M (2004). Managing people and performance: An evidence-based framework applied to health service organisations. *International Journal of Management Reviews*, vol 5, pp91-111.

25 Hillman KM, Bristow PJ, Chey T, Daffurn K, Jacques T, Norman SL et al (2001). Antecedents to hospital deaths. *Intern Med J*, vol 31, pp343-348.

26 Smith GB, Prytherch DR, Schmidt P, Featherstone PI, Knight D, Clements G et al (2006). Hospital-wide physiological surveillance-a new approach to the early identification and management of the sick patient. *Resuscitation*, vol 71, pp19-28.

27 McQuillan P, Pilkington S, Allan A, Taylor B, Short A, Morgan G et al (1998). Confidential inquiry into quality of care before admission to intensive care. *BMJ*, vol 316, pp1,853-1,858.

28 Nwulu U, Benning A, Lilford R (2009). The quality of peri-operative care in England: results of a criterion based case note audit. *In preparation*.

29 British Thoracic Society (2001). BTS Guidelines for the Management of Community Acquired Pneumonia in Adults. *Thorax*, vol 56, ppiv1-64.

30 British Thoracic Society (2004). BTS Guidelines for the Management of Community Acquired Pneumonia in Adults www.brit-thoracic.org.uk/Portals/0/Clinical%20Information/Pneumonia/Guidelines/MACAPrevisedApr04.pdf

31 The Royal Pharmaceutical Society of Great Britain, British Medical Association (2008). *British National Formulary*, 56 ed. Pharmaceutical Press.

32 The Royal Pharmaceutical Society of Great Britain, British Medical Association (2007). *British National Formulary*, 54 ed. BMJ Publishing Group.

33 The Royal Pharmaceutical Society of Great Britain, British Medical Association (2007). *British National Formulary*, 53 ed. Pharmaceutical Press.

34 Dean B, Barber N, Schachter M (2009). What is a prescribing error? *Qual Health Care*, vol 9, pp232-237.

35 Barber N, Franklin BD, Cornford T, Klecun E, Savage I (2006). *Safer, faster, better? Evaluating electronic prescribing*. Report to the Patient Safety Research Programme (Policy Research Programme of the Department of Health). PS019.

36 Cousins D, Hatoum H (1991). The development of a computerised quality assurance system for clinical pharmacy. *Int J Pharmact Pract*, vol 1, pp86-89.

37 Mohammed MA, Lilford RJ, Rudge G, Dawson JF, Deeks JJ, Girling A et al (2008). *Probing variations in hospital standardised mortality ratios in the West Midlands – A study commissioned by the NHS West Midlands Strategic Health Authority*. Birmingham, The University of Birmingham.

38 Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L, Hamilton JD (1995). The Quality in Australian Health Care Study. *Med J Aust*, vol 163, pp458-471.

39 Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG et al (1991). Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *N Engl J Med*, vol 324, pp370-376.

40 Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA et al (1991). The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. *N Engl J Med,* vol 324, pp377-384.

41 Vincent C, Neale G, Woloshynowych M (2001). Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001, vol 322, pp517-519.

42 Woloshynowych M, Neale G, Vincent C (2003). Case record review of adverse events: a new approach. *Qual Saf Health Care* vol 12, pp411-415.

43 Shirley PJ (2008). The Safer Patients Initiative: the UK experience of attempting to improve safe clinical care. *Med J Aust*, vol 189, p414.

44 The Health Foundation (2006). *Safer Patients Initiative trusts explain how they are making hospitals safer for patients*. http:// www health org uk/news/blank_1_43 html [cited 2009 Feb 01].

45 Lilford R, Mohammed MA, Spiegelhalter D, Thomson R (2004). Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *The Lancet*, vol 363, pp1,147-1,154.

46 Hackbarth A, McCannon C, Martin L, Lloyd R, Calkins D (2006). The hard count: calculating lives saved in the 100,000 lives campaign. *ACP Guide for Hospitalists*, April 2006, pp1-5.

47 Healthcare Commission (2005). *NHS national staff survey 2004: Summary of key findings*. London: The Healthcare Commission.

48 Healthcare Commission (2007). *Acute trusts – the views of staff: Key findings from the 2006 survey of NHS staff*. London: The Healthcare Commission.

49 Healthcare Commission (2006). *Making sense of your staff* survey data. London: The Healthcare Commission. http://www.carequalitycommission.org.uk/_db/_documents/Making_sense_of_your_Staff_Survey_Data_200703214701.pdf [cited 2011 Jan 28].

50 Healthcare Commission. NHS National Staff Survey 2006: Key Findings. London: The Stationery Office.

51 Dixon-Woods M, Suokas A, Pitchforth E, Tarrant C (2009). An ethnographic study of classifying and accounting for risk at the sharp end of medical wards. *Social Science & Medicine,* vol 69, pp362-369.

52 Landis J, Koch G (1977). The measurement of observer agreement for categorical data. *Biometrics,* vol 33, pp159-174.

53 Lilford R, Edwards A, Girling A, Hofer T, Di Tanna GL, Petty J et al (2007). Inter-rater reliability of case note audit: a systematic review. *J Health Serv Res Policy*, vol 12, pp173-180.

54 Dean FB, Vincent C, Schachter M, Barber N (2005). The incidence of prescribing errors in hospital inpatients: an overview of the research methods. *Drug Saf*, vol 28, pp891-900.

55 Murphy E, Dingwall R, Greatbatch D, Parker S, Watson P (1998). Qualitative research methods in health technology assessment: a review of the literature. *Health Technol Assess*, vol 2, no 16, ppiii-274.

56 Landon BE, WIlson IB, McInnes K, Landrum MB, Hirschhorn L, Marsden PV et al (2004). Effects of a Quality Improvement Collaborative on the Outcome of Care of Patients with HIV Infection: The EQHIV Study. *Ann Intern Med* 2004, vol 140, pp887-896.

57 Landon BE, Hicks LS, O'Malley AJ, Lieu TA, Keegan T, McNeil BJ et al (2007). Improving the Management of Chronic Disease at Community Health Centers. *N Engl J Med*, vol 356, pp921-934.

58 Pronovost P, Needham D, Berenholtz S, Sinopoli D, Chu H, Cosgrove S et al (2006). An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med*, vol 355, pp2,725-2,732.

59 An assessment of RPN prioritization in a failure modes effects and criticality analysis (2003). Reliability and Maintainability Symposium, 2003.

60 Shebl NA, Franklin BD, Barber N (2009). Is Failure Mode and Effect Analysis Reliable? *Journal of Patient Safety*, vol 5, pp86-94.

61 National Institute for Health and Clinical Excellence, National Patient Safety Agency. *Technical patient safety solutions for medicines reconciliation on admission of adults to hospital*. London: NICE, NPSA.

62 National Patient Safety Agency (2007). *Recognising and responding appropriately to early signs of deterioration in hospitalised patients.* London: NPSA.

63 Wachter R, Pronovost P (2006). The 100,000 lives campaign: a scientific and policy review. *Joint Commission Journal on Quality and Patient Safety*, vol 32, p621.

64 Lilford RJ, Braunholtz D (1996). For Debate: The statistical basis of public policy: a paradigm shift is overdue. *BMJ*, vol 313, pp603-607.

65 Dixon-Woods M (2010). Why is patient safety so hard? A selective review of ethnographic studies. *Journal of Health Services Research and Policy*, vol 15, pp11-16.

66 Stetler CB, McQueen L, Demakis J, Mittman BS (2008). An organizational framework and strategic implementation for system-level change to enhance research-based practice: QUERI Series. *Implement Sci*, vol 3, p30.

67 Francis J, Perlin JB (2006). Improving performance through knowledge translation in the Veterans Health Administration. *J Contin Educ Health Prof*, vol 26, pp63-71.

68 Eccles M, Clapp Z, Grimshaw J, Adams PC, Higgins B, Purves I et al (1996). North of England evidence based guidelines development project: methods of guideline development. *BMJ,* vol 312, pp760-762.

Notes from the authors

Ethics

Ethical approval for the stakeholder interviews, case note review and ethnography was granted by the Trent Multi-centre Research Ethics Committee. Ethical approval already existed for the staff and patient surveys from North West Multi-centre Research Ethics Committee and we submitted a substantial amendment form to include the non-English sites. Local research governance was followed at each site.

Contributors

Amirta Benning, Mary Dixon-Woods, Jeremy Dawson, Nick Barber and Richard Lilford designed the study and submitted the grant proposal. Richard Lilford was chief investigator. Amirta Benning, Nick Barber, Richard Lilford, Maisoon Ghaleb, Bryony Dean Franklin designed the explicit case note review pro forma and methods for the explicit case note review. Amirta Benning, Richard Lilford and Ugochi Nwulu designed the semi-structured holistic case note review pro forma and methods for data extraction. Amirta Benning and Ugochi Nwulu were responsible for the case note review collection. Maisoon Ghaleb and Bryony Dean Franklin conducted the explicit case note review. Martin Carmalt and Thirumalai Naicker conducted the holistic case note review. Ugochi Nwulu and Maisoon Ghaleb designed the case note review database. Gavin Rudge and Amirta Benning created the queries for data extraction. Alan Girling analysed the explicit case note review data. Karla Hemming analysed the holistic case note review data. Mary Dixon-Woods supervised the management survey and ethnographic study. Karla Hemming and Sopna Choudhury performed quantitative analysis of the qualitative data from the stakeholder interviews. Anu Suokas carried out the ethnographic fieldwork. Mary Dixon-Woods and Anu Suokas analysed the qualitative data. Jeremy Dawson was responsible for all aspects of the staff and patient surveys. All authors contributed to the final manuscript.

Acknowledgements

We are grateful to the following for their expertise and contribution in the development of the explicit case note review pro forma:

- Michael Morgan, consultant physician and head of service at University Hospitals of Leicester NHS Trust; honorary professor of respiratory medicine at the University of Leicester; chair of the executive committee, British Thoracic Society
- Martyn Partridge, professor of respiratory medicine at the National Heart and Lung Institute, Imperial College, London, based at Charing Cross Hospital
- Philip Ind, respiratory consultant and honorary senior lecturer at the NHLI, Imperial College, London.

We also thank Janet Willars for conducting stakeholder interviews and for helping to facilitate focus groups.

We would also like to thank Dale Webb, Louise Thomas and Simona Arena for their insight into the implementation of SPI.

Finally, we are grateful to Peter Chilton for his assistance in the preparation of this report.

Funding

The study was funded by the Health Foundation. Additional funding for the explicit case note review in control hospitals and the holistic case note review study was provided by the National Patient Safety Agency. The Centre for Medication Safety and Service Quality is affiliated with the Centre for Patient Safety and Service Quality at Imperial College Healthcare NHS Trust, which is funded by the National Institute of Health Research.

Competing interests

None declared.

Appendix 1

Staff survey – 11 questions identified as being relevant to the SPI

Six of these 11 scores are straightforward percentages:

1. **Percentage of staff having well structured appraisals** reflects the percentage of respondents who not only say that they had received an appraisal in the previous 12 months, but that this appraisal helped them improve how to do their job, helped agree clear objectives for their work, and left them feeling that their work was valued by their organisation. These aspects of appraisal have been shown to be particularly important for organisational outcomes in many sectors, including healthcare.[21;22]

2. **Percentage of staff working in well-structured teams** is the percentage of respondents who said they worked in teams, that their teams had clear objectives, that they had to work closely with team members to achieve these objectives, and that the team met regularly to discuss their effectiveness and how it could be improved. These are features of team working that have been shown to be critical for achieving high-quality team outcomes.[23]

3. **Percentage of staff witnessing potentially harmful errors or near misses in previous month** was the percentage of respondents who said they had witnessed an error or a near miss in the previous month that could have harmed either patients or staff.

4. **Percentage of staff suffering work-related injury** is the percentage of respondents who said they had suffered injury or illness as a result of moving or handling; needlestick or sharps injuries; slips, trips or falls; or exposure to dangerous substances in the previous 12 months.

5. **Percentage of staff suffering work-related stress** is the percentage of respondents who said they had suffered injury or illness as a result of work-related stress in the previous 12 months.

6. **Percentage of staff experiencing physical violence from patients/relatives** was the percentage of respondents who said they had personally experienced physical violence at work from either patients, or relatives of patients, in the previous 12 months.

The other five scores were calculated as the mean of a number of separate questionnaire items, each scored from one to five representing answers from 'strongly disagree' through to 'strongly agree', or from 'very dissatisfied' to 'very satisfied':

7. **Intention to leave** shows the extent to which employees are considering leaving their jobs. It is based on three questionnaire items.

8. **Staff job satisfaction** is a measure of employees' overall satisfaction with their jobs, and is based on seven items.

9. **Quality of work-life balance** measures the support provided by organisations for employees to maintain a good work-life balance, and is based on three items.

10. **Support from supervisors** is a measure of the extent to which employees feel supported by their immediate managers at work, and is based on five items.

11. **Organisational climate** is a measure of the overall climate, or positive feeling, within the organisation, including factors such as trust in management, communication, staff involvement in decision making, and emphasis on quality. This is based on six items. Each of these scores has been shown to relate to performance outcomes, including quality of care, in healthcare organisations.[24]

Two further scores, 'Availability of Hand Washing Materials' and 'Fairness and Effectiveness of Incident Reporting Procedures' would have been considered, except that comparable data were not available at both survey points owing to changes in the Care Quality Commission questionnaire. A complete description of the method of calculation, and the wording of all survey items, is available from the Care Quality Commission.[49]

Appendix 2

Prescribing errors – stages of the drug use process and their definition

Need for a drug includes the following:

- Omission of drug
  Any situation in which a drug is not prescribed for a clinical condition for which a drug is indicated; this includes the erroneous omission of drugs from an inpatient drug chart or discharge prescription. Also included is the premature discontinuation of a prescribed medication
- Drug no longer needed
  Continuation of a prescribed drug for a longer duration than necessary.
- No indication for drug prescribed
  Prescription of a drug without a corresponding indication.
- Duplication of therapy
  Prescription of two or more drugs with the same therapeutic action when only one of the drugs is necessary, or the prescription of the same drug more than once.

Selection of drug includes the following:

- Prescription of drug to which patient has significant allergy
  This would include the prescription of penicillins in a patient with a confirmed penicillin allergy and the prescription of NSAIDs in an asthmatic patient who is hypersensitive to drugs of this class.
- Prescription of drug that is contra-indicated due to drug interaction
  This includes the prescription of buprenorphine in a patient receiving other opiates, and the prescription of drugs which interact with anti-retrovirals.

– Prescription of drug to which patient has clinical contra-indication
  Prescription of drugs that are contra-indicated due to pre-existing medical conditions such as diabetes, severe renal impairment or liver disease.
– Prescription of drug that was not intended.
  Any situation in which the drug prescribed was not that desired. This includes errors in medication history taking and transcription errors when rewriting drug charts or discharge prescriptions, as well as inappropriate clinical decisions.

Selection of dose includes the following:
– Failure to specify maximum dose
  Failure to specify the maximum dose for a drug prescribed to be given as required.
– Failure to take into account drug interaction
  The prescription of a drug in a dose that is not appropriate because of a concurrent drug interaction.
– Dose/rate mismatch
  Prescription of a drug to be infused on a milligram/kilogram/hour basis, where the millilitre/hour rate calculated does not correspond to the dose required.
– Total daily dose divided incorrectly
  Any situation in which the total daily dose is correct, but is divided into an incorrect number of daily doses. For example, cyclizine prescribed 150mg once daily instead of 50mg three times a day.
– Overdose
  Any situation in which the patient is prescribed too high a dose of a drug, that is not covered by the situations described above.
– Underdose
  Any situation in which the patient is prescribed too low a dose of a drug, that is not covered by the situations described above.
– Failure to specify the strength of formulation
  The prescription of a drug where there is more than one strength for one formulation and not specifying the strength intended for the prescription

Selection of formulation:

Prescription of the wrong formulation for the drug and dose regimen prescribed.

Provide information for supply includes the following:

- Product or formulation not specified
  Any situation in which the product or formulation is not specified in enough detail for a supply to be made. This includes failure to adequately specify the product formulation intended and the prescription of illegible or otherwise ambiguous medication orders.
- Strength or dose not specified
  Any situation in which the strength or dose of a preparation is not specified in sufficient detail for the appropriate product to be supplied.
- Route not specified
  Failure to state the route of administration for a drug that can be given by more than one route.
- Prescription not signed
  An inpatient or discharge prescription that has not been signed by the prescriber.
- Controlled drugs prescription requirements
  Failure to write a discharge prescription according to the controlled drugs requirements.

Appendix 3

Patient survey – five identified scores relevant to SPI

Each of these was scored between 0 and 100. The three satisfaction scores were:

1. **Overall, how would you rate the care you received?**
   (five possible responses: Excellent = 100, Very good = 75, Good = 50, Fair = 25 and Poor = 0)
2. **How would you rate how well the doctors and nurses worked together?**
   (five possible responses: Excellent = 100, Very good = 75, Good = 50, Fair = 25 and Poor = 0)
3. **Overall, did you feel you were treated with respect and dignity while you were in the hospital?**
   (Yes, always = 100; Yes, sometimes = 50; and No = 0).

The two scores related to cleanliness were:

4. **In your opinion, how clean was the hospital room or ward that you were in?**
   (possible responses: Very clean = 100, Fairly clean = 67, Not very clean = 33, and Not at all clean = 0)
5. **How clean were the toilets and bathrooms that you used in hospital?**
   (same response options, plus 'I did not use a toilet or bathroom', which was excluded from the analysis)

The Health Foundation is an independent charity working to continuously improve the quality of healthcare in the UK.

We want the UK to have a healthcare system of the highest possible quality – safe, effective, person-centred, timely, efficient and equitable. We believe to achieve this, health services need to continually improve the way they work.

We are here to inspire and create the space for people, teams, organisations and systems to make lasting improvements to health services.

Working at every level of the healthcare system, we aim to develop the technical skills, leadership, capacity, knowledge, and the will for change, that are essential for real and lasting improvement.

The Health Foundation
90 Long Acre
London WC2E 9RA
T 020 7257 8000
F 020 7257 8001
E info@health.org.uk

Registered charity number: 286967
Registered company number: 1714937

www.health.org.uk

ISBN 978-1-906461-21-8