# EVALUATING INTERVENTIONS TO MAKE HEALTHCARE SAFER: METHODOLOGICAL ANALYSIS AND CASE STUDY

**By**

**AMIRTA BENNING BSc (Hons), MSc**

**DOCTOR OF PHILOSOPHY**
**A thesis submitted to**
**The University of Birmingham**
**for the degree of DOCTOR OF PHILOSOPHY**

**Public Health, Epidemiology and Biostatistics**
**College of Medicine and Dental Sciences**
**The University of Birmingham**

**April 2011**

# ABSTRACT

This thesis describes study designs for the robust evaluation of complex patient safety interventions. Fundamentally, the study designs available to measure the effectiveness of patient safety interventions fall into two categories – those that use contemporaneous controls, and those that do not. A review of the recent literature (245 citations) revealed that most studies were single-centre (63%), and the majority of these did not use contemporaneous controls (84%); whilst in multi-centre studies (37%) the number of studies using contemporaneous controls (49%) equalled the number of studies that did not (51%). Studies that do not use contemporaneous controls dominate the literature, but they are inherently weak and subject to bias.

Furthermore, this thesis discussed a case-study for the evaluation of a highly complex patient safety intervention – the *Safer Patients Initiative* (**SPI**), which sought to generically strengthen hospitals, whilst improving frontline activities. The evaluation was a before and after study, with contemporaneous controls. It used mixed-methods, so that the triangulation of one type of research finding could be reinforced when corroborated by the finding of another type. Uniquely, it also, compared the rates of change across control and **SPI** hospitals – an approach referred to as the "difference-in-difference" method.

# DEDICATION

To my Mum and Dad – Barinder Kaur and Dalbir Singh Johal. Thank you for your never ending love, support and faith. This is for you.

# ACKNOWLEDGEMENTS

The evaluation was funded by the Health Foundation and the National Patient Safety Agency. The epistemological work forms part of the output of a Cross-Council Research Network in Patient Safety Research funded by the Medical Research Council (Reference G0300370).

I am firstly grateful to my colleague and principle supervisor, Richard Lilford for his support, guidance and sesquipedalism. You've finally turned me into an academic of sorts.

I would also like to thank Mohammed A Mohammed, my second supervisor who steered me in the development of this thesis.

I am eternally grateful to "my" team: Ugochi Nwulu, Sopna Choudhry, Leela Prabu, Christine Kambwa and Peter Chilton (honorary member) and my collaborators in making this evaluation happen.

I would also like to thank all the members of 'Team Lilford', past and present, who have supported and encouraged me at various stages throughout this work.

I would also like to acknowledge and thank the following individuals:

Sheldon Stone and Chris Fuller for access to data collected as part of the NOSEC study; and David Harrison and Kathy Rowan from ICNARC for access to data collected as part of the Case Mix Programme.

Dale Webb, Gillian Hastings Louise Thomas and Simona Arena at the Health Foundation for their insight of the intervention.

Last, but not least I want to thank my family (old and new) and friends, especially my brother Pardeep, my sister Sharon, and my husband Raj (for his grammatical superiority), for their support and encouragement.

# CONTRIBUTIONS

A complex study, involving twelve sub-studies over seven years is never the endeavour of an individual. The evaluation which makes up the majority of this thesis was awarded to the University of Birmingham in 2004, after competitive tender. I am co-investigator for the evaluation. I designed the study in collaboration with Mary Dixon-Woods (Professor of Medical Sociology, University of Leicester); Jeremy Dawson (Research Fellow, Aston University); and Nick Barber (Professor of the Practice of Pharmacy, The School of Pharmacy, University of London). Richard Lilford (Professor of Clinical Epidemiology, University of Birmingham) was chief investigator.

In addition to the successfully bidding for the original funds for the evaluation, I was also instrumental in the negotiation of funding for control sites.

I was responsible for the overall management of the evaluation, recruitment of study hospitals, completion of local R&D and ethics and collection of quantitative data. I was also responsible for a case-note "clearing house" at the University of Birmingham, including recruitment of junior staff and the holistic reviewers.

I will now detail the contributions made for each chapter of this thesis:

**Chapter 1**

This chapter provides a brief introduction to the thesis and is written solely by me.

**Chapter 2**

This is a literature review that I conducted to find out which study designs, in the recent literature, have been used to evaluate complex patient safety interventions.

**Chapter 3**

This is a description of the intervention that I developed from observations and material collected at teaching events for the intervention hospitals. These were hosted by the funders and the designers of the intervention.

**Chapter 4**

The methods refer to a series of four epistemological papers for which I am co-author. Richard Lilford conceived the Cross-Council Research Network in Patient Safety Research Methodology and formulated the first draft of the report and the papers with my assistance. Celia Brown, (Senior Research Fellow, University of Birmingham) contributed to subsequent drafts of the report and papers).

Bryony Dean Franklin (Professor of Medication Safety, The School of Pharmacy, University of London, Director, Centre for Medication Safety and Service Quality, Imperial College Healthcare NHS Trust); Jon Nicholl (Director Medical Care Research Unit Policy Research

Programme, Deputy Dean and Director of Research, School of Health and Related Research, University of Sheffield); Richard Thomson (Professor of Epidemiology & Public Health, Newcastle University); and Timothy P. Hofer (Research Scientist, Veterans Affairs, Ann Arbor, Health Services Research and Development Service Center of Excellence, University of Michigan) contributed to the research network and provided comments on drafts of the report and papers in their areas of expertise.

The framework for the evaluation is an application of the epistemology papers and was used in the grant application for the award. All collaborators had an input into this framework. The selection of controls was made by Jeremy Dawson and I carried out the actual recruitment of the study hospitals into the evaluation.

**Chapter 5**

I negotiated for the staff surveys to be conducted in non-English hospitals, as data for the English hospitals was already available to the research team. Jeremy Dawson conducted the analysis of all staff survey data.

**Chapters 6 and 11**

Along with Nick Barber, Richard Lilford, Maisoon Ghaleb (Lecturer in Pharmacy Practice/Patient Safety, University of Hertfordshire) and Bryony Dean Franklin, I designed the explicit case-note review data collection tool. This was shared with the following experts for further comment and opinion: Michael D.L. Morgan (Consultant Physician and Head of Service, Department of Respiratory Medicine, Allergy and Thoracic Surgery, Glenfield Hospital, University Hospitals of Leicester, Honorary Professor of Respiratory Medicine,

University of Leicester, Chair of Executive Committee, The British Thoracic Society), Martyn R Partridge (Professor of Respiratory Medicine, Respiratory Health Services Research Group Faculty of Medicine, Imperial College, NHLI at Charing Cross Hospital) and Philip W Ind (Respiratory Consultant, Honorary Senior Lecturer, National Heart and Lung Institute, Imperial College).

Maisoon Ghaleb and Bryony Dean Franklin conducted the explicit acute medicine case-note review. Together with Gavin Rudge (Data Scientist, University of Birmingham) I designed the database queries for final analyses and to assess the learning effect on the case-note reviewers. Alan Girling (Senior Research Fellow, University of Birmingham) carried out the statistical analysis of this sub-study.

**Chapters 7 and 10**

The holistic case-note review was my idea. Richard Lilford and I designed the semi-structured holistic case-note review pro forma and methods for data extraction. I recruited Martin Carmalt (Consultant Physician, Royal Orthopaedic Hospital, Birmingham) and Thirumalai Naicker (Honorary Research Associate, University of Birmingham) to conduct the holistic case-note review. Martin Carmalt and M Clare Derrington (Public Health Physician, Independent Contractor) carried out a separate review of deaths. Karla Heming performed the analysis of the holistic case-note review data that I had assembled.

**Chapter 8**

Ugochi Nwulu (Research Associate, The University of Birmingham), Richard Lilford and I designed the peri-operative case-note review data extraction tool. This was shared with Dion Morton (Professor of Surgery, University Hospitals Birmingham) and David Thomas (Consultant Anaesthetist, Heart of England NHS Foundation Trust) to elicit their expertise and opinion.

Ugochi Nwulu and Amit Kotecha (Surgical Registrar, Royal Orthopaedic Hospital, Birmingham) conducted the surgical case-note review.

Together with Gavin Rudge I designed the database queries for final analyses and Alan Girling carried out the statistical analysis of this sub-study.

**Chapter 9**

I negotiated access and collated the hand hygiene data. Karla Hemming conducted the analysis.

**Chapter 12**

I negotiated access to intensive care outcome data and prepared it for Karla Hemming conducted the analyses.

**Chapter 13**

I negotiated access and collated infection related data. Karla Hemming conducted the analyses of these.

**Chapter 14**

The patient surveys were my idea. I negotiated for the patient surveys to be conducted in non-English hospitals. Jeremy Dawson had access to data for English hospitals. Jeremy Dawson conducted the analysis of all the patient survey data.

**Chapter 15**

All authors contributed to the final reports and hence the discussion which is featured in this chapter. It should be noted that the discussion at the end of each chapter are my thoughts and comments, as is the discussion on lessons for the study design of future evaluations.

## Appendix B

The strategic stakeholder interviews were led by Mary Dixon-Woods and they were conducted by Janet Willars (Honorary Visiting Fellow, University of Leicester) and Anu Soukas (PhD student, University of Leicester). This sub-study has been published separately: Dixon-Woods M, Tarrant, Willars J, Suokas. *How will it work? A qualitative study of strategic stakeholders' accounts of a patient safety initiative*. Qual Saf Health 2010; 19: 74-78. It is included in this thesis for completeness and for purposes of "triangulation" as discussed in Chapter 15.

I arranged for the quantitative analysis of the qualitative data and Karla Heming (Senior Research Fellow, University of Birmingham) conducted the analysis of this.

## Appendix C

The ethnography was led by Mary Dixon-Woods. Anu Soukas conducted the observations and interviews on hospital wards. She also conducted the focus groups with the assistance of Janet Willars and Ugochi Nwulu. This is again included for purposes of "triangulation" (Chapter 15).

# PUBLICATIONS

The following publications have emanated from this research:


**Benning A,** Dixon-Woods M, Nwulu U, Ghaleb M, Dawson J, Barber N, Franklin BD, Girling A, Hemming K, Carmalt M, Rudge G, Naicker T, Kotecha A, Derrington MC, Lilford R. Multiple component patient safety intervention in English hospitals: controlled evaluation of second phase. *BMJ*. 2011; 342:d199.

**Benning A,** Ghaleb M, Suokas A, Dixon-Woods M, Dawson J, Barber N et al. Mixed method evaluation of a large-scale organisational intervention to improve patient safety in four UK hospitals: Health Foundation; February 2011.

**Benning A,** Dixon-Woods M, Nwulu U, Ghaleb M, Dawson J, Barber N, Franklin BD, Girling A, Hemming K, Carmalt M, Rudge G, Naicker T, Kotecha A, Derrington MC, Lilford R. Multiple component patient safety intervention in English hospitals: controlled evaluation of second phase. *BMJ*. 2011; 342:d199.

**Benning A,** Nwulu U, Ghaleb M, Dawson J, Barber, N, Franklin B et al. A controlled evaluation of the second phase of a of a multiple-component intervention implemented in English hospitals. Health Foundation; February 2011.

Brown C, Hofer T, **Johal A,** Thomson R, Nicholl J, Franklin BD et al., An epistemology of patient safety research: a framework for study design and interpretation. Part 1. Conceptualising and developing interventions. *Qual Saf Health Care*. 2008;17:158-62.

Brown C, Hofer T, **Johal A,** Thomson R, Nicholl J, Franklin BD et al., An epistemology of patient safety research: a framework for study design and interpretation. Part 2. Study design. *Qual Saf Health Care*. 2008;17:163-9.

Brown C, Hofer T, **Johal A,** Thomson R, Nicholl J, Franklin BD et al., An epistemology of patient safety research: a framework for study design and interpretation. Part 3. End points and measurement. *Qual Saf Health Care*. 2008;17:170-7.

Brown C, Hofer T, **Johal A,** Thomson R, Nicholl J, Franklin BD et al., 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*. 2008;17:178-81.

# TABLE OF CONTENTS

**EXECUTIVE SUMMARY..... 1**

**CHAPTER 1: INTRODUCTION..... 8**

**CHAPTER 2: LITERATURE REVIEW ..... 15**

**2.1 Introduction ..... 15**

2.1.1 Study design..... 15

2.1.2 Before and after studies ..... 17

2.1.2.1 Time series and Statistical Process Control (SPC) ..... 18

2.1.3 Studies with contemporaneous controls ..... 19

2.1.3.1 The unit of comparison ..... 20

2.1.3.2 Design elements of studies using contemporaneous controls ..... 22

2.1.4 The importance of context ..... 24

2.1.5 Literature review ..... 25

**2.2 Methods..... 26**

2.2.1 Identification of studies..... 26

2.2.2 Exclusion criteria..... 26

2.2.3 Data extraction ..... 27

**2.3 Results..... 28**

2.3.1 Study flow..... 28

2.3.2 Study designs used ..... 29

2.3.2.1 Single-centre studies ..... 29
2.3.2.2 Multi-centre studies ..... 29

**2.4 Discussion ..... 32**

2.4.1 Study designs used ..... 32
2.4.1.1 Single-centre studies ..... 32
2.4.1.2 Multi-centre studies ..... 33

2.4.2 Limitations of the literature review ..... 34

2.4.3 Implications for evaluating complex patient safety interventions ..... 35

**CHAPTER 3: CASE-STUDY - THE SAFER PATIENTS INITIATIVE ..... 38**

**3.1 Introduction ..... 38**

**3.2 Selection of participating $SPI_1$ hospitals ..... 40**

**3.3 The change package ..... 42**

**3.4 A model of the healthcare system ..... 44**

3.4.1 Framework for the SPI ..... 45

**3.5 The Safer Patients Initiative – format and structure ..... 49**

3.5.1 Performing a safety self-assessment ..... 52
3.5.1.1 Mortality diagnostic case-note review ..... 52
3.5.1.2 Adverse event trigger tool case-note review ..... 52
3.5.1.3 Medication safety self assessment survey ..... 53
3.5.1.4 HCAI rates ..... 53

3.5.2 Identifying best practice ..... 53

3.5.3 Getting connected to resources ..... 54

3.5.4 Preparing a storyboard ..... 54

3.5.5 Participants ..... 54

3.5.6 Kick-off meeting and learning sessions 2, 3 and 4 ..... 56

**3.6 The Safer Patients Initiative – the content ..... 57**

3.6.1 Enhancing leadership and other generic components ..... 58

3.6.1.1 WalkRounds ..... 59

3.6.1.2 Culture ..... 59

3.6.1.3 Leadership and measurement ..... 60

3.6.1.4 Dissemination ..... 61

3.6.2 Improving specific front-line interventions ..... 63

3.6.2.1 Identifying and responding to deteriorating patients ..... 63

3.6.2.2 Effective communication of patient safety concerns ..... 65

3.6.2.2.1 Safety briefings ..... 65

3.6.2.2.2 Situation Background Assessment Recommendation (SBAR) ..... 66

3.6.2.3 Reducing medication error ..... 67

3.6.2.3.1 Medicines reconciliation ..... 67

3.6.2.3.2 Anticoagulants ..... 69

3.6.2.4 Infection control and hand hygiene ..... 70

3.6.2.5 Peri-operative ..... 71

3.6.2.6 Critical Care Systems ..... 71

3.6.3 Tools and skills developed ..... 74

3.6.3.1 Plan-Do-Study-Act (PDSA) cycles ..... 74

3.6.3.2 Statistical Process Control (SPC) ..... 77

3.6.3.3 Reliable design ..... 79

3.6.3.4 Failure Mode and Effects Analysis (FMEA) ..... 80

**3.7 The $SPI_2$ intervention phase ..... 82**

3.7.1 Selection of participating $SPI_2$ hospitals ..... 83

**3.8 Summary ..... 84**

**CHAPTER 4: OVERVIEW OF THE EVALUATION ..... 87**

**4.1 End points and measurements ..... 87**

4.1.1 Patient outcomes ..... 88

4.1.1.1 Signal to noise ratio ..... 88

4.1.1.2 Minimising bias ..... 89

4.1.2 Surrogate end points ..... 90

4.1.2.1 Fidelity ..... 91

4.1.2.2 Intervening variables ..... 91

4.1.2.3 Clinical process measurement: error ..... 92

4.1.2.3.1 What is an error? ..... 92

4.1.2.3.2 Reporting systems ..... 93

4.1.2.3.3 Trigger tools ..... 94

4.1.2.3.4 Error rates ..... 95

4.1.3 Masking ..... 101

**4.2 Framework for the evaluation ..... 102**

**4.3 Controls and intervention hospitals ..... 109**

**CHAPTER 5: STAFF SURVEYS ..... 113**

**5.1 Introduction ..... 113**

**5.2 Methods ..... 113**

5.2.1 Statistical methods ..... 115

**5.3 Results ..... 117**

5.3.1 Results of the $SPI_1$ evaluation ..... 117

5.3.2 Results of the $SPI_2$ evaluation ..... 118

**5.4 Discussion ..... 128**

**CHAPTER 6: EXPLICIT CASE-NOTE REVIEW TO MEASURE ERROR RATES/QUALITY OF CARE IN PATIENTS WITH ACUTE RESPIRATORY DISEASE ..... 129**

**6.1 Introduction ..... 129**

**6.2 Methods ..... 129**

6.2.1 Case-note selection criteria ..... 130

6.2.2 Case-note assembly ..... 131

6.2.2.1 Case-notes collected for the $SPI_1$ evaluation ..... 131

6.2.2.2 Case-notes collected for the $SPI_2$ evaluation ..... 132

6.2.3 Sample size ..... 132

6.2.4 Eligibility criteria ..... 134

6.2.5 Explicit case-note review ..... 135

6.2.6 Statistical methods ..... 140

6.2.6.1 Statistical methods used specifically in the $SPI_1$ evaluation ..... 141

6.2.6.2 Statistical methods used specifically in the $SPI_2$ evaluation ..... 141

**6.3 Results ..... 142**

6.3.1 The sample ..... 142

6.3.1.1 Number of case-notes reviewed within the explicit review in the $SPI_1$ evaluation 142

6.3.1.2 Number of case-notes reviewed within the explicit review in the $SPI_2$ evaluation 142

6.3.2 Effects associated with the review process ..... 142

6.3.3 Reliability and anonymisation of case-notes within the $SPI_1$ evaluation ..... 144

6.3.4 Quality of medical history taking ..... 144

6.3.4.1 Within the $SPI_1$ evaluation ..... 144

6.3.4.2 Within the $SPI_2$ evaluation ..... 149

6.3.5 Observations and signs of patient deterioration ..... 152
6.3.5.1 Within the $SPI_1$ evaluation ..... 152
6.3.5.2 Within the $SPI_2$ evaluation ..... 155
6.3.6 Appropriate clinical response to abnormal vital signs ..... 160
6.3.6.1 Within the $SPI_1$ evaluation ..... 160
6.3.6.2 Within the $SPI_2$ evaluation ..... 165
6.3.7 Other features of good care for specific classes of patients ..... 169
6.3.7.1 Within the $SPI_1$ evaluation ..... 169
6.3.7.2 Within the $SPI_2$ evaluation ..... 173
6.3.8 Prescribing error ..... 176
6.3.8.1 Within the $SPI_1$ evaluation ..... 176
6.3.8.2 Within the $SPI_2$ evaluation ..... 181
6.3.9 Anti-coagulant prescribing errors ..... 184
6.3.9.1 Within the $SPI_1$ evaluation ..... 184
6.3.9.2 Within the $SPI_2$ evaluation ..... 184
6.3.10 Medicines reconciliation errors at admission ..... 187
6.3.10.1 Within the $SPI_1$ evaluation ..... 187
6.3.10.2 Within the $SPI_2$ evaluation ..... 187
**6.4 Discussion ..... 190**
**CHAPTER 7: HOLISTIC CASE-NOTE REVIEW TO MEASURE ERROR RATES/QUALITY OF CARE IN PATIENTS WITH ACUTE RESPIRATORY DISEASE ..... 193**
**7.1 Introduction ..... 193**
**7.2 Methods ..... 193**
7.2.1 Statistical methods ..... 195

7.2.1.1 Statistical methods used specifically in the $SPI_1$ evaluation ..... 196
7.2.1.2 Statistical methods used specifically in the $SPI_2$ evaluation ..... 196

**7.3 Results ..... 197**

7.3.1 The sample ..... 197
7.3.1.1 Number of case-notes reviewed within the holistic review in the $SPI_1$ evaluation 197
7.3.1.2 Number of case-notes reviewed within the holistic review in the $SPI_2$ evaluation 197

7.3.2 Reliability ..... 198
7.3.2.1 Agreement between reviewers within the $SPI_1$ evaluation ..... 198
7.3.2.2 Agreement between reviewers within the $SPI_2$ evaluation ..... 198

7.3.3 Quality of care ratings ..... 199
7.3.3.1 Quality of care within the $SPI_1$ evaluation ..... 199
7.3.3.2 Quality of care within the $SPI_2$ evaluation ..... 200

7.3.4 Errors ..... 200
7.3.4.1 Error rates within the $SPI_1$ evaluation ..... 200
7.3.4.2 Error rates with the $SPI_2$ evaluation ..... 201

**7.4 Discussion ..... 207**

**CHAPTER 8: ERROR RATES AND QUALITY OF PERI-OPERATIVE CARE ..... 209**

**8.1 Introduction ..... 209**

**8.2 Methods ..... 209**

8.2.1 Case-note selection ..... 210

8.2.2 Case-note assembly ..... 211

8.2.3 Sample size calculation ..... 212

**8.3 Results ..... 214**

8.3.1 Sample, reviewer reliability and headline message ..... 214

8.3.2 Pain relief ..... 214

8.3.3 Prophylactic antibiotics ..... 215

8.3.4 Temperature monitoring ..... 215

8.3.5 DVT prophylaxis ..... 215

**8.4 Discussion ..... 218**

**CHAPTER 9: INDIRECT MEASURE OF HAND HYGIENE ..... 220**

**9.1 Introduction ..... 220**

**9.2 Methods ..... 220**

9.2.1 Statistical methods ..... 221

**9.3 Results ..... 222**

9.3.1 Data available ..... 222

9.3.2 Soap and alcohol hand rub consumption ..... 223

**9.4 Discussion ..... 225**

**CHAPTER 10: ADVERSE EVENTS DETECTED IN ACUTE MEDICAL CASE-NOTES ..... 227**

**10.1 Introduction ..... 227**

**10.2 Methods ..... 227**

10.2.1 Additional review of deaths within the $\mathrm{SPI}_2$ evaluation ..... 228

10.2.2 Statistical methods ..... 228

**10.3 Results ..... 228**

10.3.1 Findings of the $SPI_1$ evaluation ..... 228

10.3.2 Findings of the $SPI_2$ evaluation ..... 232
10.3.2.1 Additional review of deaths within the $SPI_2$ evaluation ..... 232
**10.4 Discussion ..... 236**
**CHAPTER 11: RATES OF MORTALITY AMONG ACUTE MEDICAL CARE PATIENTS ..... 238**
**11.1 Introduction ..... 238**
**11.2 Methods ..... 238**
**11.3 Results ..... 238**
11.3.1 Mortality within the $SPI_1$ evaluation ..... 238
11.3.2 Mortality within the $SPI_2$ evaluation ..... 241
**11.4 Discussion ..... 243**
**CHAPTER 12: INTENSIVE CARE UNIT (ICU): MORTALITY, MORBIDITY AND LENGTH OF STAY ..... 244**
**12.1 Introduction ..... 244**
**12.2 Methods ..... 244**
12.2.1 Statistical methods ..... 246
**12.3 Results ..... 247**
12.3.1 Data available ..... 247
12.3.2 Observed to expected mortality ..... 248
12.3.3 Median length of stay ..... 248
12.3.4 APACHE II and ICNARC risk prediction scores ..... 249

**12.4 Discussion** ..... **251**

**CHAPTER 13: C. DIFF AND MRSA INFECTION RATES** ..... **252**

**13.1 Introduction** ..... **252**

**13.2 Methods** ..... **252**

13.2.1 Statistical methods ..... 253

**13.3 Results** ..... **254**

13.3.1 Data ..... 254

13.3.2 C. diff ..... 255

13.3.3 MRSA ..... 255

**13.4 Discussion** ..... **259**

**CHAPTER 14: PATIENT SURVEY** ..... **260**

**14.1 Introduction** ..... **260**

**14.2 Methods** ..... **260**

14.2.1 Statistical methods used in the $\mathrm{SPI}_1$ evaluation ..... 261

14.2.2 Statistical methods used in the $\mathrm{SPI}_2$ evaluation ..... 262

**14.3 Results** ..... **262**

14.3.1 Findings of the patient survey in the $\mathrm{SPI}_1$ evaluation ..... 262

14.3.2 Findings of the patient survey in the $\mathrm{SPI}_2$ evaluation ..... 266

**14.4 Discussion** ..... **268**

**CHAPTER 15: OVERALL DISCUSSION** ..... **270**

**15.1 Main findings of the evaluation ..... 270**

15.1.1 Participant experience ..... 271

15.1.2 Controls *versus* SPI ..... 272

15.1.3 Interpretation ..... 273

15.1.4 Theory building ..... 280

15.1.5 Next steps ..... 281

**15.2 Lessons for the study design of future evaluations ..... 283**

15.2.1 Strengths and weaknesses ..... 283

**15.3 Conclusion ..... 287**

**REFERENCE LIST ..... 289**

**APPENDICES ..... 308**

**Appendix A: Results of the literature review ..... 309**

**Appendix B: Strategic stakeholder interviews ..... 373**

**Appendix C: Ethnographic study ..... 386**

**Appendix D: Staff survey questions ..... 402**

**Appendix E: Prescribing errors - stage of drug process and their definition ..... 404**

**Appendix F: Difference at baseline for quality ratings, errors and adverse events ..... 407**

**Appendix G: Details of fitted models ..... 409**

**Appendix H: Patient survey questions ..... 411**

# LIST OF FIGURES

Figure 2.1: Basic study designs. Taken from Brown et al., (2008b) ..... 17

Figure 2.2: Clustered sample size calculations for given ICC values (box). The calculations are based on detecting a difference in error rate between control (0.1) and intervention (0.05), with alpha = 5% and power = 80%. Taken from Brown et al., (2008b). ..... 21

Figure 2.3: An overview of the literature search process ..... 28

Figure 3.1: The Health Foundation's SPI Patient Safety Change Package. Taken from the kick off meeting (Institute for Healthcare Improvement and Health Foundation, 2005a) ..... 43

Figure 3.2: A model of the healthcare system ..... 44

Figure 3.3: Embellished Donabedian's causal chain (blue) linking SPI (green) to structure, process and outcomes ..... 46

Figure 3.4: The Phase $SPI_1$ time-line (adapted from IHI learning materials) (Institute for Healthcare Improvement and Health Foundation, 2005a) ..... 51

Figure 3.5: SBAR tool (taken from (Marshall et al., 2009) ..... 67

Figure 3.6: A summary of the hospitals participating in the $SPI_1$ and $SPI_2$ phases of the intervention ..... 85

Figure 4.1: Intellectual framework to classify links between care and poor outcome ..... 89

Figure 4.2: Causal chain linking SPI to outcomes. The evaluation sub-studies were made at points across the chain to provide information on context, fidelity, and effectiveness of SPI ..... 104

Figure 4.3: Schematic diagram demonstrating staggering of the data collection enabling the use of pre-$SPI_2$ intervention data as control data for the evaluation of $SPI_1$ ..... 112

Figure 5.1: An outline of the staff survey sub-study ..... 114

Figure 6.1: An outline of the case-note review of patients admitted with acute respiratory disease ..... 130

Figure 6.2: Prescription error rates and date of review ..... 143

Figure 8.1: An outline of the case-note review of surgical patients ..... 209

Figure 8.2: Sample sizes for 80% power (at 5% significance)..... 213

Figure 9.1: An outline of the indirect measure of hand hygiene sub-study ..... 221

Figure 9.2: Rate of soap consumption per 1,000 bed days over time in control and $SPI_2$ hospitals..... 224

Figure 9.3: Rate of alcohol hand rub (AHR) consumption per 1,000 bed days over time in control and $SPI_2$ hospitals ..... 225

Figure 12.1: An outline of the ICU outcomes sub-study ..... 246

Figure 13.1: An outline of the sub-study to measure HCAI rates ..... 253

Figure 13.2: Rate of Clostridium difficile (C. diff) cases per 1,000 bed days in control and $SPI_2$ hospitals..... 257

Figure 13.3: Rate of Methicillin-resistant Staphylococcus aureus (MRSA) cases per 100,000 bed days in control and $SPI_2$ hospitals..... 258

Figure 14.1: A summary of the patient survey sub-study ..... 261

# LIST OF TABLES

Table 2.1: Controlled study design matrix.....24

Table 2.2: Study designs used for the evaluation of complex patient safety interventions....31

Table 3.1: Hospitals that participated in $SPI_1$.....42

Table 3.2: Key elements of the SPI .....47

Table 3.3: Measurements to track IOM dimensions (STEEP), taken from materials presented at kick-off session (Institute for Healthcare Improvement and Health Foundation, 2005a)...61

Table 3.4: The three faces of performance management.....76

Table 4.1: Comparison of holistic and explicit methods of data extraction in case-note review .....99

Table 4.2: Summary of sub-studies comprising the evaluation of the SPI .....107

Table 4.3: Hospital characteristics of $SPI_2$ hospitals and matched control hospitals .....111

Table 5.1: Staff survey scores in $SPI_1$ hospitals at the two periods* .....120

Table 5.2: Staff survey scores in control and $SPI_1$ hospitals at the two periods*.....123

Table 5.3: Staff survey scores in control and $SPI_2$ hospitals at the two periods*.....125

Table 6.1: 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 .....134

Table 6.2: Vital signs that should be recorded.....138

Table 6.3: Appropriate clinical response for abnormal observations.....138

Table 6.4: Medical history taking (% of patients asked) .....146

Table 6.5: Medical history taking – differences between control and intervention hospitals, changes over time and the effect of $SPI_1$.....148

Table 6.6: Medical history taking (% of patients asked) .....150

Table 6.7: Medical history taking – differences between control and $SPI_2$ hospitals, changes over time and the effect of $SPI_2$ ..... 151

Table 6.8: Vital signs – percentage compliance with standards within the $SPI_1$ evaluation . 153

Table 6.9: Vital signs- percentage compliance with standards within the $SPI_2$ evaluation .. 156

Table 6.10: Vital signs– differences between control and $SPI_2$ hospitals, changes over time and the effect of $SPI_2$ ..... 158

Table 6.11: Appropriate clinical response – compliance with standards ..... 161

Table 6.12: Appropriate clinical response - differences between control and intervention hospitals, changes over time and the effect of $SPI_1$ ..... 163

Table 6.13: Appropriate clinical response ..... 166

Table 6.14: Appropriate clinical response – difference between control and $SPI_2$ hospitals, changes over time and the effect of $SPI_2$ (formal analyses for three items only, because of sparse data for other items) ..... 168

Table 6.15: Features of good care for specific classes of patient ..... 171

Table 6.16: Features of good care for specific classes of patient – differences between control and intervention hospitals, changes over time and the effect of $SPI_1$ ..... 172

Table 6.17: Use of steroids and antibiotics, CURB score and other standards applicable to specific cases – compliance with standards ..... 174

Table 6.18: Steroids and antibiotics, CURB score and other standards applicable to specific cases – differences between control and $SPI_2$ hospitals, changes over time and the effect of $SPI_2$ ..... 175

Table 6.19: Prescribing errors ..... 177

Table 6.20: Prescribing Errors – differences between control and intervention hospitals, changes over time and the effect of $SPI_1$ ..... 179

Table 6.21: Examples of prescribing errors relating to each stage of the drug use process found within the $SPI_1$ evaluation ..... 180

Table 6.22: Prescribing Errors ..... 182

Table 6.23: Prescribing errors – differences between control and $SPI_2$ hospitals, changes over time and the effect of $SPI_2$ ..... 183

Table 6.24: Anti-coagulant prescribing errors ..... 185

Table 6.25: Anti-coagulant prescribing errors – analysis ..... 185

Table 6.26: Anti-coagulant prescribing errors $SPI_2$ ..... 186

Table 6.27: Reconciliation errors at admission ..... 188

Table 6.28: Reconciliation errors at admission – differences between control and intervention, changes over time and the effect of $SPI_1$ ..... 188

Table 6.29: Reconciliation errors at admission $SPI_2$ ..... 189

Table 6.30: Reconciliation errors at admission – differences between control and $SPI_2$, changes over time and the effect of $SPI_2$ ..... 189

Table 7.1: Classification of errors and adverse events ..... 195

Table 7.2: Holistic review: changes in ratings and numbers of adverse events and errors between control and $SPI_1$ hospitals ..... 203

Table 7.3: Holistic review: changes in ratings and numbers of adverse events and errors between control and $SPI_2$ hospitals (standard errors in parenthesis) ..... 204

Table 7.4: Rates (per 100 patients) of errors identified by broad category of error – holistic review ..... 205

Table 7.5: Rates per 100 patients of errors identified by broad category of error (Standard errors are in parenthesis) $SPI_2$ ..... 206

Table 8.1: Sample sizes for 80% power (at 5% significance) ..... 213

Table 8.2: Reviewer agreement in the peri-operative case-note ..... 216

Table 8.3: Rates of compliance with peri-operative care standards ..... 216

Table 8.4: Peri-operative review: changes in the level of compliance between $SPI2$ and control hospitals and the effect of $SPI_2$ ..... 217

Table 9.1: Soap and AHR consumption - median and inter-quartile ranges for control and intervention hospitals, pre- and post-intervention period (total hospital consumption rates per 1,000 bed days) ..... 224

Table 10.1: Rates per 100 patients of adverse events identified by broad category of adverse events ..... 230

Table 10.2: Preventable adverse events identified as being strongly* or certainly preventable out of the 1260 case notes reviewed in the $SPI_1$ holistic review.....231

Table 10.3: Rates (per 100 patients) of adverse events among patients admitted with acute respiratory disease .....234

Table 10.4: Preventable deaths in acute medical wards across the study epochs.....235

Table 11.1: Mortality rates .....240

Table 11.2: Analysis of mortality rates .....240

Table 11.3: Mortality among acute medical care patients whose case-notes were reviewed .....242

Table 11.4: The effect of $SPI_2$ on the mortality among acute medical care patients.....242

Table 12.1: Intensive care outcomes– median and inter-quartile ranges for control and $SPI_2$ hospitals, pre- and post-intervention period .....250

Table 13.1: Rates of C.diff (per 1,000 bed days ) and MRSA infections (per 100,000 bed days) .....256

Table 14.1: Patient survey scores in $SPI_1$ hospitals at the two periods.....264

Table 14.2: Patient survey scores in control and $SPI_1$ hospitals at the two periods .....265

Table 14.3: Patient survey scores in control and $SPI_2$ hospitals at the two periods .....267

# ABBREVIATIONS

| | |
|---|---|
| $\kappa$ | Kappa |
| °C | degree Celsius |
| A&E | Accident and Emergency |
| AHR | Alcohol Hand Rub |
| ANOVA | Analysis of Variance |
| APACHE II | Acute Physiology and Chronic Health Evaluation II |
| AR | Auto Regressive |
| BNF | British National Formulary |
| BPM | Beats per Minute |
| BTS | British Thoracic Society |
| C. diff | Clostridium difficile |
| CI | Confidence Intervals |
| CMP | Case Mix Programme |
| $CO_2$ | Carbon Dioxide |
| COPD | Chronic Obstructive Pulmonary Disease |
| CQC | Care Quality Commission |
| CURB | Confusion Urea Respiratory rate Blood pressure |
| DVT | Deep Vein Thrombosis |
| EWSS | Early Warning Score Systems |
| FMEA | Failure Mode Effects Analysis |
| FTE | Full Time Equivalent |
| GEE | Generalised Estimating Equations |
| GP | General Practitioner |
| HCAI | Healthcare Associated Infections |
| HES | Health Episode Statistics |
| HPA | Health Protection Society |
| HR | Human Resources |

| | |
|---|---|
| HTA | Health Technology Assessment |
| ICC | Intraclass Correlation Coefficient |
| ICNARC | Intensive Care National Audit and Research Centre |
| ICPS | International Classification for Patient Safety |
| ICU | Intensive Care Unit |
| IHI | Institute for Healthcare Improvement |
| IMRAD | Introduction, Methods, Results and Discussion |
| INR | International Normalised Ratio |
| IOM | Institute of Medicine |
| IOM | Institute of Medicine |
| IQR | Inter-quartile Range |
| LOS | Length of Stay |
| mL | Millilitre |
| mmHg | Millimetre of Mercury |
| MRSA | Methicillin-resistant Staphylococcus aureus |
| N | Number |
| NHS | National Health Service |
| NICE | National Institute for Health and Clinical Excellence |
| NOSEC | National Observational Study to Evaluate the “cleanyourhands” campaign |
| OR | Odds Ratios |
| PCA | Patient Controlled Analgesia |
| PDSA | Plan Do Study Act |
| PPV | Positive Predictive Value |
| PUD | Peptic Ulcer Disease |
| QALY | Quality of Life Adjusted Year |
| QUERI | Quality Enhancement Research Initiative |
| RCT | Randomised Control Trial |
| RPN | Risk Priority Number |
| SBAR | Situation Background Assessment Recommendation |

| | |
|---|---|
| SDO | Service Delivery Organisational |
| SE | Standard Errors |
| SPC | Statistical Process Control |
| **SPI** | Safer Patients Initiative |
| $\boldsymbol{SPI_1}$ | Pilot phase hospitals of the Safer Patients Initiative |
| $\boldsymbol{SPI_2}$ | Second phase hospitals of the Safer Patients Initiative |
| U&E | Urea and Electrolytes |
| UK | United Kingdom |
| US | United States |
| WHO | World Health Organisation |

# EXECUTIVE SUMMARY

## Background

Many patient safety interventions are concerned with best to how organise health systems and deliver services. They can be difficult to define and the effective element difficult to specify, and these are known as complex interventions. This thesis is concerned with the evaluation of complex patient safety interventions, and attempts to answer the following research question: ***which study design provides a robust method of evaluation of complex patient safety interventions?***

Fundamentally, study designs available to measure the effectiveness of patient safety interventions fall into two categories – those that use contemporaneous controls, and those that do not. This thesis firstly describes a review of the literature, which scopes out the patient safety evaluation literature, to determine which study designs are currently employed to determine effectiveness of interventions. It then elaborates on a case-study of an evaluation of a highly complex patient safety intervention – the *Safer Patient Initiative* (**SPI**).

## Literature review

A review of the recent literature was conducted (2005 to present), and 245 citations met the search criteria. Sixty-three percent of studies (154/245) were single-centre studies and evaluations that did not use contemporaneous controls dominated these types of studies (84% [129/154]). Multi-centre studies were used less frequently 37% (91/245); however, the use of contemporaneous controls (49% [45/91]) was greater than in single-centre studies, and equalled those evaluations that did not use contemporaneous controls (51% [46/91]). Indicating that those studies that considered a multi-centre design, also considered a more robust evaluation.

Studies that do not use contemporaneous controls dominate the literature, but they are inherently weak and are subject to bias. Robust evaluations should include contemporaneous controls, as well as before and after measurements, as they account for any underlying secular trends. Ideally evaluations should be augmented with mixed-methods, as this enables research findings of one type to be reinforced when the are corroborated by the findings of another type. These principles are applied to the case-study that constitutes the remainder of the thesis.

## Case-study - *the Safer Patients Initiative*

The intervention was delivered in two phases -the first phase ($\boldsymbol{SPI_1}$) began in 2005, when the Health Foundation (an independent charity working to improve quality and safety) worked

with the Institute for Healthcare Improvement (IHI) to develop and launch the *Safer Patients Initiative* (**SPI**). They provided direct support to four NHS hospitals as they implemented an organisation-wide patient safety programme over 18 months. The intervention sought to generically strengthen hospitals by engaging senior leaders to change the culture of the organisation whilst improving the reliability of specific front-line care processes within designated clinical areas

The stated aim was to make a 50% reduction in the total number of adverse events at the four pilot 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.

To build on the experience and learning from ***SPI$_1$***, a second phase of the intervention, known as ***SPI$_2$***, was rolled out from March 2007 to September 2008 inclusive. Following a competitive bidding process, 20 UK hospitals were selected to participate in this phase. The ***SPI$_2$*** had an investment from the Health Foundation of approximately £270,000 per hospital. The ***SPI$_2$*** intervention was similar to ***SPI$_1$***, with somewhat modified goals.

## Methods

The evaluation of **SPI** employed a before and after study designed with non-randomised contemporaneous controls. It utilised mixed-methods – using both quantitative and

qualitative studies. A series of linked sub-studies were organised around a conceptual model of the healthcare system:

- management processes were studied by means of a series of qualitative interviews with strategic stakeholders;
- intervening variables were assessed by using a quantitative staff survey, and a qualitative study using ethnographic methods on hospital wards;
- clinical processes were studied by case-note review of acutely sick patients using explicit methods and holistic methods, an explicit case-note review of peri-operative care, and by using indirect measures of hand hygiene;
- outcomes were studied by case-note review of adverse events and mortality, outcomes within the intensive care unit, healthcare associated infection rates (HCAI), and quantitative patient surveys.

The evaluation of $\boldsymbol{SPI_1}$ preceded the evaluation of $\boldsymbol{SPI_2}$ and not all sub-studies were conducted in both phases. Primarily, how practitioners responded to acutely ill and deteriorating patients, were the only quantitative sub-studies conducted for both $\boldsymbol{SPI_1}$ and $\boldsymbol{SPI_2}$ evaluation, whilst the qualitative sub-studies were only carried out within the $\boldsymbol{SPI_1}$ evaluation.

For all quantitative comparisons, a “difference-in-difference” method was used, so that the intervention effect could be estimated as a rate of change above and beyond any background rate of change.

## 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 could not 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. Within the ***SPI$_2$*** evaluation, only one dimension of the staff survey changed significantly (in favour of control hospitals).

Many aspects of evidence-based medical and peri-operative care were good at baseline, leaving little room for improvement. There was marked improvement in use of hand-washing materials, and a dramatic decrease in HCAI's across all hospitals. A significant additive effect of the **SPI** on the measures included in the evaluation was not detected.

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 **SPI** procedures, practices and principles, or viewed them as top-down rather than something that 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.

## Main findings

Many aspects of care are already good or improving across the NHS, suggesting considerable gains in quality across the board. These improvements might be due to policy activities, including some with features similar to the **SPI**, and the emergence of professional consensus on some clinical processes.

An additional effect of a large-scale organisational intervention (**SPI**) was not detected. There are a number of possibilities for this: (i) any effects were too small to detect; (ii) the null additive effect was due to sub-optimal implementation; or that there may be longer-term additive effects that take longer to surface.

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 were underestimated. It may simply take a long time for programmes such as the **SPI** to achieve this effect.

Hospitals did report positive effects from **SPI** participation, including heightened managerial awareness and/or commitment 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 may either surface at a later date or prove more sustainable than the improvements seen in control hospitals. This hypothesis can only be tested through further data collection.

## Conclusion

In this thesis I have described a robust study design to evaluate complex patient safety interventions. The evaluation has taken a mixed method approach, using both quantitative and qualitative methods. All quantitative observations have been made with contemporaneous controls, and uniquely a “difference-in-difference” method was used.

The approach has been vindicated – measures of fidelity, intervening variables, process and outcome, supported by qualitative methods – and has provided a wealth of information of the effectiveness of the intervention and theories as to why the intervention worked the way it did.

# CHAPTER 1: INTRODUCTION

How best to secure greater patient safety is one of the most challenging and pressing questions facing healthcare practitioners, managers and policy makers. Increasing effort and resources have been focused on patient safety since the publication of two key reports at the turn of the century – the UK Chief Medical Officer's *Organisation with a Memory* (2000) and the US Institute of Medicine (IOM) *To Err is Human* (2000). This increased focus has gone hand in hand with an explosion of patient safety research literature (Lilford et al., 2006).

Patient safety research builds on cognate disciplines such as psychology, sociology, organisational studies, ergonomics and education, and so the design of patient safety interventions to improve patient care requires knowledge distilled from these basic disciplines. Some of these interventions aim to replace unsafe treatments and technologies (such as medical devices or surgical) with safer alternatives, and the development and evaluation of this type of intervention are under the remit of the Health Technology Assessment (HTA). However many patient safety interventions are concerned with how best to organise health systems, and deliver services so that they are safer. Research of this type falls within health service, or Service Delivery and Organisational (SDO) research.

Patient safety interventions that target the delivery and organisation of the service tend to comprise of a number of separate components, that can be complementary or work

independently. They can also involve a number of behaviour changes; different levels/structures within an organisation; different elements being delivered at different times and intensity; and different methods of organising and delivering those elements. These types of intervention, i.e. those that are difficult to define, and where the effective element of the intervention is difficult to specify, are known as “complex interventions” (Medical Research Council, 2000).

In 2000, the Medical Research Council (MRC) published a framework for the evaluation of complex interventions, in which an analogy to a randomised control trial (RCT) of a pharmaceutical intervention was made (Campbell et al., 2000; Medical Research Council, 2000). Despite the framework being welcomed as an attempt to bring scientific rigour to the evaluation of complex interventions (Rosen et al., 2006), it also brought several criticisms that were raised at a subsequent meeting (Medical Research Council, 2008):

- the guidance was derived from the evaluation of pharmaceutical and medical devices, and so was based on the assumption that the planning of a complex intervention is a linear process (Campbell et al., 2007);
- the lack of evidence for the framework;
- the lack of preliminary stages, such as early piloting and developmental work (Hardeman et al., 2005);
- an assumption that conventional clinical trials provide a template for all the different approaches to evaluation;

- that processes were not as important/useful as outcome measurements (Oakley et al., 2006);
- an assumption that interventions need to be standardised, and a lack of recognition that in some circumstances, the effectiveness of an intervention requires local tailoring (Hawe et al., 2004);

In response to these criticisms, and from the subsequent learning that has arisen, new guidance was issued (Craig et al., 2008; Medical Research Council, 2008). However, the updated recommendations do not claim to be comprehensive, and that achievable consensus on what is best practice has yet to be accomplished. However, it recognises that the process of developing and evaluating complex interventions has several phases, which may or may not follow a linear sequence.

The initial stage is the development of the intervention, which should have a sound theoretical basis. At this stage the existing evidence for the intervention should be identified, preferably by a high quality systematic review. Also, the intervention needs to be described in sufficient detail to allow it to be fully replicated.

There should also be a feasibility and piloting stage to test whether the intervention can be delivered as intended, and to estimate the effect size, so that the sample size can be determined. In the evaluation stage, the effectiveness of the intervention should be based on experimental design, but if the bias is small and the effect size is large or too rapid, then an

observational study may be more appropriate. This is consistent with the "no one size fits all" approach advocated by Brown et al., (2008d). Outcome measurements are still preferred, but the inclusion of process measurements would offer an explanation of observed effects. Also, the inclusion of an economic evaluation to assess the cost-effectiveness of the intervention would make the findings useful to decision makers.

Finally, the findings of an evaluation need to be disseminated and made available to key stakeholders. There was also a recommendation of surveillance and monitoring of interventions, augmented with long-term follow up.

The guidance also provided further clarity on what makes an intervention complex. They concurred that there is no clear boundary between simple and complex interventions in actual fact there are few interventions that can be described as truly simple, though there are some that are highly complex as they comprise of several multi-faceted components. They offer some guidance as to the dimensions of complexity, and their implications for developing and evaluating them (Craig et al., 2008). See Box 1.1 and Box 1.2.

**Box 1.1: Dimensions of complexity (taken from Craig et al., 2008)**

- number of interacting components within the experimental and control interventions;
- number and difficulty of behaviours required by those delivering or receiving the intervention;
- number of groups or organisational levels targeted by the intervention;
- number and variability of outcomes;
- degree of flexibility or tailoring of the intervention permitted.

**Box 1.2: Implications for developing and evaluating complex studies (taken from Craig et al., 2008)**

- a good theoretical understanding is needed of how the intervention causes change, so that weak links in the causal chain can be identified and strengthened;
- lack of effect may reflect implementation failure (or teething problems) rather than genuine ineffectiveness  a thorough process evaluation is needed to identify implementation problems;
- variability in individual level outcomes may reflect higher level  processes; sample sizes may need to be larger to take account of the extra variability and cluster randomised designs considered;
- a single primary outcome may not make best use of the data  a range of measures will be needed and unintended consequences picked up where possible;
- ensuring strict standardisation may be inappropriate – the intervention may work better if a specified degree of adaption to local settings is allowed for in the protocol.

This new guidance will provide a steer for this thesis, but as with the previous MRC guidance, there is debate on its relevance (Connelly, 2007; Hawe et al., 2004; Kernick, 2008; Mackenzie et al., 2010; Shiell et al., 2008). Fundamentally, these arguments are on how knowledge is built up, and reflects the different ontological (what can be known) and epistemological (how can it be known, i.e. the theory of knowledge) that underpin the basic belief system, or world view that guides a person/investigation/discipline (Guba and Lincoln, 1994). I take a pragmatic position using different ontological and epistemological philosophies depending on the research question that needs to be answered. However, essentially I am positivist/realist, i.e. there are external objective facts but these can be interpreted through social conditioning. This positivist/realist philosophy underpins this thesis and hence the use of the MRC guidance as point of reference for it.

The research question for this thesis is: ***which study design provides a robust method of evaluation of complex patient safety interventions?***

To answer this, firstly a literature review was undertaken to assess study designs currently used in the evaluation of patient safety interventions (Chapter 2).

I then progress to describe in detail a highly complex patient safety intervention (Chapter 3) – the *Safer Patients Initiative* (**SPI**), and the framework of the methods (both quantitative and qualitative) employed to evaluate it (Chapter 4). The quantitative sub-studies are reported separately (Chapters 5 to 14), with a brief introduction, methods, results and discussion. An

overall discussion of the findings is provided in the final chapter (Chapter 15), along with an exegesis of the findings and lessons for the evaluation of highly complex patient safety interventions, such as the **SPI.**

# CHAPTER 2: LITERATURE REVIEW

## 2.1 Introduction

In Chapter 1, I briefly described the updated guidance for the evaluation of complex interventions (Craig et al., 2008; Medical Research Council, 2008). By and large the authors endorsed the use of epidemiological methods in the evaluation of complex interventions, but what is not known is what study designs have been used to date to evaluate them, and, more specifically to this thesis, what study designs have been used to evaluate complex patient safety interventions. This chapter will review the recent patient safety intervention literature to investigate this.

### 2.1.1 Study design

The gold standard for evaluating interventions is the randomised control trial (RCT) (Cochrane, 1979). In this type of study the intervention is randomly allocated between control and intervention groups before the investigation begins. However, randomisation is not always possible for either practical or ethical reasons (Black, 1996), and in these circumstances non-randomised designs have been used.

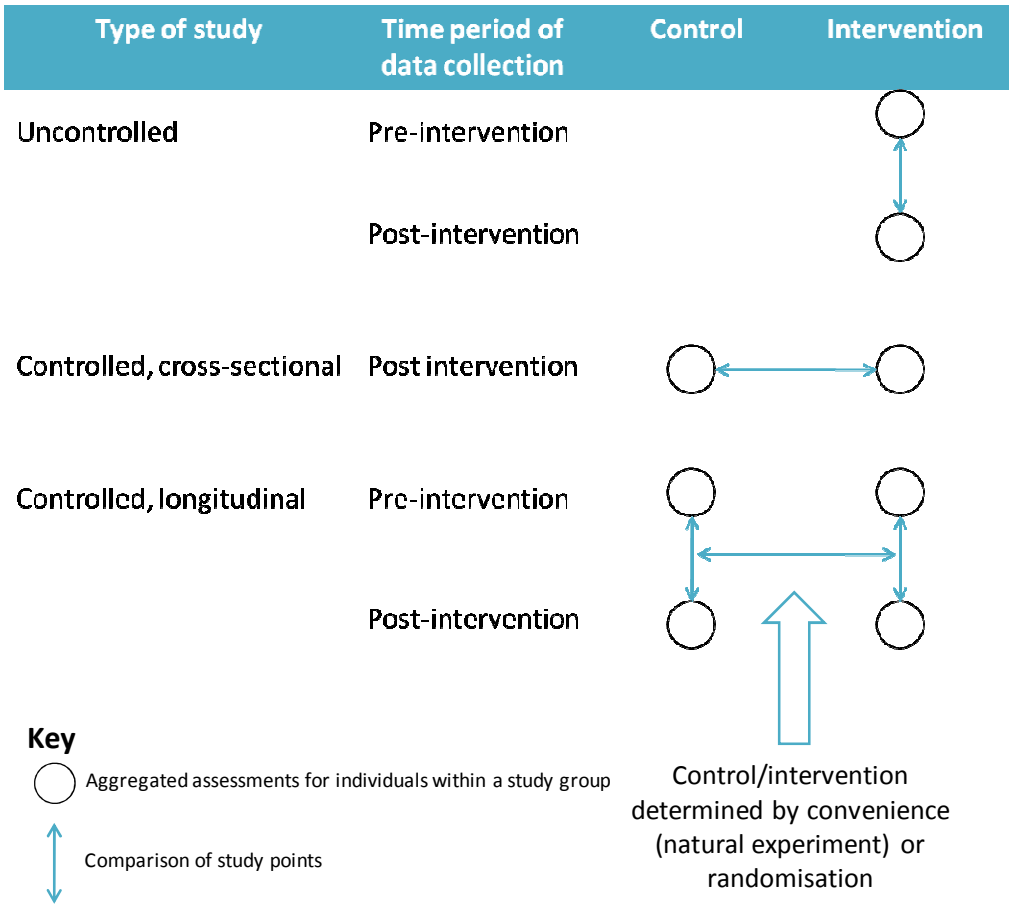
This section outlines study designs that are commonly used to evaluate complex patient safety interventions (see Figure 2.1) (Brown et al., 2008b). I shall start by describing study designs that measure endpoints "before and after" an intervention has been implemented

(section 2.1.2). These studies may involve many time points before, during and after the intervention phase, and depending on how the data is to be analysed then these designs may be referred to as time series analysis or Statistical Process Control (SPC) (see 2.1.2).

I will then go onto describe study designs that make use of contemporaneous comparisons. These can be made after an intervention has been put in place (cross-sectional studies), or can be made both before and after the intervention phase (controlled before and after studies). Designs with contemporaneous controls may be: a) non-experimental (natural experiments or "quasi-experimental" studies (Campbell and Stanley, 1963)) or b) they may involve experimental design in which intervention and controls are chosen at random, (i.e. RCT).

Studies using these two design variables (the timing of data collection, and whether the allocation to control or intervention is randomised or not) are described in section 2.1.3.

**Figure 2.1: Basic study designs. Taken from Brown et al., (2008b)**




### 2.1.2 Before and after studies

In some instances contemporaneous controls are not feasible, so neither RCTs nor natural experiments are possible. This situation can arise when there is a national directive to implement a policy across a service, or when an intervention is being implemented in a single hospital. In these instances, before and after studies are common, and are often the only practical method of evaluation. However, they are a relatively weak method to distinguish cause and effect as secular trends, while other concurrent changes may make it difficult to attribute observed changes to the intervention being studied. Before and after

evaluations also tend to overestimate the effect size; Lipsey and Wilson (1993), reviewed 45 meta-analyses of psychological, educational and behavioural interventions, noted that the observed effects from before and after studies were greater than those from controlled studies, a finding that was previously made in clinical research (Sacks et al., 1987).

Factors influencing how much confidence can be placed on the results of before and after studies include:

- the magnitude and rate of change;
- the plausibility of the intervention;
- compatibility with other contemporaneous and prior evidence (Brown et al., 2008b).

#### *2.1.2.1 Time series and Statistical Process Control (SPC)*

Time series is a type of before and after study design, it attempts to detect if the effect of an intervention is significantly greater than any secular trend (Cook and Campbell, 1979). A statistically significant interruption in a long series of observations has greater explanatory power than differences between a single before and after observation, as regression to the mean is less likely if serial observations show that the improvement was not preceded by a random “high”.

Time-series are particularly useful when it is difficult to obtain contemporaneous controls, as in the dissemination of national guidelines or mass media campaigns (Eccles et al., 2003).

Data collected several times before and after the intervention allows underlying/seasonal (cyclical) effects to be estimated. However, enough data points need to be collated prior to the intervention, to demonstrate that a stable estimate of the underlying secular trend has been obtained. Also, the time period between successive data points should be chosen with care, as data points collected close in time are likely to be more similar to one another, than to data points collected further apart, i.e. close data points autocorrelate (Eccles et al., 2003).

SPC (see 3.6.3.2) is a method of monitoring a process through the use of control charts. The advantage of using SPC is the ability to detect process changes and early trends. Users making pragmatic improvements can move away from simple bar and line graphs, mainly descriptive analyses, to a method, that when used correctly, is statistically rigorous but requires less data than when testing for significance (Benneyan et al., 2003). Although SPC is strictly a quality improvement method, it is included here as it is an approach that has been widely adopted in the patient safety field, and it possible that the findings of SPC will be reported for evaluation purposes.

### 2.1.3 Studies with contemporaneous controls

Comparative studies between hospitals exposed to the intervention and controls which are not, provide a much better basis for inferences about effectiveness, than a before and after study. However, before explaining the different design elements of these comparative studies I will discuss the unit of comparison most commonly used – a cluster.

#### *2.1.3.1 The unit of comparison*

Patient safety interventions are usually targeted at the service rather than the individual, and as interventions will affect a group of patients, cluster studies are often necessary (Donner, 1998; Edwards et al., 1999).

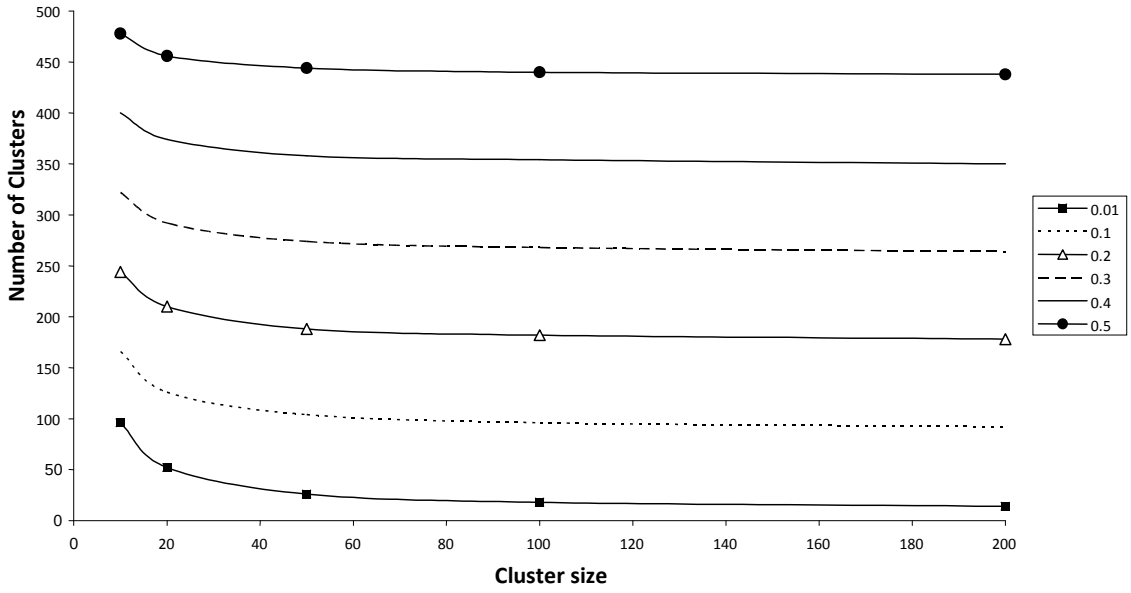
Typically clusters consist of different organisations (hospitals, general practices, etc), but other types of cluster may be used. For example, clusters can be comprised of patients treated by doctors exposed to different interventions (Landrigan et al., 2004). A drawback of cluster studies is the loss of power that results from greater similarity across individuals ***within*** a cluster, than across individuals ***between*** clusters.

As with other studies, sample size requirements in cluster studies depend on the size of effect sought, and the risks of false positive and false null study findings. However, the sample size also depends on the extent to which endpoints tend to cluster within an organisation, which is measured by the intraclass correlation coefficient (ICC). The ICC range is between 0 and 1, and will increase as the variance in endpoints between individuals within clusters falls. If the ICC is 0, there is no correlation between individuals in a cluster, and the study is effectively a parallel design. If the ICC is 1, the responses of individuals within a cluster are identical, and the effective sample size is the number of clusters (Killip et al., 2004).

ICC values are generally between 0.01 and 0.02 for healthcare organisations (Killip et al., 2004). A typical two-arm cross-sectional comparative study might include 20 clusters in each

arm with 40 individuals in each cluster. Such a study would be sufficient to detect a statistically significant reduction in error rate from 10% to 5% assuming an ICC of 0.02 (alpha = 0.05; 80% power). The total sample size of 1600 is almost twice the requirement of 868 if clustering is not taken into account. Figure 2.2 shows the trade-off between the number of clusters and cluster size required to detect a difference in error rate from 10% to 5% with different ICC values. Diminishing marginal returns to increasing cluster size on the power of study are well demonstrated (Campbell et al., 2004).

**Figure 2.2: Clustered sample size calculations for given ICC values (box). The calculations are based on detecting a difference in error rate between control (0.1) and intervention (0.05), with alpha = 5% and power = 80%. Taken from Brown et al., (2008b).**



Ukoumunne et al., (1999) identify two types of cluster design:

- cohort designs, in which repeated measures are taken from the same subjects in each cluster;
- and repeated cross-sectional designs, in which the sample of subjects within each cluster changes when each measurement is taken.

The latter design enables the repeated measurement of terminal end-points (those that only happen once) such as death, thus, the effect of a safety intervention on mortality or rates of infection can be measured in different organisations, and at different times.

#### *2.1.3.2 Design elements of studies using contemporaneous controls*

There are four types of controlled study designs (see Figure 2.1):

- cross sectional (measurement of post intervention end points) with non-randomised controls;
- cross sectional with randomised controls;
- "before and after" with non-randomised controls;
- "before and after" with randomised controls.

The advantages and disadvantages of these different types of study design are described in Table 2.1. Of these study designs, the least reliable are non-randomised post-intervention comparisons, due to the inherent difference between intervention and controls. Statistical adjustment for confounders can only be made for known and observed confounding

variables, but bias frequently originates in hidden variables (Lilford et al., 2004). Randomised studies with pre-intervention and post-intervention measures are arguably the most robust design, but non-randomised comparative studies with before and after measurements may be nearly as good, because measured rates of change can be taken into account (imperfectly) for baseline differences. The method is imperfect, since clusters with a high (or low) baseline rates may be more (or less) responsive to the intervention. In particular, clusters that actual have good practice may have little scope for improvement. Whether randomised or not, a controlled before and after design can control for secular change, since the intervention effect is estimated as a rate of change above and beyond any background rate of change, i.e. the difference-in-difference comparision. With an appropriate statistical model both secular trends and intervention effects can be modelled (Saint et al., 2003).

Sample size calculations for comparisons of rates of change in cluster studies do not need to take account of the ICC of the end points, but only of any ICC of the propensity to change, net of baseline differences. Murray and Blitstein (2003) provide theoretical justification and practical evidence that these latter ICCs, based on pre-post differences, are smaller than ICCs that must be used in cross-sectional comparisons.

However, there are caveats to using before and after controlled comparisons to minimise spurious claims about the effect of the intervention. Firstly, intervention and control groups need to be comparable, and so need to be effectively matched; secondly, baseline

measurements need to be similar, as any differences between them may indicate that they experience secular trends differently (as mentioned above) (Eccles et al., 2003).

**Table 2.1: Controlled study design matrix**

| | Allocation | |
|---|---|---|
| **Phasing** | **Randomised** | **Natural experiment** |
| **Post Intervention** | Variance in end points due to baseline difference may bias the result | Risk that comparisons will be confounded by difference at baseline |
| **Before and after** | Allows for specific comparison of change net of any baseline differences. Enables comparison to be made between organisations that change most or least | Controls for baseline difference possible |

### 2.1.4 The importance of context

Epidemiological methods provide inferences on the effectiveness of an intervention, and should remain at the core of an evaluation. However, it is recognised that a variety of methods are important in the evaluation of complex interventions. These are particularly insightful for delving into the underlying processes and the social context within which an intervention is introduced. These are an inherent part of change, and vital to understanding why an intervention does or does not work (Murtagh et al., 2007; Thomson et al., 2007), and can also explain why differences in capacity to benefit may exist. For example, there may be a "ceiling" effect, whereby non-intervention organisations are already very good, and have less headroom for further improvement than intervention sites that demonstrated positive results.

In order to make judgements regarding generalisability, it is necessary to have some knowledge of any systematic differences of those characteristics judged to be important

between study and non-study organisations. In the event that such differences exist, the theory as to why and how an intervention may work, will help in predicting the impact of such differences.

Finally, in addition to contextual differences, studies associated with greater (and significant) effects tend to be those where the intervention was implemented with greater fidelity. It is therefore important that the fidelity with which an intervention was implemented (or the method by which it may have been adapted) are clearly described, as these will properly inform judgements about the generalisability of the findings.

### 2.1.5 Literature review

I will now describe a literature review to determine what study designs have been used to evaluate patient safety interventions. As the new MRC guidance was issued in 2008 I will search the recent history of evaluations (2005 to present), and focus on those papers that report a study carried out in secondary care, as these directly relate to the subject of this thesis, an evaluation of a hospital intervention (Chapters 4 to 15).

I will use a broad search strategy that includes quality of care improvement interventions, as well as patient safety interventions. Since safety and quality are aspects on a continuum and there is no clear point at which safety becomes quality (Brown et al., 2008a).

## 2.2 Methods

### 2.2.1 Identification of studies

The following search strategy was used in OVID (MEDLINE and EMBASE) to locate primary studies of potential interest:

("patient safety" OR "quality improvement") AND ("intervention$") AND ("hospital$" OR acute)

The strategy was limited to humans, English language and for articles published from 2005 to present (13th February 2011). The terms in the search string did not map to any Medical Subject Headings (MeSH) and were therefore used as keywords limited to the abstract. A similar search strategy was used for the Web of Science database but with a restriction to articles.

MEDLINE and EMBASE databases were de-duplicated within OVID, with a further de-duplication step taken between OVID and Web of Science when all citations were imported into Reference Manager® (version 12), a software tool for publishing and managing bibliographies.

### 2.2.2 Exclusion criteria

The abstracts that this search yielded were reviewed and discarded if they described:

- a study targeting primary care;
- a literature review;
- a registry;
- a purely qualitative study;
- a descriptive study;
- the implementation of national guidelines/policy with no additional intervention;
- a clinical study;
- an opinion paper.

Full papers were retrieved of those abstracts that were of potential interest.

### 2.2.3 Data extraction

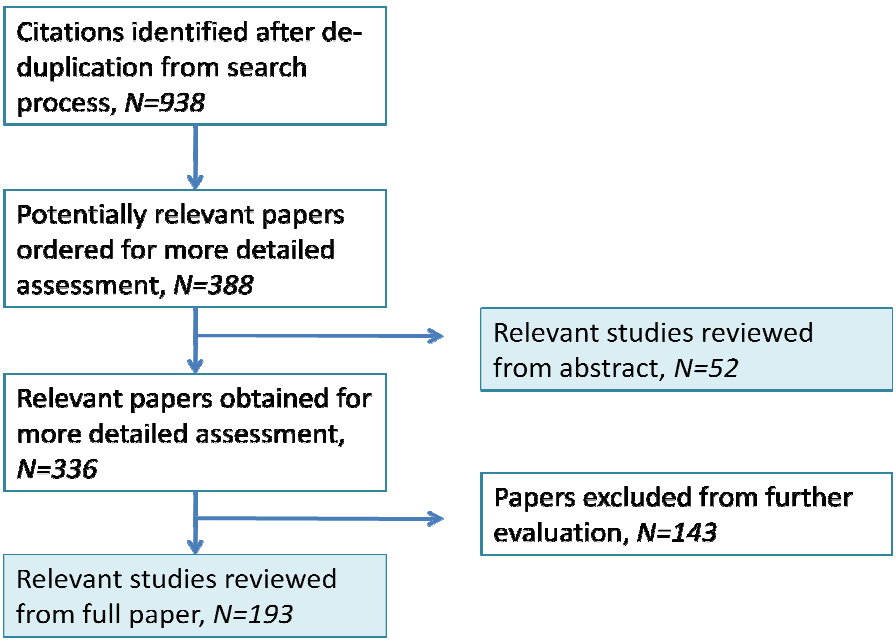
Relevant studies were initially categorised to single-site or multi-site evaluations. Following on from this, studies were then divided into those that used before and after designs, versus those that used contemporaneous controls. Among those using contemporaneous controls cross-sectional studies were distinguished from before and after controlled studies, and within both of these categories studies were further differentiated between randomised and non-randomised controls.

## 2.3 Results

### 2.3.1 Study flow

The search strategy used in this literature review provided a yield of 938 citations. From reviewing the abstracts 388 were identified as being potentially relevant. Of these 336 were available for detailed review, and a further 52 were reviewed by abstract only due to lack of availability and financial constraints. After detailed review, 143 of the 336 papers were excluded. In total, data was extracted from 245 citations (full papers and abstracts) for this literature review. The literature search process is summarised by Figure 2.3, and the full list of citations included in this literature review can be found in Appendix A.

**Figure 2.3: An overview of the literature search process**




*those studies included in the literature review are indicated by the shaded boxes, N=245*

### 2.3.2 Study designs used

See Table 2.2 for a summary.

#### *2.3.2.1 Single-centre studies*

Of the 245 papers and abstracts included in the literature review, 63% (154/245) were evaluations conducted at a single-centre. Of these single-centre evaluations, 84% (129/154) were without contemporaneous controls, with before and after studies making the largest proportion within this group (90% [116/129] before and after, *vs.* 4% [5/129] time-series, *vs.* 6% [8/129] SPC).

Contemporaneous controls were used in 16% (25/154) of single-centre studies. Of these 56% (14/25) used a cross-sectional design (43% [6/14] randomised, *vs.* 57% [8/14] non-randomised); and 44% (11/25) used an additional before and after design element (9% [1/11] randomised, *vs.* 91% [10/11] non-randomised).

#### *2.3.2.2 Multi-centre studies*

Thirty-seven percent (91/245) of the papers and abstracts included in the literature review utilised a multi-centre design. Contemporaneous controls were used in 49% (45/91) of these studies, and this was virtually equal to the number of studies that did not use them (51% [46/91]).

Of those studies that did not include contemporaneous controls, before and after studies again predominated (83% [38/46] before and after, *vs.* 15% [7/46] time series, *vs.* 2% [1/46] SPC).

Of the studies that used contemporaneous controls, cross-sectional studies were used less frequently (16% [7/45] cross-sectional, *vs*. 84% [38/45] before and after); and of those before and after studies using contemporaneous controls, 47% (18/38) used randomised controls versus 53% (20/38) that used non-randomised controls.

**Table 2.2: Study designs used for the evaluation of complex patient safety interventions**

| | No contemporaneous controls | | | Contemporaneous controls | | | | |
|---|---|---|---|---|---|---|---|---|
| | **Before & after** | **Time Series** | **SPC** | **Before & after randomised controls** | **Before & after non-randomised controls** | **Cross sectional randomised controls** | **Cross sectional non-randomised controls** | **Total** |
| **Single centre N** | 116 | 5 | 8 | 1 | 10 | 6 | 8 | 154 |
| **%** | 75 | 3 | 5 | 1 | 6 | 4 | 5 | |
| | | | | | | | | |
| **Multi-centre N** | 38 | 7 | 1 | 18 | 20 | 5 | 2 | 91 |
| **%** | 42 | 8 | 1 | 20 | 22 | 5 | 2 | |
| | | | | | | | | |
| | | | | | | | **TOTAL** | 245 |

## 2.4 Discussion

### 2.4.1 Study designs used

#### *2.4.1.1 Single-centre studies*

The evaluation method that predominates within a single-centre design is before and after studies (84% *vs.* 16% using contemporaneous controls). The dominance of this type of study design within single-centre studies is probably a reflection of the fact that it is the least resource intensive type of evaluation method to conduct.

Although before and after studies are useful to measure change in a single organisation, or when polices are introduced simultaneously across a health sector, they are considered a weak design (Eccles et al., 2003). This is as they do not provide information on cause and effect, have no estimate of any secular trends that may be occurring, and they tend to overestimate effect size (Lipsey and Wilson, 1993; Sacks et al., 1987). We have discussed in section 2.1.2 factors that increase or decrease the confidence with which cause and effect conclusions can be made.

When a before and after study is the only option available then an interrupted time-series design should be encouraged, as this affords some protection against secular trends. However, their use is underutilised, and across the whole sample (single-centre and multi-centre studies) only 12 evaluations (Gallagher et al., 2009; Hixson et al., 2005; Jeffries et al.,

2009; Krimsky et al., 2009; Levy et al., 2010; Miller et al., 2010; Pronovost et al., 2010; Riggio et al., 2009; Robb et al., 2010; Unahalekhaka et al., 2007; van Kasteren et al., 2005; Young et al., 2006) used this type of study design. This is probably because they are methodologically more demanding and, again, there may be financial constraints limiting the collection of more data.

Despite the growing use of quality improvement methods, only a small proportion of the literature was based on SPC analyses. In the whole sample (single-centre and multi-centre studies) only nine papers (Beckett and Kipnis, 2009; Cohen et al., 2005; Ernst et al., 2010; Groome et al., 2009; Iyer et al., 2011; Kotagal et al., 2009; McPhail et al., 2010; Stevens et al., 2010; von and Aslaksen, 2005) cited this method for evaluation purposes. This is probably a reflection of the fact that SPC is frequently reported as part of an intervention package. For example, McCulloch et al., (2010) described the effect of a quality improvement method known as "lean" and SPC was used to improve processes and outcomes on a surgical wards. However, in the report, chi-squared test was conducted to measure effectiveness.

#### *2.4.1.2 Multi-centre studies*

Within multi-centre evaluations, contemporaneous controls were used in studies as frequently as they were not (49% [45/91] *vs.* 51% [46/91]). This may be a reflection that those evaluations that have considered putting in place an intervention in more than one site, have also considered a more robust evaluation.

Before and after studies with contemporaneous controls feature more frequently than cross-sectional studies (84% [38/45] *vs.* 16% [14/25]) in comparison to single-centre evaluations in which the split was more evenly distributed (44% *vs.* 56%). This is again a reflection of a more considered design approach when using a more than one study centre.

Across the whole literature review only a few studies (3% [10/245]) used the weakest study design, i.e. cross-sectional evaluations using non-randomised controls. These are subject to bias and have poor generalisability, so it is encouraging to see that these are not frequently employed.

Of those multi-centre studies that used before and after design, with no contemporaneous controls, it may be that the number of intervention sites was large and it was unfeasible to obtain controls. For example, Wirtschafter et al., (2006) described a state-wide intervention promoting antenatal steroid use for foetal maturation in preterm babies.

### 2.4.2 Limitations of the literature review

This literature review lacks quality control as it was done by one individual. It would have been useful if a sample was reviewed by another person so that a measure of agreement could be made.

There were 11 papers in which the intervention being described could be argued as being a simple intervention. These included simple feedback mechanisms (Chern et al., 2005; Merle et al., 2009; Midlov et al., 2008; Quillen and Murphy, 2006); protocols and checklists (Paige

et al., 2009) and reminder systems (Agostini et al., 2007; Brackbill et al., 2010; Franklin et al., 2006; Koplan et al., 2008; Vishwanath et al., 2009). However, they could equally be defined as being complex, as they required behaviour changes and were implemented in complex systems. This difficulty in defining interventions as being simple, complex, or even highly complex, is recognised as problematic, particularly for systematic reviewers, who have called for the development of a typology of the structural characteristics of complex interventions (Shepperd et al., 2009).

### 2.4.3 Implications for evaluating complex patient safety interventions

Despite guidance on how we should evaluate complex interventions being available for over ten years (Medical Research Council, 2000; Medical Research Council, 2008), before and after study designs still dominate the literature. As stated previously, they are inherently often weak and are subject to bias.

So, which study design should be used in the evaluation of complex patient safety interventions? Firstly, contemporaneous controls should be used where possible. Ideally these should be randomised controls, but this is often not feasible. Before and after measurements are also important as they allow for any underlying secular trends to be identified, and in a cluster study, use of a difference-in-difference method enables smaller sample sizes to be used (see 2.1.3.2).

The considerations outlined above should be central to any evaluation of complex intervention but they can be further augmented by the use of mixed-methods. A mixed-

method evaluation involves measuring different end points and making both qualitative and quantitative observations. There are four key advantages to using this approach (Sandelowski, 2000):

- the ability to triangulate findings (the research findings of one type of observation are reinforced when they are corroborated by the findings of a different type). Additional credence through triangulation adds to the strength and generalisability of findings. Alternatively, where the results from different methods conflict, triangulation may prevent presumptive inferences that may have been made had results only been obtained from measurement of a single end point (Brown, 2004);
- greater understanding, enabling better interpretation and elaboration of results. A mixed-method approach allows researchers to consider *why* an intervention has or has not been effective;
- development of theory, thereby guiding generalisation and informing further studies;
- linked qualitative studies may provide evidence of problems with the intervention at an early stage, which may lead to revisions of study protocols (Murtagh et al., 2007).

Taking these considerations into account, I shall go on to describe a case-study for evaluating a complex patient safety intervention – the evaluation of the *Safer Patients Initiative* (**SPI**). However, before I do, a tenet of good practice is to fully describe an intervention to the

extent in which it can be fully replicated (Medical Research Council, 2008); something I shall do in the proceeding chapter.

# CHAPTER 3: CASE-STUDY - THE SAFER PATIENTS INITIATIVE

In the previous chapter, I detailed a literature review of the types of study design that are currently being used to evaluate patient safety interventions and I will go on to describe an evaluation of a highly complex patient safety intervention. However, I firstly need to describe the intervention being evaluated – the **SPI**. I will start by giving an overview of the provenance and the philosophy behind the **SPI**. I will then go on to describe the selection process of the first phase (***SPI*$_1$**) intervention hospitals before providing a full description of the intervention itself. Finally I will outline the second phase (***SPI*$_2$**) of the intervention and how it differed from ***SPI*$_1$**.

## 3.1 Introduction

The **SPI** was funded by the Health Foundation (Health Foundation, 2010a), an independent charity working to improve quality and safety within the UK as well as abroad. It does this through several work programmes including: demonstration projects (large-scale projects that test out new ideas for improving the quality of healthcare); a leadership scheme (investing in individuals from a wide range of organisational and professional backgrounds to help them deliver effective healthcare); and research and development (building knowledge of what works to improve quality and performance). These programmes are supported by an endowment, which was worth £675 million in 2009 (Health Foundation, 2010b).

The Health Foundation worked with the Institute for Healthcare Improvement (IHI) (Institute for Healthcare Improvement, 2010a) in developing and delivering the **SPI**. The IHI is a US-based not-for-profit organisation dedicated to improving healthcare. The **SPI** was developed under the premise that a good evidence-base existed on how to improve the safety of healthcare but an implementation gap existed. The purpose of the **SPI** was to "transform organisations to deliver safer care" (Health Foundation, 2009) by developing a strong safety culture and fostering a leadership that placed safety as a strategic priority. It had a number of features similar to the well publicised US "Saving 100,000 lives campaign" (Berwick et al., 2006; Bisognano et al., 2005); setting out to penetrate deeply into organisations, changing not only specific processes and standards, but also the attitudes, motives and behaviours of staff and how they understand the nature of their work.

***SPI1*** was mentored by the IHI over an 18-month period starting in January 2005. Hospitals were expected to embed and spread learning following the IHI mentored component. Funding of £775k (€860k; $1.2 million) per hospital was provided by the Health Foundation; this was used to secure the services of the IHI and to provide the capacity for change in the individual hospitals.

The aims of the **SPI** hospitals were to:

- achieve a 50 percent reduction in adverse events (Health Foundation, 2006; Shirley, 2008);
- develop an engaged leadership who made safety a strategic priority;

- develop individuals in safety practice methodology;
- establish an appropriate measurement system, conduct ongoing assessments of organisational safety and act upon those assessments;
- create a patient safety centric organisation;
- disseminate knowledge and expertise (Institute for Healthcare Improvement and Health Foundation, 2005a).

## 3.2 Selection of participating $SPI_1$ hospitals

The four **SPI** hospitals that participated in the first phase ($SPI_1$) were selected following a competitive bidding process to demonstrate that they would be receptive to the intervention. A review panel, including members with an international perspective as well as safety, clinical, and organisational expertise, was convened by the Health Foundation to select the hospitals. The panel used a three-stage 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 hospital visits to explore:

- the information outlined in the hospitals' applications;
- the feasibility and sustainability of the initiative 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 3.1.

**Table 3.1: Hospitals that participated in *SPI*$_1$**

| Hospital | Rural/ Urban | Bed No | Teaching status | A&E | ICU | Consultant (specialists) FTE |
|---|---|---|---|---|---|---|
| Hospital A | Urban | 625 | Associate teaching hospital | Yes | Yes | 112 |
| Hospital B | Rural | 750 | No | Yes | Yes | 120 |
| Hospital C | Urban | 903 | Principal teaching hospital | Yes | Yes | 242 |
| Hospital D | 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. The hospitals have not been named as it was agreed that anonymity would be ensured as part of evaluation participation.*

## 3.3 The change package

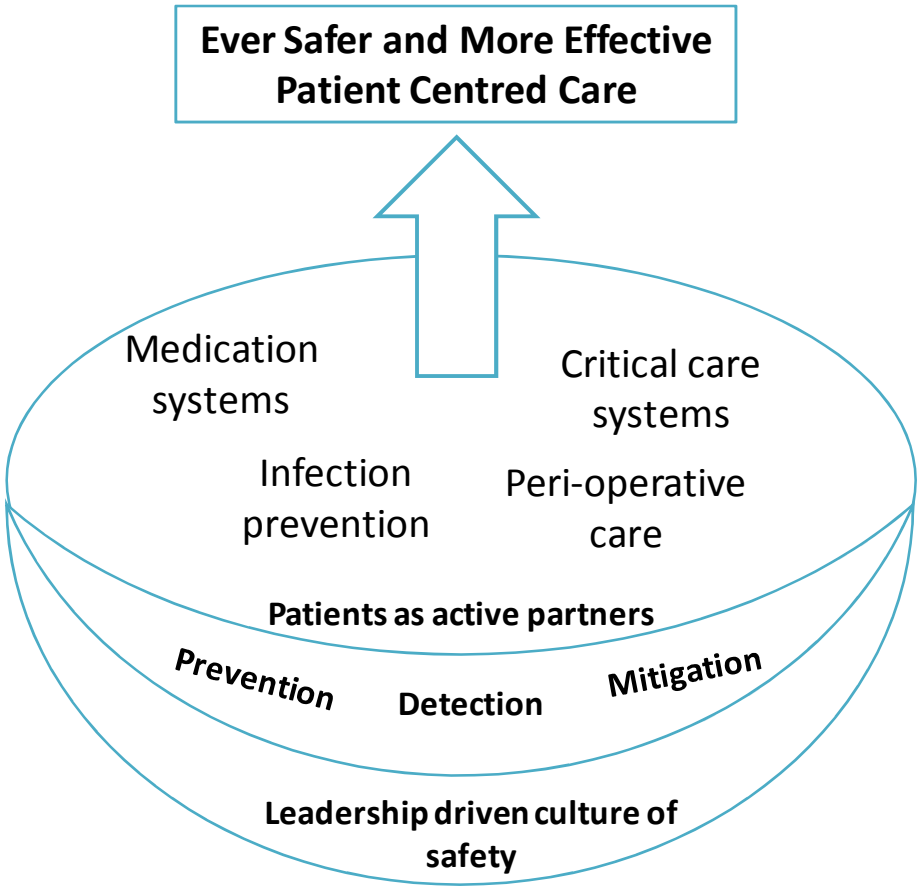
The approach taken to meet **SPI** aims was the IHI's patient safety "change package". This was a three-tiered model (Figure 3.1) in which a leadership driven culture of safety, supported the strategies of prevention (processes that avoided harm and errors), detection (processes rapidly highlighting errors) and mitigation (processes that captured or softened the impact of errors) to improve clinical outcomes.

To achieve this, **SPI** hospitals used case-note review to discover failings of care and then made evidence based improvements in these areas. These improvements were described as change concepts:

"*a general idea, with proven merit and a sound scientific of logical foundation, that can stimulate specific ideas for changes that can lead to improvement*" *(Institute for Healthcare Improvement and Health Foundation, 2005a)*

Change concepts included: standardisation, simplification, protocols/checklists, the use of forced functions, a reduction in the reliance of vigilance, and a reduction of handover points in the healthcare system. It was intended that organisations would use Plan-Do-Study-Act cycles (PDSA) (see 3.6.3.1) to measure and improve practice over time.

**Figure 3.1: The Health Foundation's SPI Patient Safety Change Package. Taken from the kick off meeting (Institute for Healthcare Improvement and Health Foundation, 2005a)**

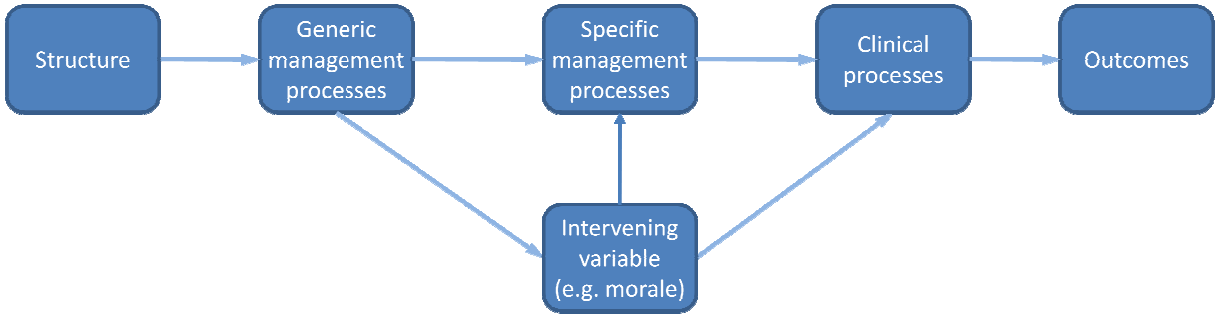


The IHI made the strong claim that change concepts taken together would produce a substantial reduction in severe events (Institute for Healthcare Improvement and Health Foundation, 2005a).

## 3.4 A model of the healthcare system

A conceptual model in which the **SPI** operates in is shown in Figure 3.2. It draws heavily on Donabedian (Donabedian, 1980), who distinguished between structure, process and outcome, and Reason (Reason, 2000), who wrote of latent errors (those that lie in the design of a system), and active errors (those lie with front-end users of the system). Behind both these concepts lies the generic idea of a service (frontline healthcare) embedded in a system.

**Figure 3.2: A model of the healthcare system**




The model starts with structure; which is defined as external factors that are outside the control of managers of a particular healthcare organisation, such as national directives, and licensing procedures. This is sometime referred to as the macro-level.

Next are the internal processes, which are within local control. These processes can be defined as: generic management/organisational processes (e.g. human resource policy, staff training, supply chain management), specific management process (e.g. such as notices to remind staff to wash hands), and clinical (frontline) processes (e.g. adoption of particular evidence-based practices; the quality of clinical-patient communication). This distinction concurs with Reason's distinction between latent errors at an organisational (meso) level and active errors which involve direct human interaction (Reason, 2000) at the micro level. Interventions focused on management processes, such as human resource polices, e.g. staff appraisal and WalkRounds (see 3.6.1.1) will generally effect intervening variables, such as staff behaviours/attitudes, morale, culture and sickness absence. Lastly there are outcomes, which can be clinical or patient reported such as satisfaction.

A systems-wide approach in which this model is considered as a whole is useful in the development phase of a new intervention, and it is imperative that once developed new interventions are studied at all levels (Brown and Lilford, 2008).

### 3.4.1 Framework for the SPI

The **SPI** can be described as a highly complex intervention (Craig et al., 2008); it consists of many components that together comprise what the IHI calls the "change package" (see 3.3). It comprises of many components that target different structures and processes in a hospital. These components can be linked to extended version of the causal chain described in section 3.4. Conceptually the components of the intervention fall into two categories; those that generically strengthen the system to reduce adverse events whatever their cause;

and those that have specific targets (Figure 3.3). These categories and component parts of the **SPI** are outlined in Table 3.2.

**Figure 3.3: Embellished Donabedian's causal chain (blue) linking SPI (green) to structure, process and outcomes**

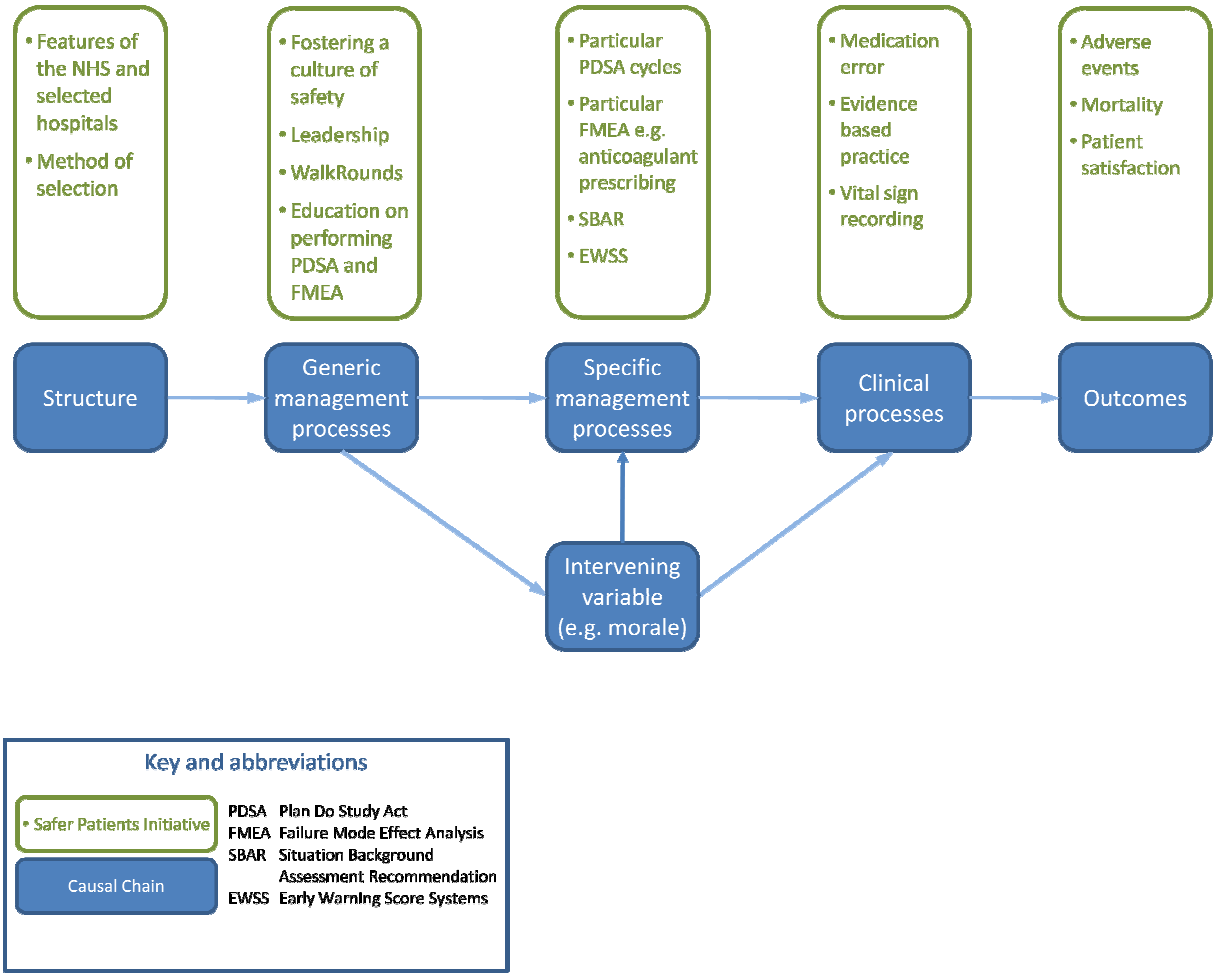


**Table 3.2: Key elements of the SPI**

| ***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.<br><br>Education so that organisations can identify problems and develop and evaluate methods to reduce risk.<br><br>Fostering an understanding of the principles of safe practice. | a. Collaborative residential learning sessions with IHI faculty (see 3.5.6);<br>b. Web-based learning and hospital visits from IHI (see 3.5);<br>c. Leadership projected in part by management WalkRounds (see 3.6.1.1);<br>d. Know-how for Plan Do Study Act (PDSA) cycles (see 3.6.3.1);<br>e. Electronic information sharing facility – for example to share results of PDSA cycles (see 3.5 and 3.6.1.3);<br>f. Participation in safety culture surveys using the Sexton tool (see 3.6.1.2) |
| ***Specific targets*** | |
| 1. Identifying and responding to deteriorating patients. To reduce:<br>Need for "crash calls";<br>Avoidable mortality. | a. Review of 50 deaths (see 3.5.1);<br>b. Tools for monitoring patients' condition and for triggering action. These tools include a proforma to record vital signs and other salient information (the Early Warning Score System [EWSS]) (see 3.6.2.1);<br>c. Promote the use of risk (severity) scores (see 3.6.2.1);<br>d. Establishing a rapid response/outreach team (see 3.6.2.1) |
| 2. Reducing medication error. | a. Medication safety assessment (involving staff in Failure Mode Effects Analysis [FMEA]) – educating staff to identify and remedy weak links in medication practice from prescribing to administration and monitoring (see 3.5.1 and 3.6.3.4);<br>b. Tool to reduce adverse events to anti-coagulant therapy (see 3.6.2.3);<br>c. Education to improve medicine reconciliation on admission (see 3.6.2.3). |

| | |
|---|---|
| 3. Communication between staff to reduce adverse events/mortality. | a. SBAR (Situation Background Assessment Recommendation) tool to ensure that information is communicated in a structured way (see 3.6.2.2);<br>b. Safety briefings – briefings at shift changes to ensure staff are aware of relevant information for patients (see 3.6.2.2). |
| 4. Infection control including Methicillin-resistant Staphylococcus Aureus. | a. Peri-operative antibiotics to reduce surgical site infection (see 3.6.2.5);<br>b. Catheter insertion and maintenance drill to prevent central line infections in intensive care (see 3.6.2.6);<br>c. Following the tenets of ventilator guidelines (bundles) to reduce ventilator acquired pneumonia, venous thromboembolism and stress ulcers in intensive care units (see 3.6.2.6);<br>d. Improve hand hygiene, for example, by means of prominently displayed posters. (see 3.6.2.4) |

I will now go on to describe the **SPI** in more detail, which was compiled using the following:

- observations whilst attending "learning sessions" (see 3.5.6);
- material presented at "learning sessions";
- materials found on the extranet for the **SPI** and the IHI website;
- original or alternative sources when further clarification was required;
- conversations and sharing of this description with the Health Foundation and the IHI.

Section 3.5 describes how the **SPI** was delivered, structured, presented and supported. Section 3.6 details the contents of the **SPI** and what was presented at the learning sessions.

## 3.5 The Safer Patients Initiative – format and structure

Phase I of the **SPI** enabled each hospital to develop its own bespoke patient safety programme under the guidance of the IHI. The structure of this collaborative model was as follows:

- a self assessment of each organisation prior to the "kick-off" meeting (see 3.5.1);
- a four day "kick-off" meeting (see 3.5.6);
- followed by three, two day "learning sessions" (see 3.5.6);
- support of IHI experts;
- hospital visits by IHI experts;

- a virtual community, which included, conference calls, email, listserv and monthly progress reports using web-based systems.

Learning sessions were comprised of plenary presentations (attended by all participants described in 3.5.5), small group discussions and team meetings. They were held over two to four days and enabled participants to gather information on patient safety and on IHI ”improvement” methodology. The sessions were also an opportunity for participants to develop detailed improvement plans for their organisations, and share information/experiences with each other.

The intervals between each meeting were described by the IHI as “action periods”. These were times of intense activity occurring on hospital, in which learning was consolidated and organisations worked towards the goals that they had set (Institute for Healthcare Improvement and Health Foundation, 2005a).

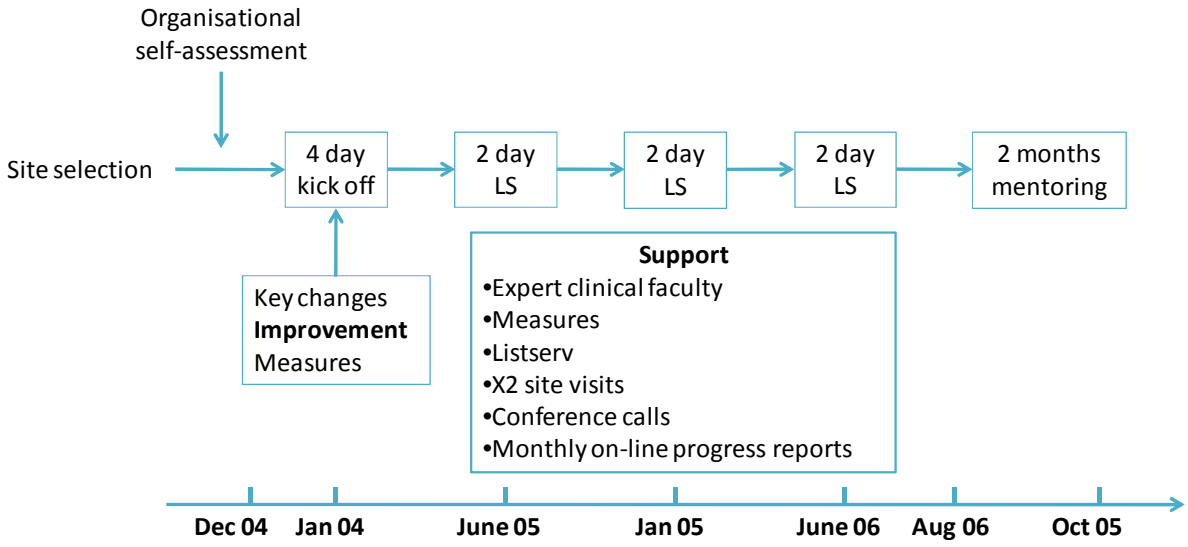
During these periods each hospital remained in contact with the IHI and other collaborative partners. The communication took the form of conference calls, listserv and sharing of information on an extranet site.

The conference calls were facilitated by the IHI with the purpose of problem solving, sharing experiences of improvements and a forum to generate new ideas. The listserv was a mailing list to which all participants subscribed. It allowed subscribers to post questions, share experiences and have discussions outside of scheduled calls and meetings.

The extranet was a website that could only be accessed with a valid username and password. Materials from meetings, monthly reports and documents developed by hospitals were posted and shared here. It also contained contact details of all team members and archived listserv messages. At the beginning of each month hospitals were expected to submit progress reports and results.

An outline of how the different aspects of the **SPI** fitted together and when they were delivered can been seen in Figure 3.4.

**Figure 3.4: The Phase *SPI*$_1$ time-line (adapted from IHI learning materials) (Institute for Healthcare Improvement and Health Foundation, 2005a)**




Learning sessions (LS)

### 3.5.1 Performing a safety self-assessment

Prior to the kick-off meeting (see 3.5.6) each hospitals performed a base-line self-assessment in the following areas (Institute for Healthcare Improvement and Health Foundation, 2004):

- mortality diagnostic case-note review;
- adverse event trigger tool case-note review;
- medication safety self assessment survey;
- healthcare acquired infection (HCAI) rates.

The purpose of each of these exercises was to emphasise any underreporting of adverse events and to provide a baseline measurement.

#### *3.5.1.1 Mortality diagnostic case-note review*

This was a case-note review of fifty consecutive deaths using a structured proforma. This proforma was developed by the Modernisation Agency and is an adaptation of an IHI 2x2 matrix tool (Institute for Healthcare Improvement, 2003). The objective of this exercise was to develop an understanding of the systems in healthcare and points at which failures occur.

#### *3.5.1.2 Adverse event trigger tool case-note review*

Using an adverse event trigger tool, each hospital was asked to review the case-notes of fifty randomly selected discharges with a minimum 24 hour length stay. The allocated review

time was 20 minutes. Again, the objective was to identify sub-standard care that could be remedied.

#### *3.5.1.3 Medication safety self assessment survey*

Each hospital was asked to perform a medication safety self assessment using an IHI tool.

#### *3.5.1.4 HCAI rates*

The hospitals were asked to collate any collected data from the previous 12 months on nosocomial infections in the following areas: ventilator acquired pneumonia; central line catheter bloodstream infections; surgical site infections; and methicillin-resistant staphylococcus aureus (MRSA).

### 3.5.2 Identifying best practice

Hospitals were encouraged to assemble and submit information on best practice that already existed within the organisation. Possible areas included:

- management of narcotics or anticoagulants;
- reconciliation of medications during admission or discharge;
- management of medicines during prescribing, dispensing and administration;
- assessment and response to the deteriorating patient;
- prevention or management of hospital acquired infection;
- medical devices management.

Any areas of confusion in the preparatory assignment were collated by the IHI prior to the first session and were addressed at the meeting.

### 3.5.3 Getting connected to resources

An important component of the initiative was to put evidence into practice, and hospitals were asked prior to the kick-off meeting to identify a person that had access and knowledge of healthcare websites such as the National electronic Library for Health, the British Medical Journal and the National Patient Safety Agency. Also all participants were required to subscribe to the IHI extranet prior to the meeting.

### 3.5.4 Preparing a storyboard

Storyboards were used as a method of sharing experiences and accomplishments. Hospitals were expected to present a storyboard at each learning session.

### 3.5.5 Participants

Hospitals selected staff ("change agents") to participate in the first meeting and to take a leading role in the **SPI**. Hospitals were advised to appoint a leadership team and frontline pilot teams from the following clinical areas: medicines management; intensive care; general ward; and peri-operative care.

The leadership team comprised of senior executives who could also serve as a "system leader" of frontline pilot teams (see below).

Each pilot team comprised of five to six members, of which four core members would attend learning sessions (Institute for Healthcare Improvement and Health Foundation, 2004). The frontline pilot teams were responsible for developing, testing, and implementing specific interventions. Each team comprised of a system leader, clinical/technical experts and day-to-day leader. More than one person could fill these roles and any individual could fill more than one role.

The system leader needed to hold influence/authority within the organisation to implement change and achieve the aims of the frontline pilot teams. It was important that they were influential in the areas that were affected by the change. They were also expected to disseminate changes throughout the organisation (see 3.6.1.4).

Clinical/technical experts were "opinion leaders" (individuals sought out for advice, who are not afraid to try changes) (Institute for Healthcare Improvement and Health Foundation, 2004) that thoroughly understood the roles, functions, operations and processes in the pilot area. They were described as "champions" interested in driving change but who still had good working relationships with colleagues and with day-to-day leader(s). Clinical/technical experts could be a nurse, doctor, pharmacist or a key lead within the pilot area.

The day-to-day leader had the task of directing, and motivating the frontline pilot team. The leader needed to ensure that appropriate data were collected and changes evaluated. The leader needed to understand the processes well enough to predict the effects of making

changes, and be able to work effectively with the system leader and the clinical/technical expert.

At each hospital, an individual was also responsible for submitting monthly progress and data reports. They needed to co-ordinate with the day-to-day leader of each frontline pilot team, and submit reports on behalf of the whole hospital.

### 3.5.6 Kick-off meeting and learning sessions 2, 3 and 4

The kick-off session was held in Bradford on 25th January 2005 (Institute for Healthcare Improvement and Health Foundation, 2005a) and lasted for four days. At the kick-off meeting the **SPI** was described in its entirety.

The meeting was launched with an orientation session in which the Health Foundation, the IHI and the hospitals were introduced to each other. This was also an opportunity for the hospitals to present their self-assessment data.

Subsequent learning sessions were 2 days long and took place in May 2005, November 2005, and June 2006 (Institute for Healthcare Improvement, 2005 (Institute for Healthcare Improvement and Health Foundation, 2005b; 2005c; 2006) The materials presented were a consolidation of those presented in the kick-off session, with the addition of emerging problems, and discussion on how to overcome barriers. In the final session, there was an emphasis on how to disseminate learning and achieve the aim of reducing adverse events by 50% by November 2006.

Learning sessions were an opportunity for hospitals to meet, share, and to discuss progress and hurdles. It was a chance for teams to have time away from their normal work duties, so that they could become focused and motivated on the tasks required of the initiative. Learning sessions were also an occasion to socialise with other hospitals and faculty (IHI, Health Foundation and guest speakers).

The learning sessions followed the same basic format. They started with a plenary session to report progress, present results, share achievements and discuss obstacles encountered. The plenary sessions were intermingled with breakout sessions on leadership, medicines management, critical care, peri-operative management, general ward and measurements. At the end of the first day teams were allocated time for hospital team meetings. The format was repeated on the second day and the session was completed with agreement on tasks to accomplish in the next action period.

## 3.6 The Safer Patients Initiative – the content

As previously mentioned the components of the intervention can be divided into generic and specific categories (Table 3.2). The generic components lend themselves to strengthening leadership within the organisation and are described in section 3.6.1, while those components that seek to improve front-line activities, are detailed in section 3.6.2. Some activities, particularly measurement, are both generic and specific activities. Thus, knowing how to do a PDSA cycles could generically strengthen the system, while the implementation

of these cycles to improve a particular area of practice is a specific front-line activity. These activities with generic and specific components are described separately in section 3.6.3.

### 3.6.1 Enhancing leadership and other generic components

Leadership was placed as being central to the success of the initiative. The leadership team (see 3.5.5) needed to make an initial assessment of patient safety issues at an executive level by reviewing board meetings. Learning from this exercise the leadership team were expected to make safety a priority, set a "patient safety" organisational tone, and establish strategic safety goals. To keep in touch with safety issues occurring at the front-line they were required to participate in WalkRounds (see 3.6.1.1) and to orchestrate a shift from a "blame culture" to a "patient safety" culture, which would be monitored by a safety climate tool (see 3.6.1.2).

The leadership team needed to assure the success of the **SPI** by aligning the goals of the initiative with that of the hospital. Changes needed to be sustained and become an integral part of the system of care. This team were expected to become actively involved in projects by removing barriers, increase resource and by providing the necessary infrastructure. They were expected to champion successful projects and develop robust dissemination strategies (see 3.6.1.4).

#### *3.6.1.1 WalkRounds*

A WalkRound was a method for executives to learn about patient safety concerns of front-line staff, and to reinforce patient safety culture (Frankel et al., 2005). A weekly WalkRound was expected to take place at each hospital and convened at the convenience of front-line staff. To keep conversations focused on patient safety executives were given a script of the type of questions that should be asked:

- "how will the next patient be harmed in your area?"
- "how does this environment fail you?"
- "the last patient hurt of a failure in care – what happened?"

During these sessions, a safety lead was expected to be present, to act on issues raised and a separate scribe to minute the conversation. The collected information was analysed and reported to the executive team, so that priorities could be assigned and plans for action decided. The outcomes (including no actions) needed to be reported back to the originator of the concerns so that interest was not lost by front-line staff.

#### *3.6.1.2 Culture*

Culture was defined as the values, behaviour and norms of a group. It was stressed that when a culture was "good" then the objectives of an organisation are more likely to be met. Patient safety culture was measured by the Sexton safety climate tool (Sexton et al., 2006). This was a 19 question survey that measured the "culture of patient safety" from three perspectives: organisational; leadership and team. The hospitals were asked to distribute the

survey instrument to all staff that had significant patient exposure, i.e. those working in a clinical area for at least four weeks and at for at least 20 hours per week. There was a recommendation that staff complete and return the survey to the co-ordinator within a week of receipt. The surveys were then sent to the IHI for analysis.

#### *3.6.1.3 Leadership and measurement*

A key aspect underpinning the intervention was that of measurement and data. All participants were expected to clearly understand the measurement philosophy but leadership were particularly required to do so. There was belief that if leadership became involved in measurement they would become empowered to lead an organisation that delivered better care. They were expected to put in place processes to accurately, and consistently collect data on a wide and balanced number of measures, principally those relating to organisational performance, which overtime could also be used to assess the sustainability of improvements and assist in improvement dissemination strategies.

Hospitals needed to decide which key performance measures they wanted to use and were guided by the Institute of Medicine (IOM) six quality dimensions (Table 3.3).

**Table 3.3: Measurements to track IOM dimensions (STEEP), taken from materials presented at kick-off session (Institute for Healthcare Improvement and Health Foundation, 2005a)**

| IOM Dimension | Measures |
|---|---|
| Safe | Adverse drug events (out and inpatient)<br>Work days lost (inpatient) |
| Timely | Third next appointment available (outpatient) |
| Effective | Hospital specific mortality rate (inpatient)<br>Functional outcomes: SF-6 (out and inpatient) |
| Efficient | Costs per capita (across region)<br>Hospital specific standardised reimbursement |
| Equitable | *Note: The sixth IOM quality dimension is "equitable" care. This will be addressed by stratifying the above measures, when possible, into subpopulations. For example, many of these measures could be stratified by gender, age, or racial groupings using available data. |
| Patient-centered | Patient satisfaction (out and inpatient)<br>Percent patients dying in hospital (across region) |

Once the measures were decided a target needed to be established (within a realistic time-frame) and an implementation strategy developed to reach these targets. Hospitals needed to make monthly measurements of the selected indicators

The hospitals were encouraged to collaborate and share their results, in a league with other hospitals that were involved with other quality improvement programmes developed by the IHI [e.g. Pursuing Perfection (Berwick and Rothman, 2002) and IMPACT (Henderson et al., 2003)].

#### *3.6.1.4 Dissemination*

The IHI had developed a framework for dissemination (Nolan et al., 2005; Rashad et al., 2006) based on the literature and interviews with organisations both within and outside of healthcare. Hospitals were expected to plan and initiate dissemination strategies of

successful projects as early as possible. These strategies needed to have clear aims and be multipronged. The strategies were requires to describe intended improvements, key performance indicators, communication methods, intended target population, time-frame of effort and an anticipation of training needs.

Day-to-day leader needed to be involved in dissemination plans, as well as an executive. Projects which had a greater chance of being accepted were those that had achieved the desired outcome, that had an advantage and that were compatible with the adopter hospital.

Hospitals were taught diffusion theory (Rogers, 2003), i.e. that in a population there are different categories of adopters, which are: innovators, early adopters, early majority, late majority, and laggard.

Innovators (2% of the adoption population) were the first individuals to adopt an innovation, and were willing to take risks - they were close to scientific sources and more likely to interact with other innovators (Rogers, 2003). Early Adopters (14%) were the second fastest category of individuals to adopt an innovation. These individuals held influence and had the highest degree of opinion leadership amongst the other adopter categories; the support of these individuals was essential for successful dissemination. The Early Majority (34%) would adopt an innovation after a varying degree of time but the time was significantly longer than with the innovators and early adopters. Late Majority (34%) individuals represented the average member of the population, and these individuals approached an innovation with a

high degree of skepticism. Laggards (16%) were the last to adopt and typically had an aversion to change and tended to focus on “traditions” (Rogers, 2003).

### 3.6.2 Improving specific front-line interventions

Specific interventions were developed for four clinical areas: general medicine, peri-operative care, critical care and medicines management. Those that targeted the general ward included better monitoring of patients using early warning score systems (see 3.6.2.1), structured methods of communicating this deterioration to doctors and outreach services (see 3.6.2.2), and the use of safety briefings at shift changes (see 3.6.2.2). Interventions developed to reduce medication error included the targeting of high risk drugs, principally anticoagulants and the development of tools to ensure that on admission (or any handover point) drugs were continued (or discontinued) appropriately, i.e. medicines reconciliation (see 3.6.2.3). Interventions to be used in critical and peri-operative care primarily targeted the prevention of hospital acquired infections (see 3.6.2.5 and 3.6.2.6).

#### *3.6.2.1 Identifying and responding to deteriorating patients*

At risk patients within a hospital could reside on general wards, and are not necessarily supported by the skills, knowledge and attitudes of those delivering services in critical care (Institute for Healthcare Improvement and Health Foundation, 2005a). Outcomes for these patients including increase length of stay and mortality, which are potentially avoidable through diligent monitoring of vital signs and appropriate response when deterioration occurs (Institute for Healthcare Improvement and Health Foundation, 2005a). Therefore,

according to the IHI, an effective programme was one able to identify these patients early, and provide appropriate care, in the correct location and in a timely manner by either early admission to a critical care unit or continued care on the ward reinforced where necessary by ICU outreach services (Institute for Healthcare Improvement and Health Foundation, 2005a).

A tool developed to detect the early deterioration of patients is the Early Warning Score System (EWSS) (McGaughey et al., 2007) and its use was advocated by the IHI. It is a simple algorithm based on the following bedside observations:

- staff member worried about the patient;
- acute change in heart rate to <40 or >130 bpm;
- acute change in systolic blood pressure to <90 mmHg;
- acute change in respiratory rate to <8 or >30 breaths per minute;
- acute change in pulse oximetry saturation to <90% despite oxygen;
- acute change in conscious state;
- acute change in urine output to <50 mL in 4 hours.

These observations were recorded onto a colour banded patient chart (to provide a visual cue) and assigned a score. If the total score exceeded a threshold level it initiated an outreach team.

The outreach approach defines critically ill patients by need or level of care required, and not by geographical boundaries of where the patient is located. The team prevents admissions to the ICU, facilitates more timely ICU discharges and enables critical care skills, knowledge and expertise to be shared.

#### *3.6.2.2 Effective communication of patient safety concerns*

Communication failure has been estimated to be a major factor in 60-70% of serious patient safety incidents (Greenberg et al., 2007; Patterson et al., 2004) and hence an improvement in communication was an **SPI** focus. There were two methods employed by the IHI to enhance patient safety communication: the first was set at team level and utilised safety briefings; whilst the second involved a structured mechanism for imparting patient information between individuals.

##### *3.6.2.2.1 Safety briefings*

The use of safety briefings originates from the aviation industry and they enabled front-line staff, without fear of reprisal, to share information about potential safety problems and concerns on a daily basis (Institute for Healthcare Improvement, 2010b). It was recommended that they were five minutes long, and that they took place at the start and end of each shift. The briefing at the start of the shift posed open questions on safety awareness, whilst the session at the end of the shift asked specific questions of safety incidences that may have occurred during that shift. All responses were recorded and it was expected that those requiring further action would be acted upon.

Initially safety briefings were conducted with nurses and then over time expanded to include all of the multi-disciplinary team. Moreover, it was intended that the drive for these briefings was from front-line staff themselves with little or no impetus from management.

#### *3.6.2.2.2 Situation Background Assessment Recommendation (SBAR)*

In other high risk industries, such as aviation, rail and the military, standardised communication is used so that information is shared consistently and imparted accurately. SBAR (see Figure 3.5) is a tool that was developed by the US Navy for standardising important and urgent communication in nuclear submarines, and has been adopted by the IHI as a model that health providers should use in clinical communication. The advantages are that hierarchies are overcome; junior staff are less likely to miss out important information and the recipient knows in which order information will be given (Lingard et al., 2005; Marshall et al., 2009).

**Figure 3.5: SBAR tool (taken from (Marshall et al., 2009)**

| | | |
|---|---|---|
| S | Situation<br>State purpose: "The reason I am calling is ....." | |
| | If urgent – say so | e.g. "Thus is urgent because the patient is unstable with a BP of 90". |
| B | Background<br>Tell the current problem | |
| | Relevant:<br>History/examination/test results<br>Management | If urgent:<br>Relevant vital signs<br>Current management |
| A | Assessment<br>State what you think is going on | |
| | e.g. "So the patient is febrile and I can't find a source of infection". | Urgent, e.g. "the patient seems to be deteriorating, I think they may be bleeding" |
| R | Request<br>State request | |
| | e.g. "I'd like your opinion on the most appropriate test" | e.g. "I need help urgently, are you able to come?" |

#### *3.6.2.3 Reducing medication error*

##### *3.6.2.3.1 Medicines reconciliation*

An area identified as having a particularly high opportunity for medication errors were "handover points" within the healthcare system. Errors include: wrong dose, wrong route, incorrect frequency being administered, and medicines being omitted. To minimise them a review or "reconciliation" of medications should be done at these interchanges.

Reconciliation was described as a three step process: verification – collection of medication history; clarification – ensuring that the medications and doses are appropriate; and

documentation – changes to prescriptions or reason for differences. During the review the following pieces of information were required to be recorded:

- current drugs prescribed by the general practitioner (GP);
- time of last dose;
- source of information (patient/family, patient's pharmacy, previous medical records, packages of patients medication, GP);
- medications ordered on admission;
- if possible document patient adherence;
- any other drugs, e.g. over the counter, herbal remedies, including dose, route and frequency.

Hospitals needed to identify who is best suited to complete the reconciliation. However, the IHI suggest that pharmacists undertake the most effective drug history.

The reconciliation form should be used every time that a prescription is changed, and whenever a patient changes service setting or level of care. Participants needed to identify the handover points that needed to be reconciled, but the most important were admission and discharge.

#### *3.6.2.3.2 Anticoagulants*

The prescribing and management of anticoagulant drug therapy is complicated and was considered an area of special concern for the **SPI** hospitals. Problems identified after the medication audit (see 3.5.1) by hospitals were:

- poor documentation of:
    - the indication for prescribing of anticoagulants;
    - the target International Normalised Ratio (INR);
    - the duration of treatment;
- lack of evidence of GP referral information;
- lack of evidence of patient counselling;
- unconsidered co-prescribing of other interacting drugs;
- lack of guidance for junior doctors on dosing schedule.

Hospitals were expected to perform a Failure Mode Effects Analysis (see 3.6.3.4) to improve core processes associated with prescribing, monitoring and discharging patients on anticoagulation therapy.

#### *3.6.2.4 Infection control and hand hygiene*

**SPI** hospitals were taught the basics of infection control and recommendations to prevent them:

- clean hands (wash hands with soap and water or waterless alcohol products);
- remove lines and drains as soon as possible;
- identify patients with potentially important communicable diseases and epidemiologically important infections and establish procedures to prevent transmission;
- isolate patients with multi-drug resistant or epidemiologically important organisms;
- vaccinate and/or treat patients and healthcare workers for important communicable diseases;
- have a robust and viable local infection control programme (Institute for Healthcare Improvement and Health Foundation, 2005a).

They were informed that patients particularly vulnerable to healthcare associated infections were those with:

- severe underlying disease;
- confinement to a bed;
- poor clinical prognosis;
- prolonged hospital stay;

- prior antibiotic therapy;
- severe burn and open wounds/drains.

#### *3.6.2.5 Peri-operative*

Within the surgical setting the intervention focused on preventing surgical site infections (Ryckman et al., 2009). The following actions were recommended (Institute for Healthcare Improvement, 2007):

- appropriate prophylactic antibiotic;
- control of glucose intra-operatively and 24 hours post-operatively in diabetic patients;
- use of clipping instead of shaving;
- normothermia (maintenance of normal core body temperature).

In addition prophylactic antibiotics were recommended as part of this bundle.

#### *3.6.2.6 Critical Care Systems*

The IHI described an "ideal" ICU model in which there are three groups of people were in play; the leadership team, the care team and patients. Leadership should be focused on infrastructure and provide realistic, but aspirational goals in quality, safety and productivity. The care team are in charge of the delivery of a reliable process, developed on the foundations of best practice. They were required to work collaboratively by including all staff, patients and family. In the centre of these two groups was the patient.

To achieve this model hospitals needed to employ multidisciplinary rounds to review patient status, schedule tests, co-ordinate care, clarify patient objectives and plan the care for that patient for that day. These rounds would enable protocol/guidelines to be delivered consistently across disciplines, highlighting any need for refinement and further education of staff.

Other key changes and measures within the ICU were:

- daily goals;
- ventilator bundles (Bonello et al., 2008; Resar et al., 2005);
- central line bundles (Bonello et al., 2008; Galpern et al., 2008; Pronovost et al., 2006a).

A bundle was defined as being:

*"a grouping of failure mode processes (bundle elements) with approximate time and space characteristics that when done collectively <u>can</u> have an enhanced effect on outcome"* (Institute for Healthcare Improvement and Health Foundation, 2005a).

The bundle consisted of a minimal number of elements that could be measured in a binary manner, (i.e. present, not present). Each element was based on "non-refutable" evidence, and be completed at a similar time and within the same environment. Although initially the emphasis of the bundle was that of process improvement they were ultimately expected to improve outcomes.

The purpose of the "ventilator bundle" was to reduce the rate of ventilator acquired pneumonia in the ICU. There were four elements to the bundle:

- raising the head of the bed;
- an assessment on the appropriateness of further sedation;
- administration of peptic ulcer disease (PUD) prophylaxis;
- administration of deep venous thrombosis (DVT) prophylaxis.

The "central line bundle" was designed to help decrease the incidence of catheter-related bloodstream infections. The key components were:

- a daily review of the necessity of the line;
- strict compliance to hand washing;
- maximal barrier precautions upon insertion, i.e. the individual carrying out the procedure should wear a cap, mask, sterile gown and gloves. The cap should cover all hair and the mask should cover the nose and mouth tightly. The patient should be covered from head to toe with a sterile drape with a small opening for the site of insertion;
- the site of insertion should be made sterile by scrubbing with chlorhexidine (2 percent in 70 percent isopropyl alcohol) skin antiseptic. The skin should be allowed to dry before insertion;
- optimal catheter site selection, with subclavian vein as the preferred site for non-tunnelled catheters.

### 3.6.3 Tools and skills developed

A key concept of the **SPI** was that of measurement. It was believed that the process of collecting information, whilst making small structured changes, would empower and transform hospitals, leaders, and front-line staff into delivering better quality and safer care. The structured methodology employed was the PDSA cycle and this is described in more detail in section 3.6.3.1.

Hospital staff were instructed in the analyses of performance data by means of *Statistical Process Control* and is described in section 3.6.3.2.

Finally hospitals were taught techniques to improve processes. These were: a) Failure Mode Effects Analysis (see 3.6.3.4), a systematic method of revealing weaknesses in a process and; b) Reliability Design (see 3.6.3.3), a means of developing robust and more stable processes through process mapping.

#### *3.6.3.1 Plan-Do-Study-Act (PDSA) cycles*

PDSA is an iterative four-step problem-solving methodology used in business process improvement (Speroff and O'Connor, 2004; Varkey et al., 2007). The first stage is to "Plan" a specific improvement. At this stage organisations must know what they were trying to accomplish. This should state clear aims and quantifiable goals. At the "Do" stage the improvement plan is implemented. The "Study" step requires there to be appropriate measurement systems in place (see 3.6.3.2) to determine if there had been a change, and if

this has resulted in an improvement. The final "Act" stage was a period of reflection to determine what further changes were needed in light of what had been learnt i.e change concepts (see 3.3). A suggested method of developing change concepts was by "divergent, and convergent thinking" used during "nominal group technique" – a structured method of brainstorming.

These stages were described by the IHI as the "Model for Improvement". Hospitals were advised to use repeated uses of the PDSA cycle and use a "1-3-5-all" strategy. Firstly they needed to have a successful prototype phase in which a small number staff executed a change, and determined if the change created improvement. Once it worked well in one setting, it was tested in three, when it worked well in three settings it was tested in five. Once it was worked well in five the change moved into an implementation phase.

The implementation phase worked in a similar fashion:

- 1-3-5-all nurses on a ward
- 1-3-5-all surgeons in a theatre or service
- 1-3-5-all pharmacists
- 1 day – 3 days – 5 days – all days
- 1 shift – 3 shift s – 5 shifts – all shifts

Hospitals were told that this strategy would see an idea move from a thought, to an actual change that could result in improvement, which could then be disseminated to other areas (see 3.6.1.4).

One of the key aspects of PDSA was making appropriate measurements. The data collected during this process was described as being for quality improvement (effectiveness) rather than that for accountability (efficiency) or research (experimental) (Brook et al., 1996). The differences between these types of data can be found in the Table 3.4 (Solberg et al., 1997):

**Table 3.4: The three faces of performance management**

| *Aspect* | *Improvement* | *Accountability* | *Research* |
|---|---|---|---|
| **Aim** | Improvement of care | Comparison, choice, reassurance, spur for change | New knowledge |
| **Methods** | | | |
| Test observability | Test observable | No test, evaluate current performance | Test blinded or controlled |
| Bias | Accept consistent bias | Measure and adjust to reduce bias | Design to eliminate bias |
| Sample size | 'Just enough' data, small sequential change | Obtain 100% of available relevant data | 'Just in case' data |
| Flexibility of hypothesis | Hypothesis flexible, changes as learning takes place | No hypothesis | Fixed hypothesis |
| Testing strategy | Sequential tests | No tests | One large test |
| Determining if a change is an improvement (see below) | Run charts or Shewhart control charts | No change focus | Hypothesis, statistical tests (t-test, F-test chi squared), p-values |
| Confidentiality of the data | Data only used by those involved with improvement | Data available for public consumption and review | Research subjects identities protected |

Three types of measurement were described:

- outcome measure that inform a system or process and describe the level of performance;
- process measures – changes to the system;
- balancing measures, which are signalling measures and usually look at a system from a different direction/dimension, e.g. financial data/intervening variables.

#### *3.6.3.2 Statistical Process Control (SPC)*

SPC is a method of monitoring a process through the use of control charts. The advantage of using SPC is the ability to detect process changes and trends early. Users making pragmatic improvements can move away from simple bar and line graphs, which are mainly descriptive, to a method, which when used correctly is statistically rigorous but requires less data than when testing for significance (Benneyan et al., 2003).

This method was developed by Shewhart (who also developed PDSA cycles) in the 1920's, whilst working at Bell Telephone Laboratories (Shewhart, 1939). Deming, a student of Shewhart, developed a theory of management based on these statistical methods and introduced them to industry in the 1930's (Deming, 1986). These methods were widely adopted by the Japanese industry in the 1950's and provides the foundations of "Six Sigma", a current business management strategy developed by Motorola (Tennant, 2001).

Shewhart discovered through experimentation that physical processes do not behave the same way as nature, i.e. it does not necessarily follow a normal distribution curve. He concluded that processes display two types of variation (Benneyan et al., 2003):

- common cause variation: this is random, unassignable variation that is inherent in the design of a process and is associated with a stable process. This is the variation that is expected to occur according to the underlying statistical distribution if parameters remain constant;
- special cause variation, is uncontrolled/unnatural variation, not inherent in the design of a process.

Central to SPC is the plotting of control charts. These are similar to run charts, i.e. a simple line graph of 15-20 data points are plotted against time. However, control charts have the additional features of a centre line which represents the median, and upper and lower process control limits (usually set at ±3 sigma). These additional features enable users to distinguish between common and special cause variation.

There are several kinds of control charts and the type used will be determined by the type of data collected, i.e. continuous (measurement) data or discrete (or count or attribute) data. For continuous data in which there is only one observation per sub group an XmR is appropriate. In instances where discrete data has been collected *p*-charts, *u*-charts, and *c*-charts are more appropriate. Details on how to construct these charts and how to interpret them, using worked examples are described by Mohammed et al., (2008a).

There are a number of rules to determine if special cause variation has occurred. Some are contentious but three are widely accepted as signifying a trend (Mohammed et al., 2008a):

- a run of eight points or more on one side of the centre line;
- two out of three consecutive points greater than 2 sigma on the same side of the centre line;
- a run of eight or more points following an upwards or downward trend.

The use of control charts reveals which type of variation is occurring but not the cause of the variation, which requires further investigation. For example, special cause variation can be deliberate and be part of an improvement strategy or it could be due to something unforeseen, such as a staff shortage. On the other hand common cause variation is not inherently good; a process maybe stable but can still be delivering poor care. Once the cause of variation has been determined practitioners needed to determine which course of action is required.

#### *3.6.3.3 Reliable design*

Reliable design stems from reliability and human factors science. It has been used in manufacturing and aviation as a way of designing systems that consistently produce a desired outcome (Nolan et al., 2004). Within healthcare, systems are chaotic, inconsistent and have been developed by trial and error. These systems were described as having a performance of $10^{-1}$ and hospitals were expected to shift their processes to a performance level of $10^{-2}$.

To improve the reliability of a process they were expected to convene a small team (knowledgeable of the process) who would map the procedure in order to identify common/important failures. The team were expected to segment parts of the process in which failures occurred, redesign and test that part whilst keeping other segments unchanged. They were expected to do this until the performance reached $10^{-2}$. In the initial stages change should be expected and users of the process could opt-in to the change. However, as the process became more robust and finalised, users had to provide valid reasons for opting out.

In making process more stable hospitals were expected to use the IHI's three-step model:

- prevention of a failure by standardising processes;
- indentifying and mitigating failures by designing a redundancy function to trap and minimise harm;
- redesigning the process based on the critical failures identified by process mapping.

#### *3.6.3.4 Failure Mode and Effects Analysis (FMEA)*

FMEA is a systematic method of evaluating a process and identifying opportunities for the failure and its likely impact (Institute for Healthcare Improvement, 2010c; Williams and Talley, 1994). Each failure received a numerical score between 1 and 10 for the following:

- likelihood the failure would occur (1 = very unlikely, 10 = very likely)?

- likelihood the failure would be detected (1 = likely, 10 = very unlikely)?
- amount of harm or damage the failure would cause (1 = no harm, 10= harmful)?

The mathematical product of these scores is known as the Risk Priority Number (RPN) and the sum of these gives the RPN for the process. The RPN for the process can be used as a measure of comparison with other processes within the same organisation.

It was recommended that hospitals should first tackle the 10 failures with the highest RPN; use these to plan improvement efforts and to track and monitor over time a reduction of RPN of a process.

In practical terms, hospitals were to recruit a multidisciplinary team involved in the process of interest. The team produced a detailed map of the process, listed all the failure modes, failure causes, failure effects and calculated the RPN.

If the failure was likely to occur the team needed to assess the cause and determine if it could be eliminated; if the failure was likely to be undetected they needed to consider adding processes upstream to identify the failure; and if the failure was likely to cause severe harm they should introduce early warning signs and extra information at the points of care for such events.

## 3.7 The $SPI_2$ intervention phase

To build on the experience and learning from $SPI_1$, a second phase of the intervention, known as $SPI_2$, was rolled out from March 2007 to September 2008 inclusive. $SPI_2$ included a further 20 UK hospitals (ten in England and ten in the other countries of the UK) that were selected following a recruitment process similar to that used for $SPI_1$.

The intervention remained much the same as that which was put in place during $SPI_1$. Again the initiative was mentored by the IHI and was designed to strengthen the organisations generically while putting in place specific front-line activities, such as the introduction of early warning score systems to improve the management of the acutely sick patient, the use of ventilator bundles to reduce ventilator-acquired pneumonia in intensive care, and the introduction of a surgical "bundle" of evidence-based standards to reduce surgical complications. There were five main differences between $SPI_1$ and $SPI_2$ in the overall management of the programme based on experiences gleaned from $SPI_1$ hospitals:

- the hospitals were required to work with a partner organisation (a "buddy system") and encouraged to hold regular meetings between the lead implementation teams (10-12 people) from each hospital. By using this system it was envisaged that hospitals would support each other, share the burden and provide mental support in quickly achieving the goals of the intervention.
- there was a longer period between dissemination of the preparatory materials (December 2006) and the first "kick-off" session where the various teams came

together with IHI to share experiences (March 2007). This gave hospitals more time for planning and developing the intervention and to obtain a baseline measurement in the safety climate survey.

- the financial package was smaller than in the case of $SPI_1$; a mean of £270k (€390k; $430k) per hospital rather than £775k (€860k; $1.2 million).
- there were four learning sessions as with $SPI_1$ but an additional reliability and capability workshop was provided.
- $SPI_2$ sought a 15% reduction in mortality rates; this was not an explicit $SPI_1$ aim.

Specific aspects of the intervention also changed:

- the reduction of adverse event target was revised from 50% to 30% as it was felt that this was a more achievable yet "aspirational" target;
- removal of the routine use of beta blockers in the surgical "bundle" as this clinical standard was contentious in the UK;

### 3.7.1 Selection of participating $SPI_2$ hospitals

As with the selection of the $SPI_1$ hospitals, $SPI_2$ hospitals were selected through a competitive bidding process. A similar format to the phase one selection was followed with initial applications reviewed by an international panel with expertise in patient safety, organisational change and improvement methodology. With applications being assessed against the following criteria:

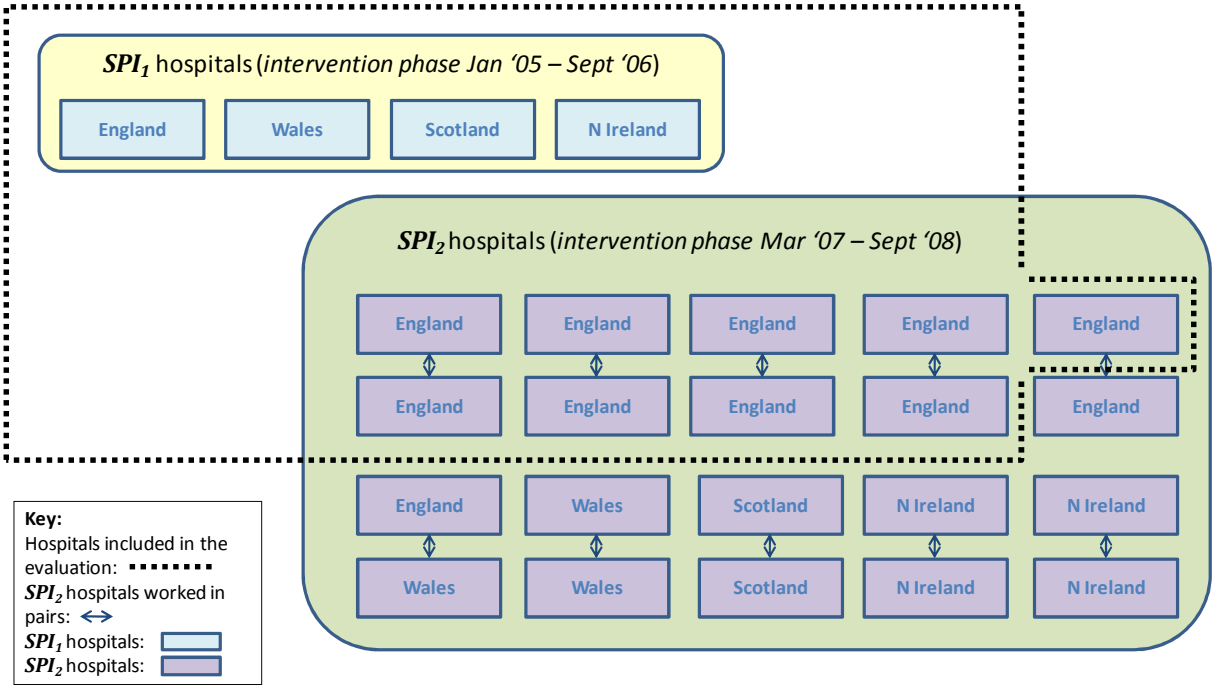
- leadership commitment;
- capacity and capability;
- openness, transparency and communication;
- collaboration.

The short-listed hospitals were subject to an on-site assessment and the final 20 hospitals were chosen by a final selection board.

## 3.8 Summary

The **SPI** comprised of two intervention phases. The first phase (***SPI$_1$***) started in January 2005 in 4 intervention hospitals, one in each of the four nations. The second phase (***SPI$_2$***) consisted of 20 intervention hospitals across the UK and began in March 2007. This is summarised in Figure 3.6. The evaluation took place in all the ***SPI$_1$*** hospitals and 9 English ***SPI$_2$*** hospitals. The rationale for this and an explanation of how control hospitals were selected for the evaluation is described in the following chapter in section 4.3.

**Figure 3.6: A summary of the hospitals participating in the $SPI_1$ and $SPI_2$ phases of the intervention**




The **SPI** intervention itself was focussed on improving the reliability of specific front-line care processes within designated clinical areas (general wards, critical care, peri-operative care and medicines management). Specific interventions targeted particular problems identified by the IHI, including medication error and identifying and responding to patient deterioration. Generically, 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, but capable of being used to pursue the goal of improving patient safety overall – 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 which is based on quality improvement methodology with a long provenance going back to Deming (Deming, 1986).

To evaluate this highly complex intervention, a mixed-method approach has been taken and this is described in Chapters 4 to 15.

# CHAPTER 4: OVERVIEW OF THE EVALUATION

This chapter will provide an outline of the methods used in the evaluation of the **SPI**, which uses a before and after study design with contemporaneous controls. The evaluation comprises of several sub-studies and uses both quantitative and qualitative methods, and relate to the causal chain described in section 3.4. The proceeding chapters will report on each individual sub study following an Introduction, Methods, Results, Discussion format. The final chapter "triangulates" the results (the reinforcement of one finding with a different type, see 2.4.3) and discusses the **SPI** evaluation overall. I will end the thesis with answering the questions of which study design provides a robust method of evaluation of complex patient safety interventions. However, firstly I will give an outline of *what* can be measured and *how* we measure it (see 4.1).

## 4.1 End points and measurements

To determine whether or not a patient safety intervention has been effective, measurements of key objects, events or abstract constructs need to be undertaken as part of an evaluation. These objects, events or constructs are referred to as end points and are discussed below.

### 4.1.1 Patient outcomes

This section will focus on clinical outcomes (e.g. morbidity, mortality, quality of life and patient satisfaction).

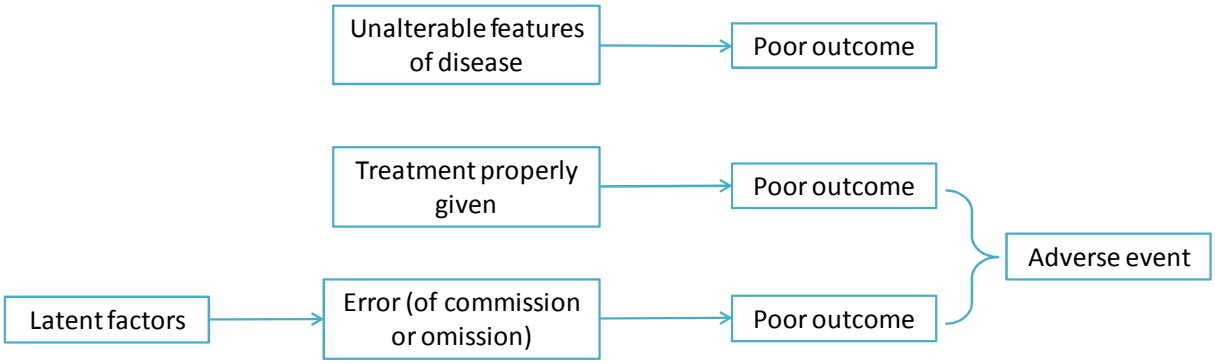
Two issues arise when patient outcomes are used as endpoints:

- the signal, (i.e. improvement caused by the intervention) to noise (stochastic variance in outcome) ratio (i.e. issues of precision);
- defining outcomes (both the numerator and denominator) consistently and minimising case-mix bias.

#### *4.1.1.1 Signal to noise ratio*

A common problem for evaluators is that the patient outcomes that may plausibly be affected by an intervention are also influenced by many other factors, making it an imprecise end point. This is illustrated in Figure 4.1 with errors forming just one component of the model. The figure also demonstrates that an evaluation of a patient safety intervention using patient outcomes is at a higher risk of a beta (or type II) statistical error – the error of accepting the null hypothesis when it is not true.

**Figure 4.1: Intellectual framework to classify links between care and poor outcome**




A way in which to improve the signal to noise ratio is by selecting as the numerator only the cases of poor outcome (harm) that were *caused* by poor care – that is, those resulting from error. This requires identification of poor outcome and then examination of the process of care to identify instances of poor care. Such methods do have limitations, as firstly judgments on preventability of poor outcomes is difficult; secondly there maybe disagreement between different reviewers as demonstrated in the Harvard Medical Practice Study (Brennan et al., 1991), and finally such judgements are affected by hindsight bias (see 4.1.3), where reviewers are skewed by awareness of the outcome (Caplan et al., 1991; Hayward and Hofer, 2001).

#### *4.1.1.2 Minimising bias*

Patient outcomes may be difficult to measure in a consistent way. For example Bruce and colleagues (2001) undertook a systematic review that identified over 40 definitions of two surgical adverse events (wound infection and anastomotic leak). They documented considerable variations in the definitions across organisations. In the context of research, it is important to ensure that the same observers make measurements across organisations

wherever possible to ensure that the results will not be biased due to differences across observers, even when they are applying the same definitions. This principle has been applied to this evaluation (see Chapters 6, 7 and 8).

A comparison of outcomes between organisations may also be affected by difference in case-mix. Patients who are sicker and/or older have more co-morbidities and are at increased risk of both worse outcomes and experiencing more errors due to the requirement for more interventions. This situation leads to case-mix bias in comparative studies even after statistical adjustment for known co-founders (Lilford et al., 2004; Lilford et al., 2003). Bias can be minimised by randomisation and by conducting controlled before and after comparisons, as described in section 2.1.3.

### 4.1.2 Surrogate end points

This section focuses on measurements occurring at the process level of the causal chain (see 4.1.3). Oakley et al., (2006) have argued that including a process evaluation alongside a more traditional outcomes-based approach would improve the science of evaluations of complex interventions. There are three types of surrogate end points considered here: the fidelity with which an intervention is implemented, the effect of the intervention on intervening variables such as morale and the effect of the intervention on clinical error rates.

#### *4.1.2.1 Fidelity*

Fidelity measures whether the intervention was implemented as intended or describes how it might have been adapted. Where the benefit of the intervention is not in doubt, measuring compliance with directive to deploy the intervention is sufficient. In the more usual situation where the benefits of a service level intervention are contested, demonstrating high fidelity at least shows that the distal benefits (in terms of patient outcomes) are plausible and vice versa. Evaluations can make use of this asymmetry in two contexts:

- where direct measurements of safety (clinical errors and/or outcome) can be made with high precision (and accuracy). In this case the fidelity measure may help explain a null result;
- where direct measurements of safety cannot be made with sufficient precision. In this case, showing that the intervention was implemented as planned provides reassurance that the desired effects on safety are plausible

#### *4.1.2.2 Intervening variables*

Intervening variables are described in section 3.4 and tend to be the target for more diffuse interventions aimed at strengthening an organisation generically (Shortell et al., 2000; Wagner et al., 2001). For example, improved human resources (HR) polices (such as staff appraisal) are expected to impact on errors by means of effects on staff motivation, and morale and reduced sickness. If the intervening variable can be measured they can be used

as surrogate end points in evaluations of complex interventions. It may thus be informative to measure the effects of such HR polices on these intervening variables. However, as with fidelity measurement, changes in surrogate measures do not prove that the "down-stream" clinical effects will be realised.

#### *4.1.2.3 Clinical process measurement: error*

##### *4.1.2.3.1 What is an error?*

Clinical error, which causes or may cause an adverse outcome is both logically and chronologically the closest surrogate to patient outcomes (Figure 3.2). Clinical error can be broadly defined as:

- the failure to apply the correct standard of care (an error of omission);
- the failure to carry out a planned action as intended (an error of omission); or
- the application of an incorrect plan (an error of commission) (Reason, 2000).

Sometimes the error and the outcome are the one and the same, for example, wrong site surgery. In other instances the error and the poor outcome occur with the same frequency such that the problems of rarity resurface. In most cases, however, the error is considerably more common than the corresponding adverse event, because many errors, even if carried through to the patient, will not be harmful (Dean et al., 2002; Taxis and Barber, 2003). Where errors are markedly more common than contingent adverse events their relative frequency affords greater precision in measurement.

There are three methods employed to measure error: reporting systems; trigger tools and error rates, which are outlined below.

#### *4.1.2.3.2 Reporting systems*

Reporting systems may be compulsory or voluntary (Boxwala et al., 2004). Denominators are not of the whole population (Pronovost et al., 2006b; Vincent, 2007), and tend to be fitted later such as a unit of time or a population of patients. Problems with this methodology include:

- differences between study units in the number and type of patients mean that there is a high likelihood of bias in an any comparative study based on reporting and retro-fitting a denominator risks a data driven (self-fulfilling) hypothesis;
- incomplete reporting which is highly selective (Begaud et al., 2002; Cullen et al., 1995; National Audit Office, 2005);
- any change in reporting may reflect propensity to report rather than a true change in the underlying problem. Moreover, a tenet of safe practice is a culture of open reporting.

Despite these problems with reporting systems reported events are still of value. National-level data from reporting systems are useful for identifying priorities for patient safety interventions or detecting previously unsuspected hazards such as side effects of drugs or device malfunction (Giles et al., 2005; World Health Organization, 2005). However, as a

method for comparing institutions or evaluating interventions, they carry a significant risk of bias.

#### *4.1.2.3.3 Trigger tools*

Trigger tools are used to identify sentinel events, such as abnormal laboratory values, the prescription of antidotes or reports of patient harm, which indicate that a preventable adverse event may have occurred (Barber et al., 2006; Resar et al., 2006). Positive triggers are subjected to further investigation to determine whether an error had occurred. Since a detailed case-note review is only required when there is a positive trigger, the method is less resource intensive than a detailed manual review of all case notes in the sample.

There are three main problems associated with the use of trigger tools:

- their lack of sensitivity, (i.e. the number of adverse events detected by the tool). For example, the first study of the use of trigger tools in the UK (Barber et al., 2006) found that the tool identified less than 1% of prescribing errors (identified through the trigger tool, spontaneous reporting, pharmacist review and retrospective case-note review). The proportion of errors identified was higher for serious errors but never exceeded 50% (Barber et al., 2006);
- their low specificity, (i.e. the proportion of triggers that are preventable) and thus positive predictive value (PPV). If the trigger is not specific the PPV will be low and scarce resources will be devoted to investigating false negative.

Estimates of PPV vary considerably but do not exceed 38% (Brown et al., 2008c); and

- the risk of bias inherent in their use, as there is a risk that the tools will identify different events with different sensitivities, depending on local factors such as the type and completeness of information held.

However, trigger tools are likely to yield less biased results when used over time within an organisation so long as there is no material change in the type of data collected. As a basis for comparison across institutions or interventions they are liable to be biased.

#### *4.1.2.3.4 Error rates*

Errors affecting clinical process require measurement of both numerator and denominators if they are to be used for comparative purposes. Errors are often denominated at the level of the patient (i.e. errors per patient). However, this may yield biased results, particularly in non-randomised and unmasked studies. This is because case-mix across time or place may yield different opportunities for error. A method to reduce this type of bias is to denominate errors on the opportunity for error, rather than the number of patients (Lilford et al., 2003). However, this method does not completely negate the bias as:

- it may be easier to detect opportunities for error in one place than in another, for example due to differences in note keeping; and
- equally performing clinicians may find some errors more difficult to avoid than others and this confounds comparisons when the opportunities for error differ

from place to place. For example, if patients with more comorbidities "cluster" under the care of particular clinicians, it may be seen that these clinicians make more error, even after case-mix adjustment (Greenfield et al., 2002).

Despite these caveats, denominating error on the opportunity for error provides at least some protection against case-mix bias. The opportunity for error method also provides a method to deal with contingent errors, where the opportunity for error arises only if certain pre-existing conditions are fulfilled.

#### 4.1.2.3.4.1 Formal methods for detection of error rates

There are four main methods for the direct measurement of error:

- case-note review – retrospective construction of a cohort;
- prospective data collection by clinical staff;
- prospective data collection by independent observers;
- prospective data collection by a participant observer, either simulated (Merien et al., 2010) or real patients.

The important distinction between all these methods and reporting systems is that the data are recorded from all cases in a predefined cohort. Evidence suggests that the four different methods above will not result in identification of the same set of errors. For example, Michel et al., (2004) show that prospective data collection by clinical staff produced a higher error rate than retrospective evaluation of case-notes in the context of accident and emergency care.

The relative strengths and weaknesses of these (and other methods) of data collection have been reviewed by Thomas and Petersen (2003), Murff et al., (2003) and the UK's National Patient Safety Agency (2005). Thomas and Petersen (2003) suggest that the relative utility of different methods depends on the type of error or incident being investigated.

#### 4.1.2.3.4.2 Holistic versus explicit review

Data on error can be collected using holistic (implicit) and/or explicit methods. The holistic (see Chapter 8) method involves experts making their own judgements about the quality of care provided and can be either structured or unstructured. In a structured holistic review, experts are presented with a series of preparatory questions designed to elicit a complete review of the important facets of care, whilst in an unstructured holistic review experts are given little guidance and typically follows the format used by expert witnesses in litigation cases.

Explicit review (see 6.2.5) involves the objective application of predetermined standards, which are developed using expert groups and/or national care protocols. Explicit reviews can be focussed (using a limited set of supported and feasible measures) or global (using a broader set of quality measures for a large number of conditions). The quality of care in USA has been studied using a global approach, applying 439 indicators for 30 conditions (McGlynn et al., 2003). These global indicators have been adapted for use in the UK (Kirk et al., 2003), using a limited set of 200 indicators across 23 conditions. However, criteria or indicators should only be applied in an explicit review if they are considered relevant to the patient in question (Kahn et al., 1990).

Both holistic and explicit methods of data collection have advantages and disadvantages, which are described in Table 4.1. Data extraction can be enhanced by combing review methods (Hutchinson et al., 2010), and so both holistic and explicit methods have been employed in this evaluation (see Chapter 6 and Chapter 7).

**Table 4.1: Comparison of holistic and explicit methods of data extraction in case-note review**

| Holistic | Explicit |
|---|---|
| ***Advantages*** | |
| Easy to develop and administer | Explicit (evidence based) criteria |
| High face validity, since experts define "good" and bad care | Reproducible |
| Self-updating through use of experts | Easy to explain low score in terms of criteria – which may narrow score of improvement efforts |
| Reflects the full scope of clinical decisions that apply to a particular patient | Can be conducted by researchers rather than clinicians, once the criteria have been agreed, reducing costs |
| Involves physicians and other expert clinicians in the quality of care process | |
| ***Disadvantages*** | |
| Requires (expensive) clinical experts | Require training of reviewers |
| More arbitrary than evidence based | Need to be updated constantly |
| Developed principally for inpatient care | Limited scope in terms of content and context (relevant publications) |
| Poor reproducibility of judgements | Possible bias if different numbers of criteria apply to patients between comparative organisations, particularly if some criteria are harder to meet than others |

#### 4.1.2.3.4.3 Validity

Any method of data collection needs to demonstrate construct validity: the results need to be an accurate reflection of the underlying concept intended for the data. A surrogate end point has construct validity if it has shown empirically to correlate with outcome or if it this is judged to be self-evident (Lilford et al., 2004).

Moreover, many actions (or failures to act) may be more ambiguous and hence their construct validity is less clear cut. For example, it is commonly believed in patient safety circles that minor prescription errors are a good surrogate for major prescription errors. This assumption is based on Heinrich's seminal work of 1931, where he outlined a ratio in which there were 29 minor injuries and 300 no-injury accidents for every major incident (Heinrich, 1931). However the empirical basis for this conclusion is poorly described and a review of studies on prescribing error found insufficient evidence to support a single ratio that could be used to validate the use of minor errors as a surrogate for major errors (Barber et al., 2006; Dean et al., 2002).

#### 4.1.2.3.4.4 Reliability

Any measure of patient safety needs to be reliable (repeatable). Inter-observer reliability implies that more than one observer/reviewer would come to the same conclusion when evaluating a single care process. The traditional methods of assessing the extent of inter-rater reliability between two or more independent reviewers is the Cohen kappa $\kappa$ (Cohen, 1960), although the sensitivity of $\kappa$ to the prevalence of error (Lilford et al., 2007) has

resulted in other methods, such as tetrachoric or polychoric correlations being advocated where possible (Hutchinson, 1993). It should be noted that the κ statistic is used for dichotomous judgements and the intra-class correlation coefficient (ICC) is used for rating scales, as used for the quality of care ratings in section 7.3.3. Since duplicating observations is expensive it is important to optimise the sample size for the measurement of reliability using formulae to calculate the precision of measurements if inter-observer agreement (Altman, 1991).

A recent review by Lilford et al., (2007) considered factors that may affect the κ obtained and reported that the following factors raised the mean κ: a) the use of explicit, rather than holistic; b) assessing outcomes rather than causality and processes. It was also noted that κ was positively correlated with the prevalence of error.

### 4.1.3 Masking

As with all other evaluations, masking patients, caregiver and observers as well as those undertaking the statistical analysis is important in minimising information bias (Schulz and Grimes, 2002). Masking is particularly important when the end-point being measured is subjective rather than objective (Morrison and Lilford, 2000). Where assessments are being made about the quality of care reviewers tend to give worse rating if an adverse outcome occurred – hindsight bias (Caplan et al., 1991). This can bias the estimated size of the problem and lead to exaggerated estimates of cost effectiveness. Evaluating all cases of care and looking for all errors, rather than first selecting adverse events and then looking into to

see whether an error occurred, can reduce such a bias – an approach taken in this evaluation.

Hindsight bias does not prejudice assessment of relative safety improvements in a comparative study if it is applied equally across comparison groups. Here, bias is a risk if the observer is aware of the group (intervention or control) to which an individual or cluster has been assigned. Observers should therefore be blinded to the allocation of the data, be it intervention, control, or before or after. This principle as has been applied in this evaluation, see 6.2.4).

Biases associated with the measurement of patient outcomes are discussed separately in section 4.1.1.2.

## 4.2 Framework for the evaluation

In section 4.1 I described what can be measured to determine if a patient safety complex intervention has been effective. However, it is important to understand the fidelity to which an intervention was implemented and to understand why it worked the way it did (see 2.1.4). In these circumstances a mixed-method evaluation, using both quantitative and qualitative studies is useful.

The evaluation of the **SPI** used mixed-methods and was based on a systems-wide approach that is described in section 3.4. In this approach the system is conceptualised as the setting in which care is delivered and five levels can be distinguished:

- structure (e.g. size of hospital and types of services provided). This has been described in Table 3.1 and Table 4.3;
- management processes (e.g. leadership style, management WalkRounds);
- intervening (mediating) variables (e.g. culture, morale, absence due to sickness) that connect management to clinical process;
- clinical processes (error rates/compliance with tenets of evidence based care);
- outcome (adverse events, mortality, patient satisfaction).

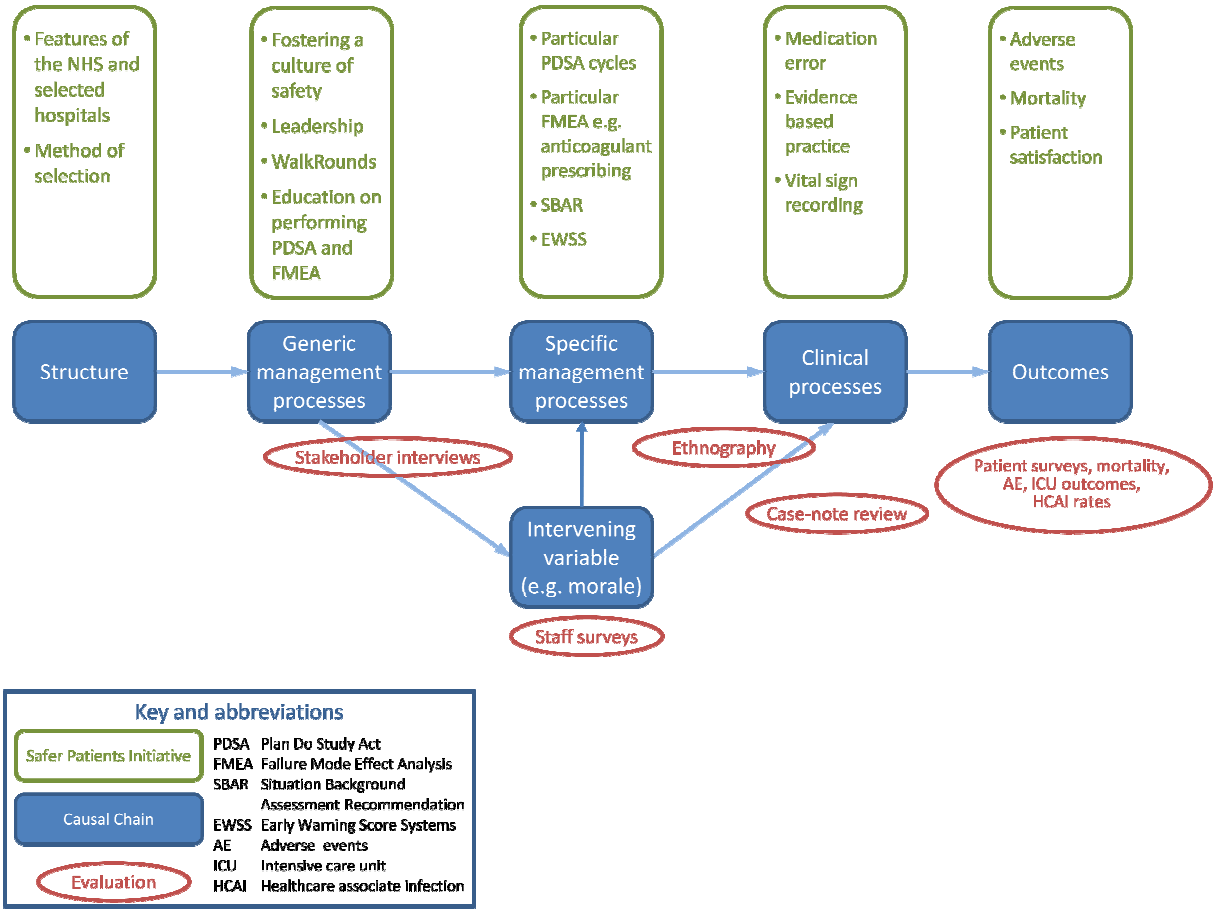
The data collection and analyses was organised around this conceptual model, using a series of linked sub-studies (Figure 4.2):

- management processes were studied by means of a series of qualitative interviews with strategic stakeholders (Appendix B);
- intervening variables were assessed by using quantitative staff survey (Chapter 5) and a qualitative study using ethnographic methods on hospital wards (Appendix C);
- clinical processes were studied by: case-note review of acutely sick patients using explicit methods (Chapter 6) and holistic methods (Chapter 7); an explicit case

note review of peri-operative care (Chapter 8); and by using indirect measures of hand hygiene (Chapter 9);

- outcomes were studied by: case-note review of adverse events (Chapter 10) and mortality (Chapter 11);, outcomes within the ICU (Chapter 12), healthcare associated infection rates (HCAI) (Chapter 13) and quantitative patient surveys (Chapter 14).

**Figure 4.2: Causal chain linking SPI to outcomes. The evaluation sub-studies were made at points across the chain to provide information on context, fidelity, and effectiveness of SPI**



The evaluation of ***SPI1*** preceded the evaluation of ***SPI2*** and not all methods were employed in all hospitals:

The qualitative studies (Appendix B and C) were conducted in the four ***SPI1*** intervention hospitals only, as the funder of the evaluation believed that "data saturation" would be reached, i.e. no further data collection would elicit new information.

The only studies conducted in both the ***SPI1*** and ***SPI2*** evaluations were the staff surveys (Chapter 5), case-note review of acutely sick patients (Chapters 6, 7, 10 and 11) and patient surveys (Chapter 14). This was because at the time of commissioning the ***SPI1*** evaluation, a decision needed to be made on what could be measured within the funding envelope. The staff survey and patient surveys utilised data already collected within the NHS, whilst the decision to collect case-notes of the acutely sick patient was based on the opportunity to measure several facets of the intervention, mainly the use of EWSS (see 3.6.2.1) and medicines management (see 3.6.2.3) within one episode of care. In addition one of the ***SPI1*** hospitals did not have an ICU, which was a target of the **SPI**. These measures were extended to the ***SPI2*** evaluation so that the data collected in the ***SPI2*** evaluation could act as control data for the evaluation of ***SPI1***, the details of which can be found in section 4.3.

For the evaluation of ***SPI2*** the funders requested that the quantitative data collection be expanded to evaluate additional aspects of the intervention not measured within the ***SPI1*** evaluation, such that, case-note review of peri-operative care (Chapter 8), indirect measures

of hand hygiene (Chapter 9), outcomes with the ICU (Chapter 12) and HCAI rates (Chapter 13) were conducted in the $\boldsymbol{SPI_2}$ evaluation only.

All the quantitative sub-studies were before and after studies in intervention and concurrent control hospitals. Use of both before and after observations across control and hospitals enabled rates of change to be compared across control and **SPI** hospitals – an approach referred to as the "difference-in-difference" method (see 2.1.3.2). The sub-studies are summarised in Table 4.2.

**Table 4.2: Summary of sub-studies comprising the evaluation of the SPI**

| Sub-study | Purpose | Location | Data collection | Analysis |
|---|---|---|---|---|
| **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 |
| **Staff survey** | Measure effects of **SPI** on staff morale, culture and opinion | Control, ***SPI₁*** and ***SPI₂*** hospitals | Validated structured questionnaire | 1) Comparison of control versus **SPI** hospitals:<br>a. At baseline<br>b. Over time, i.e. difference-in-difference<br>2) Comparisons within control and **SPI** cohorts |
| **Ethnography study at ward level** | Discover the impact of ***SPI₁*** on 'hearts and minds' of those actually 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 |
| **Quality of care: acute medical care (clinical process)** | Measure effects of **SPI** on the quality of care being delivered using independent case-note reviews in acute medical care (both explicit and holistic). | Control, ***SPI₁*** and ***SPI₂*** hospitals | Before and after intervention phase of ***SPI₁*** and ***SPI₂*** | 1) Comparison of control versus **SPI** hospitals:<br>a. At baseline<br>b. Over time, i.e. difference-in-difference (Epochs 1 vs. 2 for the ***SPI₁*** evaluation) (Epochs 1+2 vs. Epoch 3 for the ***SPI₂*** evaluation )*<br>c. Measurement of reliability and learning/fatigue effects |
| **Quality of care: peri-operative care (clinical process)** | Measure effects of **SPI** on the quality of care using independent case-note reviews in peri-operative care (explicit). | Control and ***SPI₂*** hospitals | Before and after intervention phase of ***SPI₂*** | 1) Comparison of control versus ***SPI₂*** hospitals:<br>a. At baseline<br>b. Over time, i.e. difference-in-difference (Epoch 2 vs. Epoch 3)* |

| | | | | |
|---|---|---|---|---|
| **Indirect measures (clinical process)** | Measure used hand hygiene consumables. | Control and $SPI_2$ hospitals | Trend data collected as part of the National Observation Study of Effectiveness of the national "*cleanyourhands*" Campaign study | 1) Comparison of control versus SPI hospitals:<br>a. At baseline<br>b. Over time, i.e. difference-in-difference |
| **Outcomes** | Measure effects of **SPI** on:<br>a. Adverse events among acute medical care case-notes reviewed<br>b. Mortality among acute medical care case-notes reviewed<br>c. ICU outcomes [$\boldsymbol{SPI_2}$ only]<br>d. HCAI rates [$\boldsymbol{SPI_2}$ only]<br>e. Patient satisfaction | Control, $\boldsymbol{SPI_1}$ and $\boldsymbol{SPI_2}$ hospitals | Before and after study using:<br>a) and b) Case-notes<br>c) d) and e) Routine data<br>f) Validated structured survey | 1) Comparison of **SPI** versus control hospitals:<br>a. At baseline<br>b. Over time, i.e. difference-in-difference<br>2) Comparisons within **SPI** and control cohorts<br>3) Measurement of reliability and learning/fatigue effects |

*Sub-studies involving case-note review that overlapped with $\boldsymbol{SPI_1}$ have two pre-intervention phases (Epochs 1+2), while sub-studies specific to $\boldsymbol{SPI_2}$ have only one pre-intervention phase (Epoch 2). In all cases Epoch 3 is the post-intervention phase.

## 4.3 Controls and intervention hospitals

The **SPI** hospitals were selected by the Health Foundation (see 3.2 and 3.7.1). For the ***SPI1*** evaluation all four intervention hospitals were evaluated. In the ***SPI2*** phase of the intervention 20 hospitals across the UK were selected to participate, however, the evaluation focused on the ten English ***SPI2*** hospitals only so that routinely collected data in England (staff survey, hospital mortality, ICU outcomes, HCAI rates and patient survey) could be accessed.

Although the ***SPI2*** hospitals worked in pairs each hospital formed a unit of analysis for the statistical power calculation and for the evaluation. Participation in the evaluation was not compulsory and one of the ten ***SPI2*** hospitals declined to partake as they felt they did not have the capacity to assist in the evaluation, leaving nine available for the study.

The selection of the control hospitals for the ***SPI1*** evaluation capitalised on the evaluation of ***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 the ***SPI2*** evaluation 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 hospitals and nine were ***SPI2*** matched control hospitals (for the ***SPI1*** evaluation these are described as the "18 controls").

***SPI2*** controls were selected using the following criteria (it should be noted that some of the matching was done at trust level and some at hospital level):

- only non-specialist acute trusts in England were considered;
- control and ***SPI2*** trusts should have a similar directorate structure (as described in the NHS national staff survey);
- the trusts should have the same foundation or non-foundation status (to gain foundation status a trust must satisfy the government that it has the management capacity to warrant greater operational autonomy);
- the trust should be similarly located in either urban or rural settings;
- once these criteria were satisfied, the hospital within a trust with the most similar size (usually within 1000 staff) to the ***SPI2*** hospital was selected as the control hospital;
- if a trust had more than one hospital, quantitative data collection was focused on the largest hospital with an ICU, as the ICU was targeted for improvement in the **SPI**.

Characteristics of ***SPI2*** hospitals and the matched controls can be found in Table 4.3.

**Table 4.3: Hospital characteristics of *SPI2* hospitals and matched control hospitals**

| | *SPI2* hospitals | | | Matched control hospitals | | |
|---|---|---|---|---|---|---|
| **Pair n°** | **Bed numbers (current)** | **Urban/ rural†** | **Teaching status ‡** | **Bed numbers (current)** | **Urban/ rural** | **Teaching status** |
| **1** | 411 | Rural | Affiliated | 475 | Rural | Nil |
| **2** | 455 | Urban | Nil | 511 | Urban | Nil |
| **3** | 620 | Urban/ rural | Nil | 618 | Urban | Teaching hospital |
| **4** | 634 | Urban | Nil | 723 | Urban/ rural | Nil |
| **5** | 688 | Urban | Teaching hospital | 447 | Urban/ rural | Affiliated |
| **6** | 804 | Urban | Teaching hospital | 789 | Urban | Affiliated |
| **7** | 668 | Urban | Teaching hospital | 988 | Urban | Affiliated |
| **8** | 523 | Urban | Teaching hospital | 532 | Urban/ rural | Nil |
| ***9** | 566 | Urban | Affiliated | 1,036 | Urban | Affiliated |

** The discrepancy in size has arisen as matching was based number of staff at trust level, which was the best available data at the time of matching. The bed numbers in this table are those at hospital level and have only recently become available.*

*† Based on visual inspection of population density map*

*‡ Based on hospital website*

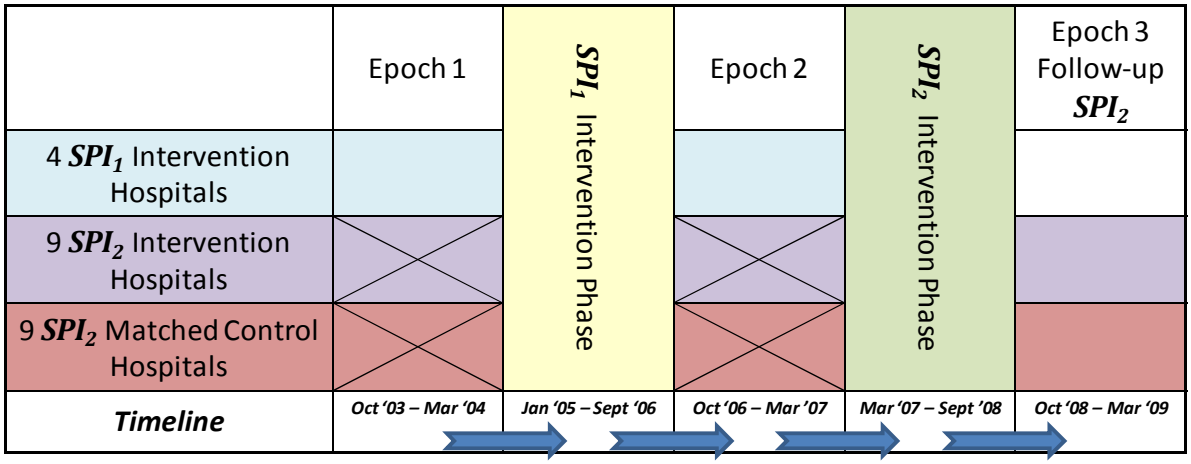
*The hospitals have not been named as it was agreed that anonymity would be ensured as part of evaluation participation.*

The method by which ***SPI2*** hospitals could serve as controls for ***SPI1*** is explained by Figure 4.3.

Although four ***SPI1***, nine ***SPI2*** and nine control hospitals agreed to participate in the evaluation further consent for each sub-study was required. In some instances this was not granted. In addition some of the study hospitals did not participate in the national routine data collection exercises that were external to the evaluation but were accessed for analyses. Whilst yet others failed to supply case-notes for specific analysis. It is for these

reasons that discrepancies exist in the number of hospitals agreeing to participate in the evaluation and the number included in each sub-study.

**Figure 4.3: Schematic diagram demonstrating staggering of the data collection enabling the use of pre-$SPI_2$ intervention data as control data for the evaluation of $SPI_1$.**




Note:
a) Those time periods in which a cross appears pertain to the control data for the $SPI_1$ evaluation
b) No $SPI_1$ data was collected during the $SPI_2$ follow-up phase;
c) The "epochs" relate to the times when patients whose notes were reviewed were treated. The reviews of the notes themselves followed the epochs but they were overlapped so that any learning/fatigue effects on part of the reviewers could be controlled for.

# CHAPTER 5: STAFF SURVEYS

## 5.1 Introduction

Strategic support for any programme may not reflect views at the "sharp end" of practice. Thus, intervening variables such as staff morale, attitudes and various factors relevant to 'culture' that might be affected by the **SPI** were assessed by staff survey. The aim of this sub-study was to measure the effects of **SPI** on intervening variables.

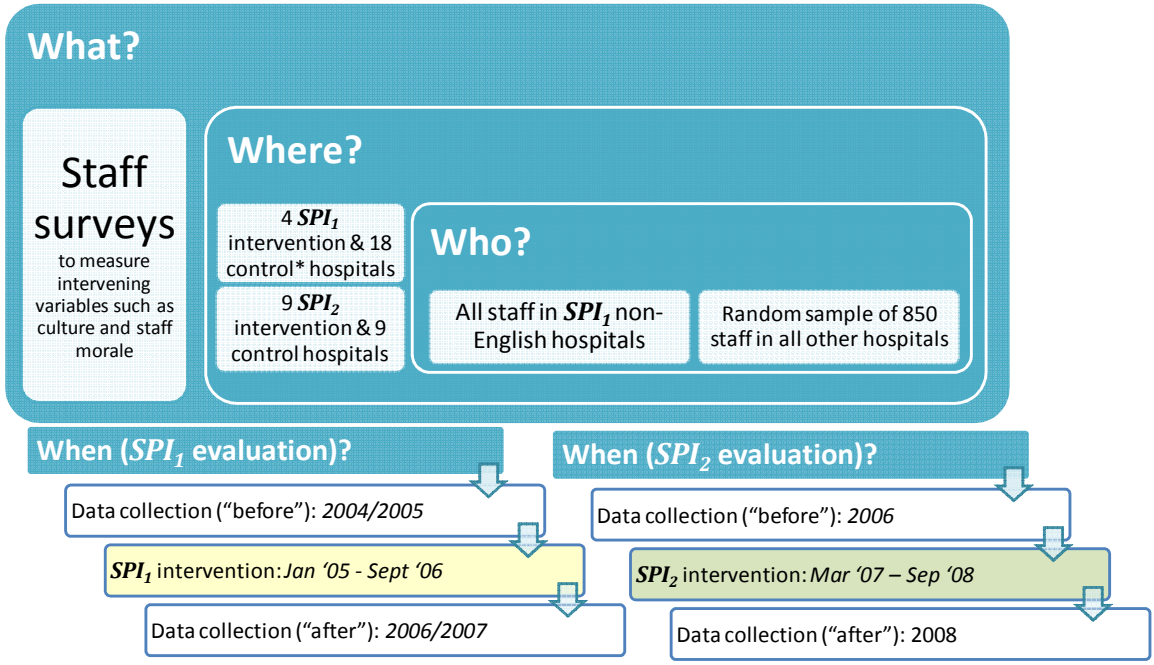
## 5.2 Methods

This sub-study was conducted in both the ***SPI$_1$*** and ***SPI$_2$*** evaluations. The methods for this sub-study are summarised in Figure 5.1.

All hospitals in England participate in the National Staff Survey, a yearly survey run by the Care Quality Commission (CQC) (formerly the Healthcare Commission). For the evaluation of ***SPI$_1$*** arrangements were made to conduct the same survey, using the same survey methods and questionnaire in the three non-English hospitals. 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 (note that this was 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 ***SPI$_1$***

hospitals because it was undergoing a merger and a period of major reorganisation. Data were therefore available for three of the four ***SPI*$_1$** hospitals in the second round.

**Figure 5.1: An outline of the staff survey sub-study**




*The controls for the ***SPI*$_1$** evaluation comprise of the pre-intervention data collected for the ***SPI*$_2$** evaluation. For a full explanation see section 3.2.

Questionnaires were sent to all staff in the four ***SPI*$_1$** 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 hospital) would yield 95% confidence intervals of no greater than 10% for all scores within a single organisation.

In the ***SPI*$_2$** evaluation all nine control hospitals and nine intervention hospitals were included in both the 2006 and 2008 National Staff Surveys, conducted between October and

December in each of these years, and so data from these surveys were used to test for effects of the intervention. Again questionnaires were sent to a simple random sample of 850 staff in each hospital. The detail of the survey methods is not repeated here but is available from the staff survey advice centre website (www.nhsstaffsurveys.com).

### 5.2.1 Statistical methods

Approximately 28 survey scores are regularly reported by the CQC. Thirteen of these (Box 5.1) were identified at the start of the evaluation as being of likely relevance to the **SPI** programme, 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 D.

**Box 5.1: Staff survey variables deemed relevant to the SPI**

| |
|---|
| 1. Well structured appraisals (Guzzo et al., 1985; West et al., 2006) |
| 2. Working in well-structured teams (Borrill et al., 2000) |
| 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 (Michie and West, 2004) |
| 12. Fairness and effectiveness of incident reporting procedures* |
| 13. Availability of hand washing materials* |

** These scores were not included in the 2004 and 2006 staff survey and consequently were unavailable for the* $SPI_1$ *evaluation. However, these were included in proceeding surveys and hence included in the* $SPI_2$ *evaluation.*

Differences between the control and **SPI** hospitals in terms of changes between the two survey periods were tested using a generalised linear mixed model with **SPI**/control and survey period as fixed factors (with interaction), and hospital as a random factor. In order 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 formal statistical correlation for multiple observations was not applied but statistical significance is claimed for p-values less than 0.01, and 99% confidence intervals are used throughout.

## 5.3 Results

### 5.3.1 Results of the $SPI_1$ evaluation

For the first staff survey, the mean response rate was 45% (7826 of 17,507 returned) in the four $SPI_1$ hospitals; for the second it was just 35% (4191/11,922) across the three hospitals that participated (see Table 5.1). In the eighteen 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%.

There was no significant difference in response rates between control and $SPI_1$ hospitals at baseline. Table 5.1 shows the values of the 11 survey scores in each of the four $SPI_1$ hospitals for the two surveys, along with details of response rates. Table 5.2 shows the *changes* in both control and $SPI_1$ 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 (Healthcare Commission, 2007a). At baseline, the percentage of staff reporting "*well-structured appraisals within the previous 12 months*" was significantly lower in $SPI_1$ hospitals than control hospitals. "*Job satisfaction"* and "*support from supervisors"* were also significantly lower in $SPI_1$ hospitals than control hospitals ($p<0.01$). None of the other baseline differences was statistically significant.

Only one of the eleven scores shows a statistically significant difference ($p<0.01$) in changes between the control hospitals and $SPI_1$ hospitals between the two surveys. "*Organisational climate*", which refers to extent of positive feeling within the organisation relating to communication, staff involvement, innovation and patient care (Michie and West, 2004), was similar between the control and $SPI_1$ hospitals at baseline (3.11 versus 3.08 on a scale where 1 is very negative and 5 is very positive). This score decreased by 0.22 in the control hospitals but only by 0.15 in the $SPI_1$ hospitals ($p<0.001$). The effect size for this difference in change between the control and $SPI_1$ hospitals after covariates are taken into account was modest, at 0.08 points on a 5 point scale where there was a range at baseline of 0.5 points between hospitals.

### 5.3.2 Results of the $SPI_2$ evaluation

In the nine $SPI_2$ hospitals, the overall response rate for the first, "before", survey was 53% (3957 of 7402 valid questionnaires returned). This remained the same (53%) for the second, "after", survey (3940/7448). In the nine control hospitals, the response rates were 50% (3634/7301) and 49% (3616/7424) respectively. Table 5.3 shows the changes in both control and $SPI_2$ hospitals on each of the 13 scores identified, along with the differences between the groups in these changes (with associated 99% confidence intervals).

Comparison with control hospitals showed that national changes in the NHS over this period resulted in generally more positive scores at the second survey than at the first (Healthcare Commission, 2007a).

Again, only "*organisational climate*" (out of 13 scores) shows a statistically significant ($p<0.01$) change over time between the control hospitals and $\boldsymbol{SPI_2}$ hospitals. On this occasion it was significantly lower in the control hospitals than the $\boldsymbol{SPI_2}$ hospitals at baseline (2.79 versus 2.91 on a scale where 1 is very negative and 5 is very positive). Thus, although the increase in this score in control hospitals was higher than in $\boldsymbol{SPI_2}$ hospitals (0.08 compared with 0.01), the score was still higher in the $\boldsymbol{SPI_2}$ hospitals at the second survey. The effect size for this difference in change between the control and $\boldsymbol{SPI_2}$ hospitals after covariates are taken into account was modest, at 0.07 points on a 5 point scale where there was a range at baseline of 0.55 points between hospitals.

**Table 5.1: Staff survey scores in $SPI_1$ hospitals at the two periods***

| | Hospital 1 | | | | Hospital 2 | | | | Hospital 3 | | | | Hospital 4** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | Survey 1 [2004/2005] score (SE) | N | Survey 2 [2006/2005] score (SE) | N | Survey 1 [2004/2005] score (SE) | N | Survey 2 [2006/2007] score (SE) | N | Survey 1 [2004/2005] score (SE) | N | Survey 2 [2006/2007] score (SE) | N | Survey 1 [2004/2005] score (SE) |
| % staff having well structured appraisals within previous 12 months (Guzzo et al., 1985; West et al., 2006) | 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 (Borrill et al., 2000) | 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 (Michie and West, 2004) | 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 (Michie and West, 2004) | 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 (Michie and West, 2004) | 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 (Michie and West, 2004) | 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 (Healthcare Commission, 2006; Michie and West, 2004)) | 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 1-5, and are based on the mean of between three and six questions, each of which was scored between 1 and 5 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)

Shaded areas relate to post-intervention epochs.

**Table 5.2: Staff survey scores in control and *SPI*$_1$ hospitals at the two periods***

| | Control hospitals | | | | | *SPI*$_1$ hospitals | | | | | Range at base-line | Difference in change (99% CI) | P-value |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | Survey 1 [2004] score (SE) | N | Survey 2 [2006] score (SE) | Absolute % change | N | Survey 1 [2004/ 2005] score (SE) | N | Survey 2 [2006/ 2007] 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 (Healthcare Commission, 2006)* | *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 (Healthcare Commission, 2006) | 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 (Healthcare Commission, 2006) | 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 (Healthcare Commission, 2006) | 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 (Healthcare Commission, 2006) | 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 1-5, and are based on the mean of between three and six questions, each of which was scored between 1 and 5 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 $\boldsymbol{SPI_1}$ 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.

Standard Error (SE)

Shaded areas relate to post-intervention epochs

**Table 5.3: Staff survey scores in control and *SPI2* hospitals at the two periods***

| | Control hospitals | | | | | *SPI2* hospitals | | | | | Range at baseline | Difference in change (99% CI) | p-value |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | Survey 1 [2006] score (SE) | N | Survey 2 [2008] score (SE) | Absolute % change | N | Survey 1 [2006] score (SE) | N | Survey 2 [2008] score (SE) | Absolute % change | | | |
| % staff having well structured appraisals within previous 12 months (Guzzo et al., 1985; West et al., 2006) | 3477 | 28 (1) | 3429 | 28 (1) | -1 | 3783 | 28 (1) | 3734 | 26 (1) | -2 | 20-39 | 3 (-3, 9) | |
| % staff working in well structured teams (Borrill et al., 2000) | 3498 | 36 (1) | 3408 | 37 (1) | 1 | 3781 | 38 (1) | 3747 | 38 (1) | 0 | 32-42 | 4 (-4, 12) | 0.205 |
| *% staff witnessing potentially harmful errors or near misses in previous month* | *3602* | *37 (1)* | *3532* | *33 (1)* | *-4* | *3918* | *41 (1)* | *3851* | *40 (1)* | *-1* | *32-47* | *-4 (-10, 3)* | *0.167* |
| *% staff suffering work related injury in previous 12 months* | *3524* | *19 (1)* | *3490* | *16 (1)* | *-3* | *3848* | *19 (1)* | *3796* | *18 (1)* | *-1* | *16-23* | *-2 (-5, 2)* | *0.182* |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *% staff suffering work related stress in previous 12 months* | *3575* | *33 (1)* | *3532* | *27 (1)* | *-6* | *3882* | *32 (1)* | *3842* | *27 (1)* | *-6* | *26-40* | *-1 (-6, 5)* | *0.670* |
| *% staff experiencing physical violence from patients/relatives in previous 12 months* | *3598* | *1 (1)* | *3536* | *11 (1)* | *-1* | *3884* | *11 (1)* | *3849* | *11 (1)* | *0* | *7-16* | *-1 (-3, 3)* | *0.645* |
| *Intention to leave* (Michie and West, 2004)) | *3557* | *3.26 (0.02)* | *3544* | *3.40 (0.02)* | *0.14* | *3880* | *3.31 (0.01)* | *3865* | *3.42 (0.01)* | *0.11* | *3.07-3.50* | *-0.04 (-0.12, 0.04)* | *0.198* |
| Staff job satisfaction (Michie and West, 2004) | 3593 | 3.34 (0.01) | 3568 | 3.44 (0.01) | 0.10 | 3902 | 3.40 (0.01) | 3898 | 3.49 (0.01) | 0.09 | 3.23-3.50 | -0.02 (-0.08, 0.04) | 0.422 |
| Quality of work life balance (Michie and West, 2004) | 3568 | 2.77 (0.02) | 3536 | 2.56 (0.02) | -0.22 | 3868 | 2.68 (0.02) | 3857 | 2.51 (0.02) | -0.17 | 2.46-2.97 | 0.05 (-0.04, 0.14) | 0.142 |
| Support from supervisors (Michie and West, 2004) | 3583 | 3.39 (0.02) | 3551 | 3.56 (0.02) | 0.17 | 3894 | 3.43 (0.01) | 3869 | 3.61 (0.01) | 0.18 | 3.22-3.53 | 0.00 (-0.08, 0.07) | 0.889 |
| Organisational climate (Healthcare Commission, 2006; Michie and West, 2004) | 3578 | 2.79 (0.01) | 3551 | 2.87 (0.01) | 0.08 | 3861 | 2.91 (0.01) | 3886 | 2.92 (0.01) | 0.01 | 2.52-3.07 | -0.07 (-0.14, 0.00) | 0.009 |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| † Fairness and effectiveness of incident reporting procedures (Michie and West, 2004) | 3555 | 3.36 (0.01) | 3487 | 3.41 (0.01) | 0.05 | 3861 | 3.41 (0.01) | 3803 | 3.45 (0.01) | 0.04 | 3.27-3.54 | -0.01 (-0.05, 0.04) | 0.664 |
| † Availability of hand washing materials (Michie and West, 2004) | 2939 | 4.58 (0.01) | 3126 | 4.75 (0.01) | 0.17 | 3231 | 4.51 (0.01) | 3418 | 4.67 (0.01) | 0.16 | 4.32-4.72 | -0.01 (-0.07, 0.04) | 0.587 |

* 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 seven are on a scale of 1-5, and are based on the mean of between three and six questions, each of which was scored between 1 and 5 for each respondent. For six of these seven scores, 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 **SPI** and control 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.

† These scores were not included in the $\boldsymbol{SPI_1}$ evaluation as they were not included in survey 1.

## 5.4 Discussion

In both the ***SPI₁*** and control hospitals the response rate declined from the first survey to the second survey, which was reflected in a national decrease in response rates among acute hospitals in England over the same period (from 57% to 52%) (Healthcare Commission, 2005; Healthcare Commission, 2007b). However, response rates were somewhat lower in the non-English hospitals, which may be due to the lack of the national profile for the survey outside England. The ***SPI₂*** response rates remained stable between the two surveys for both control and intervention hospitals.

Overall, the staff survey shows little change between the first and second survey in both control and **SPI** hospitals (both ***SPI₁*** and ***SPI₂***). Control and **SPI** hospitals were also mostly indistinguishable at baseline.

The **SPI** had little impact on the culture of the organisation, as within both evaluations only one (*"organisational climate"*) of the 11 dimensions of staff satisfaction changed significantly over time but to a small degree. However, this change favoured the intervention hospitals in the ***SPI₁*** evaluation; but within the ***SPI₂*** evaluation controls were favoured.

# CHAPTER 6: EXPLICIT CASE-NOTE REVIEW TO MEASURE ERROR RATES/QUALITY OF CARE IN PATIENTS WITH ACUTE RESPIRATORY DISEASE

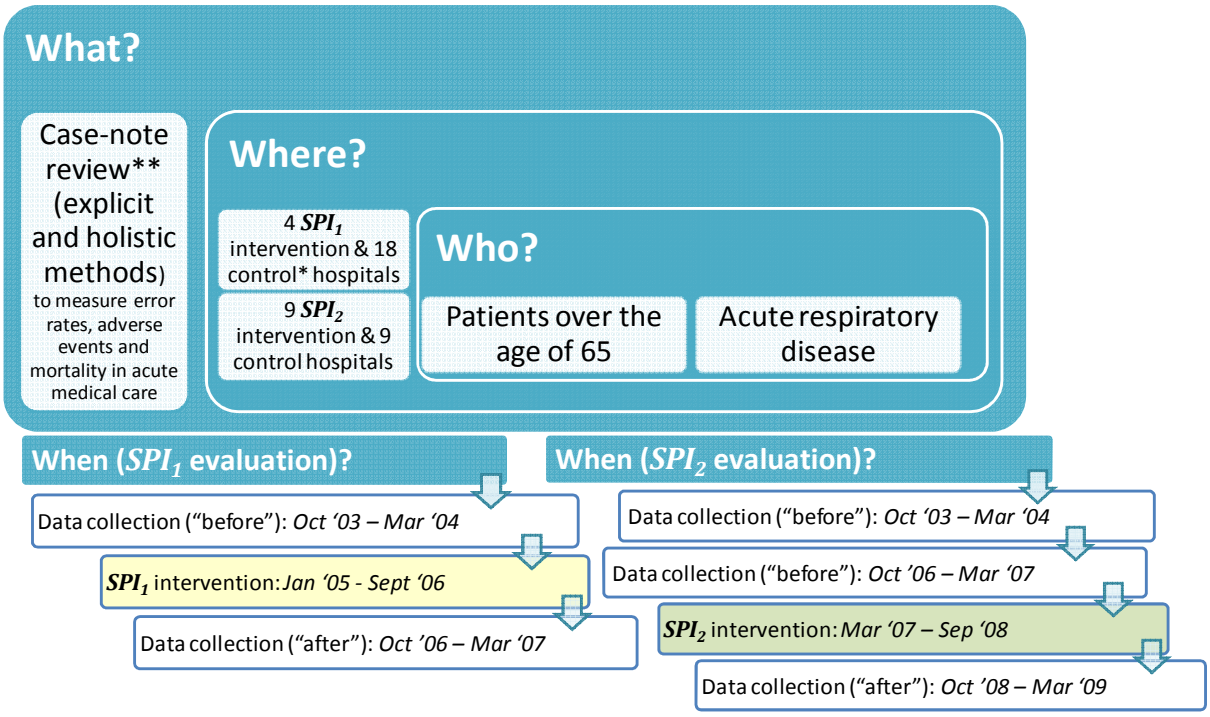
## 6.1 Introduction

Specific **SPI** targets included improvements to the monitoring of the acutely sick patient and a reduction in the number of medication errors (see 3.6.2.1 and 3.6.2.3). An anamnestic method of measuring the effect of the **SPI** in these areas is by retrospective case-note review. In this chapter I will describe an explicit method of review (adherence to predetermined standards of care) of patients admitted with acute respiratory disease. The areas of review include both those specifically targeted by the **SPI**, and those that might plausibly be expected to improve if an overall shift in organisational systems and culture related to patient safety had occurred.

## 6.2 Methods

An overview of this sub study can be found in Figure 6.1. This sub-study was conducted in both the $SPI_1$ and $SPI_2$ evaluations. The case-notes collected as part of this sub-study are also utilised in the holistic review, measurement of adverse events and measurement of mortality within this cohort of patients. These sub-studies are reported as separately in Chapters 8, 11 and 12.

**Figure 6.1: An outline of the case-note review of patients admitted with acute respiratory disease**




*The controls for the ***SPI*$_1$** evaluation comprise of the pre-intervention data collected for the ***SPI*$_2$** evaluation. For a full explanation see section 3.2.
** The sub-studies for the explicit review, holistic review, adverse events and mortality are reported separately in Chapters 7, 8, 11 and 12 respectively

### 6.2.1 Case-note selection criteria

Patients over the age of 65 years with acute respiratory disease admitted to acute medical wards were selected as the focus for study for the following reasons:

- improving recognition and response to acute deterioration in a patient's condition was a specific SPI target, and patients admitted with acute respiratory disease are at high risk of such deterioration (Hillman et al., 2001; Smith et al., 2006);
- a number of specific evidence-based guidelines exist for this condition;

- there is a high incidence of co-morbidities in people aged over 65, making this a high-risk population where the opportunity for error is high and hence where there should be opportunity for improvement.

### 6.2.2 Case-note assembly

#### *6.2.2.1 Case-notes collected for the $SPI_1$ evaluation*

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 ***$SPI_1$*** intervention period (Figure 6.1) were collected. The number of ***$SPI_1$*** hospitals was fixed and the control observations were spread across a greater number of hospitals to provide a more robust sample. The aim was to analyse 100 case-notes from each ***$SPI_1$*** hospital per epoch (800 in total) and 15 from each control hospital per epoch (540 in total).

Epoch 1 extended from October 2003 to September 2004 in the ***$SPI_1$*** 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.

The case-notes from the first eight or nine patients who fulfilled the eligibility criteria (see 6.2.4) were selected from each ***$SPI_1$*** hospital, in each month from each epoch. In the control hospitals, the first two or three of such cases were selected.

#### *6.2.2.2 Case-notes collected for the $SPI_2$ evaluation*

Case-notes from both the nine control and nine ***$SPI_2$*** hospitals from time periods that both preceded (Epochs 1 and 2) and followed (Epoch 3) the ***$SPI_2$*** intervention period were collected (Figure 6.1). The pre-implementation observations were spread over two epochs (Epoch 1 [October 2003 to March 2004] and Epoch 2 [October 2006 to March 2007]) so that the hospitals participating in the ***$SPI_2$*** evaluation could also serve as controls for the preceding ***$SPI_1$*** evaluation. Epoch 3 (October 2008 to March 2009) was therefore the post-***$SPI_2$*** period. (The temporal change between Epochs 1 and 2 was included as a fixed effect in the statistical models. Each six month time period was made to correspond across the calendar to control for seasonal effects). The aim was to analyse, using review against explicit criteria, 15 case-notes from each control and ***$SPI_2$*** hospital per epoch (810 in total).

Epoch 1 extended from October 2003 to March 2004 in the ***$SPI_2$*** and control hospitals, and from October 2006 to March 2007 in Epoch 2, and from October 2008 to March 2009 in Epoch 3.

### 6.2.3 Sample size

The target sample size for both the ***$SPI_1$*** and ***$SPI_2$*** evaluations (540 and 810 case-notes respectively) would give 80% power to detect effects summarised in Table 6.1. For example, for a standard (such as measurement of respiratory rate at least 6 hourly) 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 binary data analysis 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 6.1 since the typical patient is associated with more than one medication prescription and thus has several opportunities for error. However, 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.

**Table 6.1: 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**

| 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 |

### 6.2.4 Eligibility criteria

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 6.2.1).

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.

The quality of anonymisation was audited by asking the reviewer to note if the hospital of origin, the year of origin and the patient name had been recognised.

### 6.2.5 Explicit case-note review

A set of explicit criteria to define medical care for respiratory patients was developed with reference to British Thoracic Society (BTS) guidelines (2001) and (2004), the British National Formulary [BNF (versions 53 (2007a), 54 (2007b) and 56 (2008)) – the editions that covered the study period] and expert opinion (consultant respiratory physicians from a teaching and a general hospital – see contributions). The areas of review and source of guidelines were:

- quality of medical history-taking. Eleven items (Box 6.1) were identified as constituting the ideal history for a patient admitted with acute respiratory disease (expert opinion);
- proportion of routine investigations (urea and electrolytes, chest x-ray and full blood count) ordered within 6 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) was evaluated on admission and then for the

first and subsequent 6 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 6.3) (BTS);
- investigating features of good care for specific classes of patients by:
    - calculating the CURB score to determine the severity of community acquired pneumonia and hence appropriate antibiotic selection (Box 6.1) (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" (Dean et al., 2000).*

Errors were identified using a previously developed pro forma (Barber et al., 2006) and categorised according to stage of the drug use process (Appendix B) (Cousins and Hatoum,

1991). **SPI** had identified reductions in the number of adverse effects related to anticoagulant therapy as a key aim 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 3.2) and, therefore, failures to continue to prescribe medicines on the transition from primary to secondary care where no explanation for this was recorded in the notes were examined.

**Box 6.1: 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

**Table 6.2: 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 | - | ✓ |

**Table 6.3: Appropriate clinical response for abnormal observations**

| Abnormal vital sign | Appropriate clinical response |
|---|---|
| Oxygen saturation <90%, at any time | Full blood gases within 2 hours<br>Given oxygen (if not on oxygen)<br>Doctor called or transferred to ICU (if on oxygen) |
| Blood pressure systolic <90 mmHg | At least next 6 hours, hourly observations<br>Blood culture |
| Sputum present | Sputum culture |
| Respiratory rate >20 breaths per minute at any time after admission | Given oxygen (if not on oxygen)<br>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<br>Repeat chest x-ray<br>White cell counted/repeated<br>Appropriate addition of further antibiotics |

**Box 6.2: Assessment of severity of community acquired pneumonia using the CURB score**

**CURB score**
**C**onfusion: new mental confusion (defined as an Abbreviated Mental Test score of 8 or less);
**U**rea: raised >7 mmol/l;
**R**espiratory rate: raised > 30/min;
**B**lood 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 severe 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 β-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 over a period from November 2006 to November 2009. Ideally reviews would be conducted in a random sequence once all records had been collected. However, due to the time taken to collect the case-notes and the reporting requirements this was not possible. Therefore, to control for any learning and/or fatigue effect on the part of the reviewer, 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.

Inter-observer agreement on prescribing error was evaluated for the $SPI_1$ evaluation only and was measured by assigning every tenth case-note to a second observer, who is also a qualified pharmacist, and who assessed cases in batches blinded to the others assessment, but compared and discussed results after each batch.

### 6.2.6 Statistical methods

Generalised linear mixed models were used to analyse the effect of the **SPI** intervention. 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 (see Chapter 12). 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 (GEE). 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.

#### *6.2.6.1 Statistical methods used specifically in the $SPI_1$ evaluation*

Fixed effects were included: (a) for differences in pre-intervention levels between control and $\boldsymbol{SPI_1}$ 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 2); and (c) the effect of the **SPI**, interpreted as the difference between the temporal changes pre/post intervention experienced in the control and $\boldsymbol{SPI_1}$ hospitals.

#### *6.2.6.2 Statistical methods used specifically in the $SPI_2$ evaluation*

Within all models, pre-intervention levels were estimated by pooling data from the first two epochs and post-intervention levels were estimated using data from the third epoch. Fixed effects were included: (a) for differences in pre-intervention levels between control and $\boldsymbol{SPI_2}$ hospitals ("baseline comparisons") (b) for temporal changes between Epochs 1 and 2 across all hospitals; (c) the temporal change experienced in the control hospitals between the pre-intervention period (i.e. Epochs 1 and 2 pooled together) and the post-intervention period (Epoch 3); and (d) the effect of the **SPI**, interpreted as the difference between the temporal changes pre/post intervention experienced in the control and $\boldsymbol{SPI_2}$ hospitals.

## 6.3 Results

### 6.3.1 The sample

#### *6.3.1.1 Number of case-notes reviewed within the explicit review in the $SPI_1$ evaluation*

The smallest $SPI_1$ hospital could not identify the target numbers of case-notes, leading to a slight shortfall in the intended $SPI_1$ sample size of 400 case-notes in each epoch: 381 (Epoch 1, before $SPI_1$) and 380 (Epoch 2, after implementation of $SPI_1$). The corresponding numbers for control hospitals are 236 case-notes in Epoch 1 and 240 in Epoch 2.

#### *6.3.1.2 Number of case-notes reviewed within the explicit review in the $SPI_2$ evaluation*

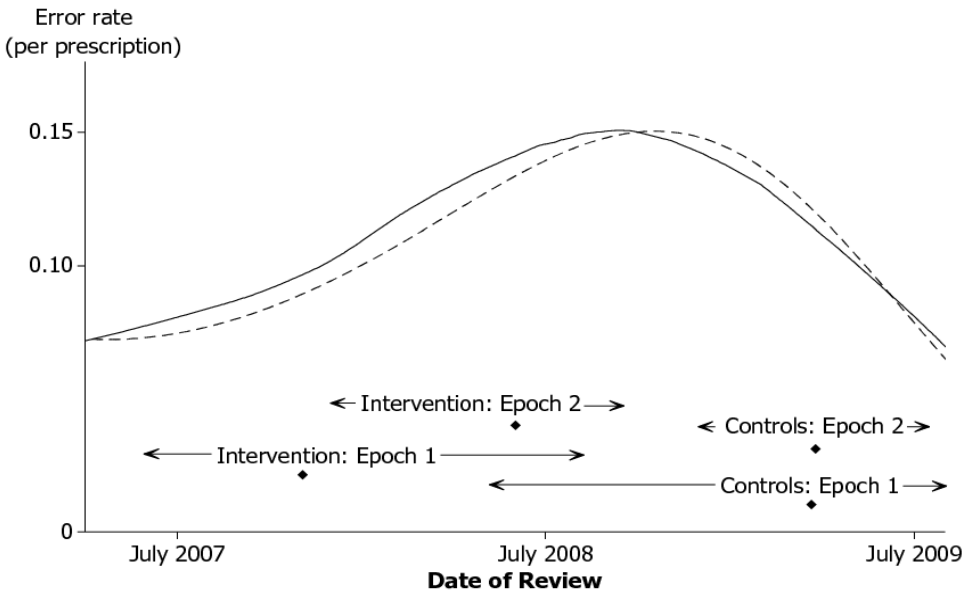
The intended sample size of 405 from the $SPI_2$ hospitals was not met - 347 case-notes were reviewed. These case-notes were split approximately equally across the epochs: 116 from Epoch 1, 117 from Epoch 2 and 114 from Epoch 3. Control hospitals yielded 355 case-notes out of the intended sample size of 405: 120 from Epoch 1, 123 from Epoch 2 and 112 from Epoch 3.

### 6.3.2 Effects associated with the review process

Case-note reviews took place in the period November 2006 to August 2009. The review of $SPI_1$ hospital case-notes was done first, and was 90% complete by August 2008. By contrast,

90% of the ***SPI$_2$*** and 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, though there was considerable overlap between the dates of the two sets of reviews (see Figure 6.2). 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 6.3.8, 6.3.9 and 6.3.10 below. Elsewhere they were found to be significant for just two items not associated with prescribing error (see 6.3.4 below).

**Figure 6.2: Prescription error rates and date of review**




The rates of detected errors (per prescription) are represented by a locally smoothed version of the raw error-rates (solid line) and also by the cubic polynomial (broken line) used in the analysis. Median dates of review (♦), and intervals showing when most (80%) of the reviews were done (← →), are given for each combination of Arm and Epoch. Date of review for Epoch 3 is not shown here but occurred during Dec 2008 to Nov 2009 and learning/fatigue effects from this time have also been incorporated in the analysis.

### 6.3.3 Reliability and anonymisation of case-notes within the $SPI_1$ evaluation

The comparison of prescribing error results between two observers showed substantial inter-observer agreement (Landis and Koch, 1977), with a κ 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.

During the review the primary reviewer was able to detect the hospital of origin in 1% of cases (11/1154), the epoch in 14% of cases (158/1154) and the patients' name in 4% of cases (42/1154).

### 6.3.4 Quality of medical history taking

#### *6.3.4.1 Within the $SPI_1$ evaluation*

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 $SPI_1$ hospitals; neither is there significant evidence of temporal improvement for any item (see Table 6.5).

There was some evidence of a temporal effect in the review process (learning/fatigue effect) for Item 2 ("*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 2, 3, 5, 6, 7 and 9 ($p<0.001$ in all cases), typically reducing the odds of the question being asked by about 5% per year of age.

**Table 6.4: Medical history taking (% of patients asked)**

| | Control Hospitals (N=18) | | | | $SPI_1$ Hospitals (N=4) | | | |
|---|---|---|---|---|---|---|---|---|
| | Epoch 1 | | Epoch 2 | | Epoch 1 | | Epoch 2 | |
| Number 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 |

Entries are percentages with Binomial standard errors (SE). Shaded areas relate to post-intervention epochs

**Table 6.5: Medical history taking – differences between control and intervention hospitals, changes over time and the effect of $SPI_1$**

| | Baseline Comparisons | | Changes in Controls | | Effect of $SPI_1$ | |
|---|---|---|---|---|---|---|
| | | | | | | |
| | OR (99% CI) | p-value | OR (99% CI) | p-value | OR (99% CI) | p-value |
| | | | | | | |
| 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

#### *6.3.4.2 Within the $SPI_2$ evaluation*

Baseline comparisons showed no significant differences between control and $SPI_2$ hospitals. An effect of **SPI** was not apparent and was not statistically significant for any of the items measured. For two items ("*Exercise Tolerance*" and "*Occupation*") measured in relation to history taking there was significant evidence of an improvement overtime in both control and $SPI_2$ hospitals (see Table 6.7). There was some evidence of a reviewer learning/fatigue effect for "*Exercise Tolerance*" ($p<0.001$) "*Chest Pain*" ($p=0.010$) and "*Occupation*" ($p=0.001$).

Again, several of the questions were asked less often for older patients. Age was a significant predictor for items 3, 6 and 7 ($p\leq0.001$ in all cases), typically reducing the odds of the question being asked by about 5% per year of age.

**Table 6.6: Medical history taking (% of patients asked)**

| | Control Hospitals (N=9) | | | | | | $SPI_2$ Hospitals (N=9) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Epoch 1 | | Epoch 2 | | Epoch 3 | | Epoch 1 | | Epoch 2 | | Epoch 3 | |
| Number of Patients | 120 | | 123 | | 112 | | 116 | | 117 | | 114 | |
| | | | | | | | | | | | | |
| | % | SE | % | SE | % | SE | % | SE | % | SE | % | SE |
| | | | | | | | | | | | | |
| 1. Duration of "presenting" symptom | 92.5 | 2.4 | 91.1 | 2.6 | 95.5 | 2.0 | 96.6 | 1.7 | 98.3 | 1.2 | 99.1 | 0.9 |
| 2. Normal exercise tolerance | 26.5 | 4.1 | 31.7 | 4.2 | 38.4 | 4.6 | 38.8 | 4.5 | 38.3 | 4.6 | 33.9 | 4.5 |
| 3. Presence/absence shortness of breath | 88.3 | 2.9 | 91.1 | 2.6 | 88.4 | 3.0 | 91.4 | 2.6 | 93.2 | 2.3 | 92.0 | 2.6 |
| 4. Presence/absence orthopnoea | 23.3 | 3.9 | 28.1 | 4.1 | 17.0 | 3.6 | 32.8 | 4.4 | 29.3 | 4.2 | 18.0 | 3.7 |
| 5. Presence/absence cough | 88.3 | 2.9 | 89.4 | 2.8 | 86.6 | 3.2 | 91.4 | 2.6 | 91.5 | 2.6 | 83.9 | 3.5 |
| 6. If cough, was it productive | 78.3 | 3.8 | 84.6 | 3.3 | 77.7 | 4.0 | 87.1 | 3.1 | 88.0 | 3.0 | 76.8 | 4.0 |
| 7. Smoking history taken | 73.9 | 4.0 | 81.3 | 3.5 | 66.1 | 4.5 | 77.6 | 3.9 | 79.5 | 3.7 | 74.1 | 4.2 |
| 8. Presence/absence of haemoptysis | 22.2 | 3.9 | 28.1 | 4.1 | 16.1 | 3.5 | 25.2 | 4.1 | 23.3 | 3.9 | 26.1 | 4.2 |
| 9. Chest pain (of any type) | 68.1 | 4.3 | 71.5 | 4.1 | 54.5 | 4.7 | 54.3 | 4.6 | 65.5 | 4.4 | 59.8 | 4.7 |
| 10. Occupation/previous occupation | 44.4 | 4.6 | 37.7 | 4.4 | 53.6 | 4.7 | 34.8 | 4.5 | 38.5 | 4.5 | 38.4 | 4.6 |
| 11. Pets | 3.4 | 1.7 | 3.3 | 1.6 | 0.9 | 0.9 | 1.7 | 1.2 | 2.6 | 1.5 | 6.3 | 2.3 |
| | | | | | | | | | | | | |
| % over all items | | 55.7 | | 58.2 | | 54.1 | | 57.5 | | 59.0 | | 57.4 |

Entries are percentages with binomial standard errors.

Shaded areas relate to post-intervention epochs.

**Table 6.7: Medical history taking – differences between control and $SPI_2$ hospitals, changes over time and the effect of $SPI_2$**

| | Baseline Comparisons | | Changes in Controls | | Effect of $SPI_2$ | |
|---|---|---|---|---|---|---|
| | | | | | | |
| | OR (99% CI) | p-value | OR (99% CI) | p-value | OR (99% CI) | p-value |
| | | | | | | |
| 1. Duration of "presenting" symptom | 3.2 (0.7, 14.0) | 0.040 | 1.6 (0.4, 7.3) | 0.391 | 1.7 (0.07, 40.3) | 0.672 |
| 2. Normal exercise tolerance | 1.4 (0.8, 2.4) | 0.125 | 2.2 (1.1, 4.4) | 0.005 | 0.7 (0.3, 1.7) | 0.312 |
| 3. Presence/absence shortness of breath | 1.3 (0.5, 3.5) | 0.480 | 0.8 (0.3, 2.3) | 0.539 | 1.3 (0.3, 5.7) | 0.701 |
| 4. Presence/absence orthopnoea | 1.3 (0.7, 2.5) | 0.330 | 0.6 (0.3, 1.5) | 0.159 | 0.9 (0.3, 2.6) | 0.749 |
| 5. Presence/absence cough | 1.2 (0.5, 2.9) | 0.506 | 0.7 (0.2, 1.8) | 0.286 | 0.7 (0.2, 2.4) | 0.407 |
| 6. If cough, was it productive | 1.4 (0.7, 2.9) | 0.208 | 0.7 (0.3, 1.6) | 0.307 | 0.7 (0.2, 2.1) | 0.418 |
| 7. Smoking history taken | 1.1 (0.5, 2.1) | 0.841 | 0.6 (0.3, 1.2) | 0.061 | 1.5 (0.5, 4.0) | 0.313 |
| 8. Presence/absence of haemoptysis† | 0.9 (0.4, 1.9) | 0.686 | 0.6 (0.2, 1.4) | 0.106 | 2.2 (0.7, 6.5) | 0.061 |
| 9. Chest pain (of any type) | 0.6 (0.4, 1.1) | 0.041 | 0.7 (0.4,1.4) | 0.193 | 2.1 (0.9, 5.2) | 0.028 |
| 10. Occupation/previous occupation | 0.9 (0.5, 1.7) | 0.696 | 2.0 (1.0, 4.0) | 0.010 | 0.6 (0.3, 1.5) | 0.178 |
| 11. Pets | 0.9 (0.2, 4.7) | 0.872 | 0.3 (0.02, 5.6) | 0.299 | 8.3 (0.3, 210.0) | 0.093 |

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

### 6.3.5 Observations and signs of patient deterioration

#### *6.3.5.1 Within the $SPI_1$ evaluation*

Compliance in recording patient observations did improve markedly at both 6 and 12 hours after admission and this was statistically significant for all but one item. Though both control and $\boldsymbol{SPI_1}$ hospitals improved, the improvement was greater in $\boldsymbol{SPI_1}$ hospitals (see Table 6.8) though statistically significant only for the recording of respiratory rate at 12 hours.

.

**Table 6.8: Vital signs – percentage compliance with standards within the $SPI_1$ evaluation**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Epoch 1** | | **Epoch 2** | | **Epoch 1** | | **Epoch 2** | | | | |
| | % | SE | % | SE | % | SE | % | SE | OR (99% CI) | p-value | OR (99% CI) | p-value |
| ***On Admission*** | | | | | | | | | | | | |
| 1. Temperature | 97.9 | 0.9 | 99.2 | 0.6 | 99.0 | 0.5 | 99.2 | 0.5 | 5.1 (0.3, 89.5) | 0.144 | 0.2 (0.01, 8.5) | 0.289 |
| 2. Respiratory rate | 96.2 | 1.3 | 98.8 | 0.7 | 90.8 | 1.5 | 98.4 | 0.6 | 4.7 (0.6, 36.5) | 0.052 | 1.5 (0.2, 16.0) | 0.677 |
| 3. Cyanosis/Oxygen saturation | 98.7 | 0.7 | 98.8 | 0.7 | 97.6 | 0.8 | 99.2 | 0.5 | 1.7 (0.2, 18.2) | 0.578 | 2.7 (0.1, 55.2) | 0.385 |
| 4. Confusion/Mental state | 57.9 | 3.2 | 64.6 | 3.1 | 66.7 | 2.4 | 68.9 | 2.4 | 1.2 (0.7, 2.1) | 0.307 | 1.2 (0.6, 2.6) | 0.437 |
| 5. Pulse | 98.7 | 0.7 | 99.2 | 0.6 | 99.0 | 0.5 | 99.5 | 0.4 | 3.3 (0.2, 68.7) | 0.306 | 0.5 (0.01, 23.7) | 0.614 |
| 6. Blood pressure | 98.7 | 0.7 | 99.2 | 0.6 | 99.0 | 0.5 | 99.5 | 0.4 | 3.3 (0.2, 68.7) | 0.306 | 0.5 (0.01, 23.7) | 0.614 |
| ***At 6 Hours*** | | | | | | | | | | | | |
| 7. Temperature | 62.4 | 3.2 | 73.8 | 2.8 | 75.9 | 2.2 | 86.1 | 1.8 | 1.7 (1.0, 2.9) | 0.008 | 1.0 (0.5, 2.2) | 0.976 |
| 8. Respiratory rate | 44.0 | 3.3 | 72.5 | 2.9 | 42.8 | 2.5 | 81.6 | 2.0 | 3.6 (2.1,6.2) | < 0.001 | 2.0 (1.0, 4.2) | 0.015 |
| 9. Pulse | 67.1 | 3.1 | 77.1 | 2.7 | 82.7 | 1.9 | 88.2 | 1.7 | 1.7 (1.0, 3.0) | 0.012 | 1.0 (0.4, 2.3) | 0.973 |
| 10. Oxygen saturation | 61.1 | 3.2 | 75.0 | 2.8 | 76.6 | 2.2 | 87.6 | 1.7 | 1.9 (1.1, 3.3) | 0.002 | 1.3 (0.6, 2.8) | 0.425 |
| ***At 12 Hours*** | | | | | | | | | | | | |
| 11. Temperature | 58.5 | 3.2 | 70.3 | 3.0 | 70.6 | 2.3 | 82.4 | 2.0 | 1.8 (1.0, 3.0) | 0.005 | 1.3 (0.6, 2.8) | 0.314 |
| 12. Respiratory rate | 39.7 | 3.2 | 68.8 | 3.0 | 37.0 | 2.5 | 77.9 | 2.1 | 3.7 (2.2, 6.2) | < 0.001 | 2.1 (1.0, 4.3) | 0.008 |
| 13. Pulse | 61.5 | 3.2 | 73.8 | 2.8 | 75.3 | 2.2 | 83.2 | 1.9 | 1.9 (1.1, 3.2) | 0.002 | 1.2 (0.5, 2.5) | 0.618 |
| 14. Oxygen saturation | 55.6 | 3.3 | 73.3 | 2.9 | 64.0 | 2.5 | 81.8 | 2.0 | 2.3 (1.4, 3.9) | < 0.001 | 1.4 (0.7, 2.9) | 0.234 |

| ***Routine Investigations*** | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15. Urea & Electrolytes | 99.6 | 0.4 | 98.8 | 0.7 | 98.7 | 0.6 | 98.7 | 0.6 | 0.6 (0.02, 14.1) | 0.665 | 0.8 (0.02, 39.5) | 0.865 |
| 16. Chest X-ray | 96.6 | 1.2 | 97.9 | 0.9 | 94.5 | 1.2 | 93.7 | 1.2 | 2.4 (0.5, 11.8) | 0.164 | 0.5 (0.1, 3.0) | 0.291 |
| 17. Full Blood Count | 98.7 | 0.7 | 98.3 | 0.8 | 99.0 | 0.5 | 98.2 | 0.7 | 1.2 (0.1, 10.5) | 0.789 | 0.2 (0.01, 5.4) | 0.223 |

Entries are percentages, with Binomial standard errors (SE)
Shaded areas relate to post-intervention epochs

#### *6.3.5.2 Within the $SPI_2$ evaluation*

There is no significant evidence for an effect associated with **SPI** (Table 6.9 and Table 6.10). However, compliance in taking patient observations at 6 and 12 hours after admission also improved in both groups of hospitals when Epochs 1 and 2 are compared to Epoch 3. This was most evident for "*Respiratory Rate*" where practice continued to improve across all three epochs. In addition improvement took place between the first two epochs on these and most of the other 6 and 12 hour items ($p<0.010$ for all items except for 6 hour "*Pulse"*, for which $p=0.016$).

**Table 6.9: Vital signs- percentage compliance with standards within the $SPI_2$ evaluation**

| | Control Hospitals (N=9) | | | | | | $SPI_2$ Hospitals (N=9) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Epoch 1 | | Epoch 2 | | Epoch 3 | | Epoch 1 | | Epoch 2 | | Epoch 3 | |
| | | | | | | | | | | | | |
| | % | SE | % | SE | % | SE | % | SE | % | SE | % | SE |
| ***On Admission*** | | | | | | | | | | | | |
| 1. Temperature | 96.7 | 1.6 | 99.2 | 0.8 | 99.1 | 0.9 | 99.1 | 0.9 | 99.1 | 0.9 | 96.5 | 1.7 |
| 2. Respiratory rate | 95.8 | 1.8 | 99.2 | 0.8 | 100.0 | 0.0 | 96.5 | 1.7 | 98.3 | 1.2 | 100.0 | 0.0 |
| 3. Cyanosis/Oxygen saturation | 98.3 | 1.2 | 98.4 | 1.1 | 100.0 | 0.0 | 99.1 | 0.9 | 99.1 | 0.9 | 100.0 | 0.0 |
| 4. Confusion/Mental state | 53.3 | 4.6 | 71.5 | 4.1 | 74.1 | 4.2 | 62.6 | 4.5 | 57.3 | 4.6 | 80.7 | 3.7 |
| 5. Pulse | 98.3 | 1.2 | 99.2 | 0.8 | 100.0 | 0.0 | 99.1 | 0.9 | 99.1 | 0.9 | 100.0 | 0.0 |
| 6. Blood pressure | 98.3 | 1.2 | 99.2 | 0.8 | 100.0 | 0.0 | 99.1 | 0.9 | 99.1 | 0.9 | 100.0 | 0.0 |
| ***At 6 Hours*** | | | | | | | | | | | | |
| 7. Temperature | 61.7 | 4.5 | 69.9 | 4.2 | 69.6 | 4.4 | 63.2 | 4.5 | 77.8 | 3.9 | 68.1 | 4.4 |
| 8. Respiratory rate | 40.8 | 4.5 | 69.1 | 4.2 | 72.3 | 4.2 | 47.4 | 4.7 | 76.1 | 4.0 | 77.9 | 3.9 |
| 9. Pulses | 69.2 | 4.2 | 73.2 | 4.0 | 75.0 | 4.1 | 64.9 | 4.5 | 81.2 | 3.6 | 79.6 | 3.8 |
| 10. Oxygen saturation | 61.7 | 4.5 | 71.5 | 4.1 | 74.1 | 4.2 | 60.5 | 4.6 | 78.6 | 3.8 | 79.6 | 3.8 |
| ***At 12 Hours*** | | | | | | | | | | | | |
| 11. Temperature | 58.3 | 4.5 | 70.7 | 4.1 | 68.8 | 4.4 | 58.8 | 4.6 | 69.8 | 4.3 | 72.6 | 4.2 |
| 12. Respiratory rate | 35.0 | 4.4 | 69.9 | 4.2 | 73.2 | 4.2 | 44.7 | 4.7 | 67.5 | 4.3 | 78.8 | 3.9 |
| 13. Pulse | 63.3 | 4.4 | 76.4 | 3.8 | 75.0 | 4.1 | 59.6 | 4.6 | 70.9 | 4.2 | 79.6 | 3.8 |
| 14. Oxygen saturation | 54.2 | 4.6 | 75.6 | 3.9 | 74.1 | 4.2 | 57.0 | 4.7 | 70.9 | 4.2 | 79.6 | 3.8 |

| ***Routine Investigations*** | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15. U & E | 99.2 | 0.8 | 98.4 | 1.1 | 99.1 | 0.9 | 100.0 | 0.0 | 99.1 | 0.9 | 100.0 | 0.0 |
| 16. Chest X-ray | 96.7 | 1.6 | 97.6 | 1.4 | 97.3 | 1.5 | 96.5 | 1.7 | 98.3 | 1.2 | 100.0 | 0.0 |
| 17. Full Blood Count | 98.3 | 1.2 | 97.6 | 1.4 | 99.1 | 0.9 | 99.1 | 0.9 | 99.1 | 0.9 | 100.0 | 0.0 |

Shaded areas relate to post-intervention epochs

**Table 6.10: Vital signs– differences between control and $SPI_2$ hospitals, changes over time and the effect of $SPI_2$**

| | Baseline Comparisons | | Changes in Controls | | Effect of $SPI_2$ | |
|---|---|---|---|---|---|---|
| | | | | | | |
| | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** |
| ***On Admission*** | | | | | | |
| 1. Temperature | 2.2 (0.2, 21.1) | 0.381 | 0.7 (0.02, 24.0) | 0.823 | 0.1 (0.002, 4.1) | 0.108 |
| 2. Respiratory rate | 0.7 (0.1, 3.9) | 0.617 | - | - | - | - |
| 3. Cyanosis/Oxygen saturation | 1.6 (0.1, 18.2) | 0.605 | - | - | - | - |
| 4. Confusion/Mental state | 0.9 (0.5, 1.7) | 0.674 | 1.8 (0.8,3.7) | 0.045 | 1.7 (0.6, 4.5) | 0.187 |
| 5. Pulse | 1.1 (0.1, 14.9) | 0.942 | - | - | - | - |
| 6. Blood pressure | 1.1 (0.1, 14.9) | 0.942 | - | - | - | - |
| ***At 6 Hours*** | | | | | | |
| 7. Temperature | 1.3 (0.7, 2.4) | 0.323 | 1.4 (0.7, 2.8) | 0.239 | 0.8 (0.3, 1.9) | 0.457 |
| 8. Respiratory rate | 1.3 (0.7, 2.5) | 0.281 | 2.1 (1.0, 4.3) | 0.010 | 1.0 (0.4, 2.8) | 0.907 |
| 9. Pulse | 1.1 (0.6, 2.1) | 0.604 | 1.3 (0.6, 2.8) | 0.327 | 1.2 (0.4, 3.3) | 0.662 |
| 10. Oxygen saturation | 1.2 (0.7, 2.2) | 0.433 | 1.4 (0.7, 3.0) | 0.223 | 1.2 (0.4, 3.1) | 0.703 |
| ***At 12 Hours*** | | | | | | |
| 11. Temperature | 1.0 (0.6, 1.8) | 0.934 | 1.2 (0.6, 2.4) | 0.583 | 1.2 (0.5, 2.9) | 0.685 |
| 12. Respiratory rate | 1.2 (0.6, 2.3) | 0.524 | 2.4 (1.1, 5.0) | 0.002 | 1.2 (0.4, 3.1) | 0.713 |
| 13. Pulse | 0.8 (0.5, 1.4) | 0.394 | 1.2 (0.6, 2.5) | 0.510 | 1.5 (0.6, 4.1) | 0.268 |
| 14. Oxygen saturation | 1.0 (0.6, 1.7) | 0.953 | 1.4 (0.7, 2.9) | 0.231 | 1.4 (0.5, 3.6) | 0.430 |

| ***Routine Investigations*** | | | | | | |
|---|---|---|---|---|---|---|
| 15. U & E | 0.9 (0.03, 28.8) | 0.944 | 0.6 (0.01, 27.7) | 0.762 | - | - |
| 16. Chest X-ray | 1.1 (0.2, 5.1) | 0.904 | 0.7 (0.1, 5.6) | 0.641 | - | - |
| 17. Full Blood Count | 1.6 (0.2, 16.9) | 0.609 | 1.7 (0.1, 40.4) | 0.663 | - | - |

No items showed significant variation between hospitals within arms. Blanks are associated with 100% compliance for which logistic regression analysis is impossible.

### 6.3.6 Appropriate clinical response to abnormal vital signs

#### *6.3.6.1 Within the $SPI_1$ evaluation*

The data are summarised in Table 6.11 and Table 6.12. 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 0.10 (= 0.09 in both cases) for only two items.

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

**Table 6.11: Appropriate clinical response – compliance with standards**

| | Control Hospitals (N=18) | | | | | | $SPI_1$ Hospitals (N=4) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 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 ICU (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 by 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 |

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.
Shaded areas relate to post-intervention epochs

**Table 6.12: Appropriate clinical response - differences between control and intervention hospitals, changes over time and the effect of $SPI_1$**

| | Comparisons at Epoch 1 | | Change in Controls | | Effect of $SPI_1$ | |
|---|---|---|---|---|---|---|
| | $SPI_1$/Control | | Epoch 2/ Epoch 1 | | Ratio of temporal changes | |
| | OR (99% CI) | p-value | OR (99% CI) | p-value | OR (99% CI) | p-value |
| **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 oxygen (if not on oxygen) | 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 ICU (if on oxygen) | 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 |

#### *6.3.6.2 Within the $SPI_2$ evaluation*

Again the data was sparse (Table 6.13 and Table 6.14), and formal analysis was possible for only three items. No significant conclusions were indicated.

**Table 6.13: Appropriate clinical response**

| | Control Hospitals (N=18) | | | | | | | | | $SPI_2$ Hospitals (N=4) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Epoch 1 | | | Epoch 2 | | | Epoch 3 | | | Epoch 1 | | | Epoch 2 | | | Epoch 3 | | |
| | N | % | SE | N | % | SE | N | % | SE | N | % | SE | N | % | SE | N | % | SE |
| ***Oxygen saturation <90 at any time*** | | | | | | | | | | | | | | | | | | |
| Full blood gases within 2 hours | 13 | 61.5 | 13.5 | 10 | 50.0 | 15.8 | 0 | - | - | 2 | 50.0 | 35.4 | 10 | 70.0 | 14.5 | 4 | 25.0 | 21.7 |
| Given oxygen (if not on oxygen) | 12 | 66.7 | 13.6 | 7 | 57.1 | 18.7 | 1 | 0.0 | 0.0 | 4 | 75.0 | 21.7 | 9 | 77.8 | 13.9 | 2 | 50.0 | 35.4 |
| Doctor called or transferred to ICU (if on oxygen) | 8 | 25.0 | 15.3 | 6 | 50.0 | 20.4 | 0 | - | - | 2 | 50.0 | 35.4 | 5 | 80.0 | 17.9 | 2 | 50.0 | 35.4 |
| ***Blood pressure systolic <90*** | | | | | | | | | | | | | | | | | | |
| At least next 6 hours, hourly observations | 7 | 28.6 | 17.0 | 8 | 25.0 | 15.3 | 8 | 50.0 | 17.7 | 4 | 50.0 | 25.0 | 6 | 16.7 | 15.2 | 2 | 100.0 | 0.0 |
| Blood culture | 4 | 50.0 | 25.0 | 5 | 40.0 | 21.2 | 8 | 37.5 | 17.1 | 4 | 25.0 | 21.7 | 5 | 80.0 | 17.9 | 2 | 100.0 | 0.0 |
| ***Sputum Present*** | | | | | | | | | | | | | | | | | | |
| Sputum culture | 70 | 41.4 | 5.9 | 72 | 48.6 | 5.9 | 69 | 24.6 | 5.2 | 71 | 36.6 | 5.8 | 78 | 46.2 | 5.7 | 62 | 29.0 | 5.8 |
| ***Respiratory rate >20 at any time after admission*** | | | | | | | | | | | | | | | | | | |
| Given oxygen (if not on oxygen) | 3 | 0.0 | 0.0 | 0 | . | . | 0 | - | - | 2 | 0.0 | 0.0 | 1 | 0.0 | 0.0 | 0 | - | - |
| Doctor called (if on oxygen) | 5 | 0.0 | 0.0 | 1 | 100.0 | 0.0 | 0 | - | - | 3 | 0.0 | 0.0 | 2 | 0.0 | 0.0 | 3 | 0.0 | 0.0 |
| ***Temperature over 38°C – any episode*** | | | | | | | | | | | | | | | | | | |
| If yes, blood culture | 16 | 68.8 | 11.6 | 14 | 71.4 | 12.0 | 15 | 73.3 | 11.4 | 19 | 73.7 | 10.1 | 25 | 76.0 | 8.5 | 13 | 61.5 | 13.5 |

| ***Failure to improve by 48 hours or subsequent deterioration*** | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Review by consultant | 11 | 100.0 | 0.0 | 12 | 100.0 | 0.0 | 10 | 100.0 | 0.0 | 9 | 100.0 | 0.0 | 10 | 100.0 | 0.0 | 3 | 100.0 | 0.0 |
| Repeat chest x-ray | 10 | 100.0 | 0.0 | 9 | 100.0 | 0.0 | 9 | 100.0 | 0.0 | 8 | 100.0 | 0.0 | 8 | 100.0 | 0.0 | 3 | 100.0 | 0.0 |
| White cell counted/repeated | 11 | 100.0 | 0.0 | 12 | 100.0 | 0.0 | 11 | 100.0 | 0.0 | 8 | 100.0 | 0.0 | 10 | 100.0 | 0.0 | 3 | 100.0 | 0.0 |
| Appropriate addition of further antibiotics | 9 | 100.0 | 0.0 | 5 | 100.0 | 0.0 | 8 | 75.0 | 15.3 | 7 | 85.7 | 13.2 | 6 | 100.0 | 0.0 | 2 | 50.0 | 35.4 |
| ***Follow up*** | | | | | | | | | | | | | | | | | | |
| Arrange follow up? | 45 | 71.1 | 6.7 | 47 | 61.7 | 7.1 | 38 | 42.1 | 8.0 | 49 | 59.2 | 7.0 | 52 | 63.5 | 6.6 | 44 | 38.6 | 7.3 |

**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.

Shaded areas relate to post-intervention epochs

**Table 6.14: Appropriate clinical response – difference between control and $SPI_2$ hospitals, changes over time and the effect of $SPI_2$ (formal analyses for three items only, because of sparse data for other items)**

| | Baseline Comparisons | | Changes in Controls | | Effect of $SPI_2$ | |
|---|---|---|---|---|---|---|
| | | | | | | |
| | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** |
| ***Sputum Present*** | | | | | | |
| Sputum Culture | 0.8 (0.4, 1.6) | 0.411 | 0.5 (0.2, 1.2) | 0.040 | 1.7 (0.5, 6.0) | 0.250 |
| ***Temperature >38°C*** | | | | | | |
| If Yes, Blood Culture | 1.0 (0.2, 4.5) | 0.969 | 0.9 (0.1, 7.1) | 0.874 | 0.6 (0.04, 9.6) | 0.636 |
| ***Appropriate Follow-up*** | | | | | | |
| Clinical Review Arranged if appropriate | 0.7 (0.3, 1.7) | 0.343 | 0.3 (0.1, 1.0) | 0.009 | 1.2 (0.3, 5.4) | 0.698 |

### 6.3.7 Other features of good care for specific classes of patients

#### *6.3.7.1 Within the $SPI_1$ evaluation*

The effect of **SPI** is not apparent, and is not statistically significant for any item (Table 6.15 and

Table 6.16). The baseline comparisons showed no significant differences between control and $\boldsymbol{SPI_1}$ hospitals. However, it should be noted that the use of CURB score (a clinical prediction rule for predicting mortality from community-acquired pneumonia and infection at any site) has improved significantly in control hospitals over time, albeit from a very low base.

A strong negative age-effect was apparent for item 3 (*"severity of pneumonia"*) yielding a reduction in odds of compliance of about 8% per year of age.

**Table 6.15: Features of good care for specific classes of patient**

| | Control Hospitals (N=18) | | | | | | $SPI_1$ Hospitals (N=4) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 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 |

Shaded areas relate to post-intervention epochs

**Table 6.16: Features of good care for specific classes of patient – differences between control and intervention hospitals, changes over time and the effect of $SPI_1$**

| | Baseline Comparisons | | Changes in Controls | | Effect of $SPI_1$ | |
|---|---|---|---|---|---|---|
| | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** |
| 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 ($p<0.01$)

#### *6.3.7.2 Within the $SPI_2$ evaluation*

There is no significant evidence that the **SPI** had an effect (Table 6.17 and Table 6.18). Use of the CURB score again improved significantly over time (OR=7.3; 1.4 - 37.7), though from a very low base and differences were not statistically significant between control and $SPI_2$ hospitals. A negative age-effect (p<0.001) was apparent for item 3 (*"severity of pneumonia"*) yielding a reduction in odds of compliance of about 6% per year of age.

.

**Table 6.17: Use of steroids and antibiotics, CURB score and other standards applicable to specific cases – compliance with standards**

| | Control Hospitals (N=9) | | | | | | | | | $SPI_2$ Hospitals (N=9) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Epoch 1 | | | Epoch 2 | | | Epoch 3 | | | Epoch 1 | | | Epoch 2 | | | Epoch 3 | | |
| | N | % | SE | N | % | SE | N | % | SE | N | % | SE | N | % | SE | N | % | SE |
| | | | | | | | | | | | | | | | | | | |
| 1. Asthma or COPD given steroids within 24 hrs | 70 | 84.3 | 4.4 | 63 | 91.8 | 3.5 | 56 | 92.9 | 3.5 | 59 | 91.5 | 3.7 | 74 | 93.2 | 2.9 | 53 | 94.3 | 3.2 |
| 2. Peak flow record | 10 | 80.0 | 12.6 | 11 | 63.6 | 14.5 | 5 | 40.0 | 21.9 | 24 | 79.2 | 8.3 | 18 | 94.4 | 5.4 | 8 | 75.0 | 15.3 |
| 3. Severity of pneumonia patients recorded in notes? | 52 | 73.1 | 6.2 | 68 | 70.6 | 5.6 | 57 | 77.2 | 5.6 | 49 | 77.6 | 5.9 | 45 | 77.8 | 6.3 | 60 | 70.0 | 6.0 |
| 4. CURB score recorded in notes? | 52 | 1.9 | 1.9 | 67 | 22.4 | 5.1 | 56 | 21.4 | 5.5 | 50 | 2.0 | 2.0 | 44 | 25.0 | 6.1 | 60 | 41.7 | 6.4 |
| 5. Was appropriate antibiotic treatment given? | 51 | 94.1 | 3.3 | 68 | 92.6 | 3.2 | 53 | 96.2 | 2.6 | 49 | 91.8 | 4.0 | 42 | 100.0 | 0.0 | 55 | 94.5 | 3.1 |

COPD: Chronic Obstructive Pulmonary Disease

CURB: Confusion/Urea / Respiratory rate / Blood pressure score

Shaded areas relate to post-intervention epochs

**Table 6.18: Steroids and antibiotics, CURB score and other standards applicable to specific cases – differences between control and $SPI_2$ hospitals, changes over time and the effect of $SPI_2$**

| | Baseline Comparisons | | Changes in Controls | | Effect of $SPI_2$ | |
|---|---|---|---|---|---|---|
| | | | | | | |
| | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** |
| 1. Asthma or COPD given steroids within 24 hrs | 1.8 (0.6, 5.7) | 0.183 | 0.9 (0.2, 4.8) | 0.813 | 0.6 (0.05, 6.8) | 0.568 |
| 2. Peak flow record | 1.1 (0.03, 40.9) | 0.954 | 0.1 (0.001, 13.5) | 0.255 | 29.7 (0.1, 15943) | 0.165 |
| 3. Severity of pneumonia patients recorded in notes? | 0.9 (0.3, 3.2) | 0.829 | 0.9 (0.3, 3.1) | 0.821 | 0.7 (0.1, 3.0) | 0.478 |
| 4. CURB score recorded in notes? | 1.4 (0.4, 4.9) | 0.453 | 7.3 (1.4, 37.7) | 0.002 | 2.1 (0.4, 11.1) | 0.236 |
| 5. Was appropriate antibiotic treatment given? | 1.4 (0.2, 10.5) | 0.676 | 1.5 (0.1, 15.7) | 0.665 | 0.5 (0.02, 10.0) | 0.519 |

No items showed significant variation between hospitals within arms.

COPD: Chronic Obstructive Pulmonary Disease

CURB: Confusion/Urea / Respiratory rate / Blood pressure score

### 6.3.8 Prescribing error

#### *6.3.8.1 Within the $SPI_1$ evaluation*

The results are presented in Table 6.19 and Table 6.20. There are more prescriptions per patient in the $SPI_1$ hospitals (29.3), compared to the control hospitals (24.9) and a small increase (about 2 per patient) across epochs in both arms. Unadjusted analysis suggests an increases 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 in the period 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. $SPI_1$ hospitals were reviewed while it was increasing. 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. Thus 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**. (Examples of prescribing errors are given in Table 6.21).

**Table 6.19: Prescribing errors**

| | Control Hospitals (N=18) | | | | *SPI1* Hospitals (N=4) | | | |
|---|---|---|---|---|---|---|---|---|
| | Epoch 1 | | Epoch 2 | | Epoch 1 | | Epoch 2 | |
| | | | | | | | | |
| Number of patients* | 233 | | 239 | | 381 | | 378 | |
| Number 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 |

*Number of patients with medication charts available to review

Shaded areas relate to post-intervention epochs

**Table 6.20: Prescribing Errors – differences between control and intervention hospitals, changes over time and the effect of $SPI_1$**

| | Baseline Comparisons | | Changes in Controls | | Effect of $SPI_1$ | |
|---|---|---|---|---|---|---|
| | **Rate Ratio (99% CI)** | **p-value** | **Rate Ratio (99% CI)** | **p-value** | **Rate Ratio (99% CI)** | **p-value** |
| 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 6.21: Examples of prescribing errors relating to each stage of the drug use process found within the $SPI_1$ evaluation**

| 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<br>• 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 2 puffs four times a day. This is drug duplication as both of these drugs have the same pharmacological action<br>• 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 4 puffs four times a day .This was a wrong dose (overdose) as the maximum should have been 2 puffs four times a day<br>• Paracetamol 1g oral to be given "when required" prescribed without indicating the maximum daily frequency/dose |
| Selection of formulation | • Seretide 250 2 puffs inhaler twice a day prescribed without specifying whether evohaler or accuhaler<br>• Dipyridamole 200mg orally twice a day prescribed without indicating that modified release formulation intended |
| Provide information needed for supply | • Co-amoxiclav 625mg three time a day prescribed without indicating the route of administration<br>• Clopidogrel 75mg oral once a day prescribed and given without having a signature of prescriber |

#### *6.3.8.2 Within the $SPI_2$ evaluation*

A reviewer learning/fatigue effect was significant (p=0.009) in the review of prescribing errors (Table 6.22 and Table 6.23), with a decreasing rate of error detection with time of review; this was allowed for in the analysis. No significant time effects for **SPI** arm, Time or **SPI** were detected (Table 6.23).

**Table 6.22: Prescribing Errors**

| | Control Hospitals (N=9) | | | | | | $SPI_2$ Hospitals (N=9) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Epoch 1 | | Epoch 2 | | Epoch 3 | | Epoch 1 | | Epoch 2 | | Epoch 3 | |
| | | | | | | | | | | | | |
| Number of patients[†] | 120 | | 122 | | 112 | | 113 | | 117 | | 114 | |
| Number of prescriptions | 2953 | | 3269 | | 2871 | | 2529 | | 2938 | | 2656 | |
| Prescriptions per patient | 24.6 | | 26.8 | | 25.6 | | 22.4 | | 25.1 | | 23.3 | |
| ***Errors*** | | | | | | | | | | | | |
| Total | 345 | | 298 | | 216 | | 251 | | 266 | | 167 | |
| | | | | | | | | | | | | |
| | **Rate** | **SE** | **Rate** | **SE** | **Rate** | **SE** | **Rate** | **SE** | **Rate** | **SE** | **Rate** | **SE** |
| ***Unadjusted Rates*** | | | | | | | | | | | | |
| Error rate per prescription | 0.120 | 0.017 | 0.092 | 0.013 | 0.076 | 0.012 | 0.106 | 0.015 | 0.088 | 0.013 | 0.060 | 0.010 |
| | | | | | | | | | | | | |
| ***Rates adjusted for date of review*** | | | | | | | | | | | | |
| Overall rate (all errors) | 0.103 | 0.017 | 0.086 | 0.013 | 0.089 | 0.015 | 0.098 | 0.015 | 0.084 | 0.013 | 0.073 | 0.014 |
| | | | | | | | | | | | | |
| ***By 5 most prevalent stages of the drug use process*** | | | | | | | | | | | | |
| Need for drug therapy | 0.006 | 0.002 | 0.018 | 0.003 | 0.018 | 0.004 | 0.014 | 0.003 | 0.011 | 0.002 | 0.015 | 0.002 |
| Selection of dose | 0.057 | 0.009 | 0.031 | 0.005 | 0.032 | 0.006 | 0.039 | 0.006 | 0.039 | 0.006 | 0.030 | 0.006 |
| Selection of drug | 0.002 | 0.001 | 0.002 | 0.001 | 0.001 | 0.001 | 0.004 | 0.001 | 0.002 | 0.001 | 0.001 | 0.001 |
| Selection of formulation | 0.010 | 0.003 | 0.005 | 0.001 | 0.005 | 0.002 | 0.007 | 0.002 | 0.008 | 0.002 | 0.009 | 0.003 |
| Provide information needed for supply | 0.029 | 0.006 | 0.029 | 0.005 | 0.035 | 0.007 | 0.032 | 0.006 | 0.026 | 0.005 | 0.025 | 0.006 |

[†]The number of patients are those with medication charts available for review. Shaded areas relate to post-intervention epochs

**Table 6.23: Prescribing errors – differences between control and $SPI_2$ hospitals, changes over time and the effect of $SPI_2$**

| | Baseline Comparisons | | Changes in Controls | | Effect of $SPI_2$ | |
|---|---|---|---|---|---|---|
| | | | | | | |
| | **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) | 0.860 | 0.9 (0.6, 1.4) | 0.662 | 0.9 (0.5, 1.5) | 0.444 |
| | | | | | | |
| ***By 5 most prevalent stages of the drug use process*** | | | | | | |
| Need for drug therapy | 1.0 (0.6, 1.6) | 0.825 | 1.5 (0.7, 2.9) | 0.157 | 0.8 (0.3, 2.1) | 0.626 |
| Selection of dose | 0.9 (0.6, 1.4) | 0.595 | 0.8 (0.5, 1.3) | 0.166 | 1.0 (0.5, 2.0) | 0.982 |
| Selection of drug | 1.2 (0.5, 2.9) | 0.670 | 0.7 (0.1, 4.2) | 0.643 | 0.7 (0.05, 9.3) | 0.687 |
| Selection of formulation | 1.1 (0.6, 2.0) | 0.788 | 0.8 (0.3, 2.0) | 0.506 | 1.6 (0.5, 5.3) | 0.277 |
| Provide information needed for supply | 1.0 (0.6, 1.8) | 0.842 | 1.1 (0.7, 1.9) | 0.556 | 0.7 (0.3, 1.5) | 0.220 |

### 6.3.9 Anti-coagulant prescribing errors

#### *6.3.9.1 Within the $SPI_1$ evaluation*

A specific breakdown of errors relating to anti-coagulant administration was carried out because this treatment was particularly stressed by IHI (Table 6.24 and Table 6.25). No differences were observed, but the denominators are small, especially in control hospitals.

#### *6.3.9.2 Within the $SPI_2$ evaluation*

A total of ten errors were recorded. Six occurred in $SPI_2$ hospitals before the introduction of the intervention, the other four in control hospitals in Epoch 3. The breakdown is shown in Table 6.26, but no further analysis was possible.

**Table 6.24: Anti-coagulant prescribing errors**

| | Control Hospitals (N=18) | | | | *SPI1* Hospitals (N=4) | | | |
|---|---|---|---|---|---|---|---|---|
| | Epoch 1 | | Epoch 2 | | Epoch 1 | | Epoch 2 | |
| | | | | | | | | |
| Number of patients* | 52 | | 93 | | 167 | | 224 | |
| Number of prescriptions | 83 | | 132 | | 274 | | 362 | |
| ***Number 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 |

* Number of patients receiving anti-coagulant therapy
Shaded areas relate to post-intervention epochs

**Table 6.25: Anti-coagulant prescribing errors – analysis**

| | Baseline Comparisons | | Changes in Controls | | Effect of *SPI1* | |
|---|---|---|---|---|---|---|
| | **Rate Ratio (99% CI)** | **p-value** | **Rate Ratio (99% CI)** | **p-value** | **Rate Ratio (99% CI)** | **p-value** |
| 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 |

**Table 6.26: Anti-coagulant prescribing errors *SPI2***

| | Control Hospitals (N=9) | | | *SPI2* Hospitals (N=9) | | |
|---|---|---|---|---|---|---|
| | **Epoch 1** | **Epoch 2** | **Epoch 3** | **Epoch 1** | **Epoch 2** | **Epoch 3** |
| | | | | | | |
| Number of patients | 28 | 42 | 62 | 31 | 64 | 74 |
| Number of prescriptions | 43 | 61 | 92 | 54 | 92 | 99 |
| Number of errors | 0 | 0 | 4 | 1 | 5 | 0 |

Shaded areas relate to post-intervention epoch

### 6.3.10 Medicines reconciliation errors at admission

#### *6.3.10.1 Within the $SPI_1$ evaluation*

The results can be found in Table 6.27 and Table 6.28. 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 $SPI_1$ hospitals.

#### *6.3.10.2 Within the $SPI_2$ evaluation*

The results can be found in Table 6.29 and Table 6.30. Again, there is no significant evidence that the **SPI** has an effect (p=0.914).

**Table 6.27: Reconciliation errors at admission**

| | Control Hospitals (N=18) | | *SPI1* Hospitals (N=4) | |
|---|---|---|---|---|
| | **Epoch 1** | **Epoch 2** | **Epoch 1** | **Epoch 2** |
| | | | | |
| Number 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 Number of Errors when Error is present (SE) | 1.6 (0.2) | 2.2 (0.3) | 2.3 (0.4) | 2.1 (0.2) |

*Number of patients with admission medication charts available to review

**Table 6.28: Reconciliation errors at admission – differences between control and intervention, changes over time and the effect of *SPI1***

| | Comparisons at Epoch 1 | | Change in Controls | | Effect of *SPI1* | |
|---|---|---|---|---|---|---|
| | ***SPI1*/Control** | | **Epoch 2/ Epoch 1** | | **Ratio of temporal changes** | |
| | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** |
| | | | | | | |
| 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

**Table 6.29: Reconciliation errors at admission $SPI_2$**

| | Control Hospitals (N=9) | | | $SPI_2$ Hospitals (N=9) | | |
|---|---|---|---|---|---|---|
| | **Epoch 1** | **Epoch 2** | **Epoch 3** | **Epoch 1** | **Epoch 2** | **Epoch 3** |
| | | | | | | |
| Number of admissions | 120 | 122 | 112 | 113 | 117 | 114 |
| | | | | | | |
| ***Admissions with reconciliation errors*** | | | | | | |
| N | 8 | 14 | 10 | 7 | 8 | 6 |
| % (SE) | 6.7 (2.3) | 11.5 (2.9) | 8.9 (2.7) | 6.2 (2.3) | 6.8 (2.3) | 5.3 (2.1) |
| | | | | | | |
| Mean number of errors when error is present (SE) | 1.4 (0.3) | 2.1 (0.3) | 2.2 (0.6) | 2.9 (1.1) | 2.3 (0.7) | 2.3 (0.6) |

Shaded areas relate to post-intervention epochs

**Table 6.30: Reconciliation errors at admission – differences between control and $SPI_2$, changes over time and the effect of $SPI_2$**

| | Baseline comparisons | | Change in Controls | | Effect of $SPI_2$ | |
|---|---|---|---|---|---|---|
| | | | | | Ratio of temporal changes | |
| | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** |
| | | | | | | |
| Admissions with reconciliation errors | 1.2 (0.3, 4.8) | 0.727 | 1.5 (0.6, 3.8) | 0.292 | 0.9 (0.3, 3.2) | 0.914 |

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

## 6.4 Discussion

This sub-study was a substantial undertaking with 1,463 episodes of care reviewed by the retrospective explicit case-note review method. Case-notes were reviewed with respect to areas that have been targeted by the **SPI** – observations of the deteriorating patient and prescribing errors (particularly anticoagulation prescribing and medicines reconciliation). Observations were also made in areas which might be expected to improve if the overall goal of strengthening the system and achieving cultural and organisational realignments around safety had been achieved, such as the quality of history taking, appropriate clinical response to abnormal vital signs (e.g. taking a blood culture if body temperature exceeds 38°C) and other features of good care (e.g. prescribing steroids within 24 hours for asthma and COPD patients).

At baseline there was no difference between **SPI** and control hospitals. Observations of patients taken at 6 and 12 hours improved in both intervention and control hospitals over time in both evaluations. There was only one measure in which the $SPI_1$ intervention was significant (p=0.008) and that was the recording of respiratory rate at 12 hours. Across all measures there was no difference between intervention and control hospitals for both evaluations and overall no **SPI** effect was observed.

The observed prescribing error rates were quite high when compared with the only other study using the same methods in two hospitals (7.4 and 8.6%) (Barber et al., 2006). After adjustment for learning/fatigue effects the evaluation found no improvement in prescribing

over epochs, and there was no difference between control and intervention hospitals, thus 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 (Dean et al., 2005). Errors associated with anticoagulation therapy are *potentially* a particular cause of concern (Leape et al., 1991) and was therefore a specific **SPI** target. No trends towards fewer errors over time in anticoagulation therapy were found (and discovered only one adverse event associated with this class of drug, see Table 10.2). Medicines reconciliation was another key **SPI** target, but there was no trend towards improvements either over time, or between the **SPI** hospitals and controls.

A number of clinical processes that were *not* specific **SPI** targets showed that there were no significant differences between control and **SPI** 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 here there was room for improvement.

An important methodological finding of this sub-study is the learning/fatigue effect observed on behalf of the reviewer, particularly with semi-structured items such as prescribing which requires piecing together disparate pieces of information in the case-notes. If this had not be taken into account in the analysis prescribing errors would have increased overtime as

the reviewer became more adept at spotting this type of error and then decreased as the reviewer became weary of the process. Ideally all case-notes would be collected and randomised for review so that this type of bias is spread over time and between intervention and control arms but for pragmatic reasons, such as in this instance this is not always possible and hence this type of adjustment needs to be made. In hindsight, a panel of reviewers were required so that the review process was not channelled through an individual and to minimise bias of learning/fatigue further.

# CHAPTER 7: HOLISTIC CASE-NOTE REVIEW TO MEASURE ERROR RATES/QUALITY OF CARE IN PATIENTS WITH ACUTE RESPIRATORY DISEASE

## 7.1 Introduction

In the previous chapter, I described the review of acute respiratory episodes of care using an explicit review method. In this chapter, I will describe an alternative audit method utilising the same cases-notes – that of a semi-structured holistic (implicit) review. In a structured holistic review, experts are presented with a series of pre determined questions designed to extract a complete assessment. The differences between explicit and holistic reviews are described in Table 4.1. However, holistic reviews can be considered as complimentary rather than an alternative to explicit reviews, as it reflects the full scope of clinical decisions that apply to a particular patient.

## 7.2 Methods

The case-note selection and eligibility criteria have previously been described in Chapter 6. This sub-study was conducted in both the $\boldsymbol{SPI_1}$ and $\boldsymbol{SPI_2}$ evaluation (Figure 6.1). This review was undertaken by a specialist in general medicine, who has considerable experience in case-note review, and has investigated hospitals who were outliers on hospital mortality

statistics (Mohammed et al., 2008b). To measure inter-observer reliability, a subset of case-notes was independently re-evaluated by an experienced trainee in respiratory medicine.

Using expert clinical judgement, an overall quality score was assigned, graded on a scale from 1 (unsatisfactory, an error had occurred) to 10 (very best care). A specific score for each of the three stages of care – (a) admission; (b) management; and (c) pre-discharge – was allocated on a scale from 1 (unsatisfactory) to 6 (excellent care). Reviewers classified errors and adverse events using standard definitions found in Box 7.1 (Brennan et al., 1991; Leape et al., 1991; Reason, 2000; Vincent et al., 2001; Wilson et al., 1995; Woloshynowych et al., 2003).

**Box 7.1: 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"*

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 per 100 patients. Average ratings and average numbers of errors were calculated for both control and intervention groups, for

all study epochs (with standard errors). Errors were further classified by broad categories (Table 7.1).

**Table 7.1: Classification of errors and adverse events**

| Category | Nature of the problem |
|---|---|
| Diagnosis/Assessment Admission Error | • Failure to diagnose promptly /correctly<br>• Failure to assess patient's overall condition adequately (including comorbidities) |
| Hospital-acquired infection | • Hospital-acquired infection |
| Technical/management | • Technical problem relating to a procedure<br>• Problem in management/monitoring (including nursing and other professional care) |
| Medication/Maintenance/Test results | • Failure to give correct/monitor the effect of medication<br>• Failure to maintain correct hydration/electrolytes<br>• 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".*

### 7.2.1 Statistical methods

No formal adjustments were made for multiple comparisons, although 99% confidence intervals are quoted in all cases. Inter-observer reliability was calculated by ICC for the four scores. Inter-observer reliability was assessed between the two errors by using the $\kappa$ statistic.

#### *7.2.1.1 Statistical methods used specifically in the $SPI_1$ evaluation*

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).

#### *7.2.1.2 Statistical methods used specifically in the $SPI_2$ evaluation*

A mixed modelling approach was used to test for differences in between Epochs 1 and 2, and Epoch 3. Random effects were included to allow for within hospital correlation, using an exchangeable correlation structure. Covariates included a binary variable "After" indicating whether the observation was before or after the intervention period; a binary variable "Intervention" indicating whether the hospital was a control or $SPI_2$ hospital; a binary variable "Epoch 1 (or 2)" indicating whether the observation was from the pre-intervention phase; and an interaction between "After" and "Intervention", to evaluate the estimated difference in change between the control and $SPI_2$ hospitals (between Epoch 3 and the average of the pre-intervention epochs). All models were adjusted for age and sex of patients.

## 7.3 Results

### 7.3.1 The sample

#### *7.3.1.1 Number of case-notes reviewed within the holistic review in the $SPI_1$ evaluation*

The number of case-notes reviewed by the holistic method differs, being higher than the number reviewed by the explicit review method (see 6.3.1). This is because the reviews were commissioned at different times. In the four $SPI_1$ hospitals, 390 and 381 case notes were holistically reviewed from Epoch 1 and Epoch 2 respectively (roughly divided equally between the four hospitals). For the eighteen control hospitals, 243 and 246 case notes were reviewed from Epoch 1 and Epoch 2 (range 8 to 15 cases per hospital).

#### *7.3.1.2 Number of case-notes reviewed within the holistic review in the $SPI_2$ evaluation*

In the nine $SPI_2$ hospitals, 359 case-notes were holistically reviewed (roughly divided equally between the nine hospitals). For the nine control hospitals, 366 cases notes were holistically reviewed (again roughly equally divided between the 9 hospitals). For both the control and $SPI_2$ hospitals, roughly equal numbers of cases notes were reviewed from each of the three epochs (243 cases notes were reviewed from Epoch 1; 246 from Epoch 2; and 236 from Epoch 3). This means that a total of 489 cases notes were reviewed from the pre-intervention period and 236 cases notes were reviewed from the post-intervention period. A

small number of case-notes analysed by explicit review did not get included in the holistic review, and vice-versa, due to logistical problems and time constraints. For this reason the homology between the two sets of notes is not complete. For example there were 31 deaths among the explicit case-notes reviewed, and 30 among the implicit case-notes.

### 7.3.2 Reliability

#### *7.3.2.1 Agreement between reviewers within the $SPI_1$ evaluation*

In total, 122 case notes were reviewed by both reviewers. Measures of reliability between the two holistic reviewers were, as expected for holistic reviews, low (Lilford et al., 2007) (ICC's 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). 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 was low ($\kappa$= 0.15, se 0.08).

#### *7.3.2.2 Agreement between reviewers within the $SPI_2$ evaluation*

Seventy-four case-notes were reviewed by two reviewers. Again, measures of reliability between the two holistic reviewers were low (Lilford et al., 2007) (ICC's were 0.05 (99% CI: -0.25, 0.34) for admission rating; 0.05 (99% CI: -0.25,0.34) for the management rating; 0.37

(99% CI: 0.08,0.60) for the pre-discharge care rating; and 0.31 (99% CI: 0.02, 0.56) for the overall care rating). Again 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 errors identified by the two reviewers was small ($\kappa$ = 0.08, se 0.09).

### 7.3.3 Quality of care ratings

#### *7.3.3.1 Quality of care within the $SPI_1$ evaluation*

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 1 (below best practise) to 6 (excellent care); and the average score for over all care was 7.4 (0.06), on a scale of 1 (unsatisfactory) to 10 (very best care). Admission, management and pre-discharge care ratings were higher in the $SPI_1$ hospitals compared with the control hospitals, during both Epoch 1 and Epoch 2 (Table 7.2), 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 admission rating increased significantly between epochs (increase 0.28, p=0.001) However, differences *in changes* across control and $SPI_1$ hospitals were not significant for any of the four ratings (Table 7.2).

#### *7.3.3.2 Quality of care within the $SPI_2$ evaluation*

The average quality of care scores during Epoch 1 (with standard errors) for admission, management and pre-discharge ratings were 4.89 (0.08), 4.15 (0.12), and 4.20 (0.12) respectively; and the average score for overall care was 7.56 (0.09). During Epoch 1, all of the four quality of care ratings were higher in the $SPI_2$ hospitals compared with the control hospitals (Table 7.3), although not significantly so. However, during both Epoch 2 and Epoch 3, all four quality of care ratings were higher in the control hospitals compared to the $SPI_2$ hospitals (although again not significantly so). In the control hospitals, all ratings tended to increase with time; whereas in the $SPI_2$ hospitals, all ratings decreased between Epoch 1 and Epoch 3, (although once again not significantly so). However, differences *in changes* across control and $SPI_2$ hospitals were not significant for any of the four ratings (Table 7.3).

An analysis of the differences between $SPI_2$ hospitals, and control hospitals at baseline and changes between Epoch 3 was also conducted, the details of which can be found in Appendix C.

### 7.3.4 Errors

#### *7.3.4.1 Error rates within the $SPI_1$ evaluation*

The numbers of errors per 100 patients were lower in the $SPI_1$ hospitals compared to the control hospitals, for both for Epoch 1 and Epoch 2 (Table 7.2). In the control hospitals, there was approximately one error for every two patients, whereas in the $SPI_1$ hospitals there was

circa one error for every three patients. The numbers of errors decreased in Epoch 2 (for both the control and $SPI_1$ hospitals), although this difference was not significant. Again differences in changes across control and $SPI_1$ hospitals were not significant for errors.

A total of 425 errors were identified (Table 7.4). 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 $SPI_1$ hospitals in Epoch 2. Rates of other errors also differed between control and $SPI_1$ hospitals and between Epoch 1 and Epoch 2, although no differences in changes were significant.

#### *7.3.4.2 Error rates with the $SPI_2$ evaluation*

Over all hospitals and all epochs, the average number of errors observed was 41 (se 2.17) per 100 patients, which equates to approximately one error in every 2.5 case-notes reviewed. In the control hospitals, the average number of errors per 100 patients decreased over the three epochs from 52.4 (se 5.6) errors per 100 patients in the first epoch to 30.7 (se 5.3) in the third epoch (Table 7.5). Whereas, in the $SPI_2$ hospitals, the average number of errors per 100 patients was relatively stable over epochs: from 35.9 (se 4.9) in the first epoch to 38.5 (se 5.0) in the third. Again differences in changes in average number of errors before and after the intervention across control and $SPI_2$ hospitals were not significant (Rate Ratio 1.47; 0.74-0.90).

A total of 153 errors were identified in the control hospitals and 145 errors identified in the ***SPI₂*** hospitals (Table 7.5). The most frequent categories of errors related to *"diagnosis, assessment or admission"*, or were errors relating to *"poor clinical reasoning"*. Errors relating to both these types were more frequent in the control hospitals in Epoch 1, but were less frequent during Epochs 2 and 3. Rates of other errors also differed between control and ***SPI₂*** hospitals and between Epoch 1 and Epoch 2, although no differences in changes were significant.

**Table 7.2: Holistic review: changes in ratings and numbers of adverse events and errors between control and $SPI_1$ hospitals**

| | Control Hospitals | | | | $SPI_1$ Hospitals | | | | Difference in change |
|---|---|---|---|---|---|---|---|---|---|
| | Epoch 1 | | Epoch 2 | | Epoch 1 | | Epoch 2 | | (99% CIs)* |
| | | | | | | | | | |
| Number of Patients | 243 | | 246 | | 390 | | 381 | | |
| | | | | | | | | | |
| | Ave | SE | Ave | SE | Ave | SE | Ave | SE | |
| ***Quality Ratings*** | | | | | | | | | |
| Admission Rating† | 4.9 | 0.08 | 4.9 | 0.08 | 5.0 | 0.07 | 5.3 | 0.05 | 0.23 (-0.14, 0.60) |
| Management Rating† | 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† | 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‡ | 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 | |
| Number Errors per 100 patients$^{\Phi}$ | 44.44 | 3.8 | 42.3 | 3.8 | 29.7 | 2.5 | 24.4 | 2.3 | -2.42 (-17.99, 13.31) |
| Number Adverse Events per 100 patients$^{\Phi}$ | 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) |

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

†Score scale: 1 (below best practice) to 6 (excellent care)

‡Score scale: 1 (unsatisfactory) to 10 (very best care)

$^{\Phi}$Patients could experience more than one error and more than one adverse event

Shaded areas relate to post-intervention epochs

**Table 7.3: Holistic review: changes in ratings and numbers of adverse events and errors between control and $SPI_2$ hospitals (standard errors in parenthesis)**

| | Control Hospitals | | | $SPI_2$ Hospitals | | | Difference in Change (99% CIs)* |
|---|---|---|---|---|---|---|---|
| | **Epoch 1** | **Epoch 2** | **Epoch 3** | **Epoch 1** | **Epoch 2** | **Epoch 3** | |
| Number of patients | 126 | 126 | 114 | 117 | 120 | 122 | |
| **Quality Ratings** | | | | | | | |
| Admission Rating† | 4.76 (0.13) | 4.94 (0.12) | 4.97 (0.10) | 5.03 (0.10) | 4.93 (0.11) | 4.87 (0.10) | -0.26 (-0.77, 0.24) |
| Management Rating† | 3.98 (0.17) | 4.18 (0.17) | 4.29 (0.16) | 4.35 (0.16) | 4.03 (0.17) | 4.25 (0.16) | -0.18 (-0.92, 0.56) |
| Pre-discharge Rating† | 4.13 (0.16) | 4.25 (0.14) | 4.32 (0.13) | 4.28 (0.15) | 4.16 (0.15) | 4.25 (0.14) | -0.10 (-0.74, 0.54) |
| Overall Care Rating‡ | 7.42 (0.13) | 7.62 (0.12) | 7.77 (0.11) | 7.72 (0.11) | 7.46 (0.12) | 7.47 (0.11) | -0.41 (-0.94, 0.11) |
| **Errors / Adverse Events** | | | | | | | **Rate Ratios** |
| Number Errors Φ | 52.4 (5.6) | 39.7 (5.2) | 30.7 (5.3) | 35.9 (4.9) | 45.0 (5.7) | 38.5 (5.0) | 1.47 (0.74, 2.90) |
| Number Adverse Events Φ | 4.76 (2.21) | 3.97 (1.74) | 3.51 (1.73) | 0.85 (0.85) | 5.00 (1.99) | 0 (--) | |

* Differences in changes are estimated from a mixed effects model (see methods for details) and represent a difference in change between Epoch 3 and Epochs 1 and 2.

† Score scale: 1 (below best practice) to 6 (excellent care)

‡ Score scale: 1 (unsatisfactory) to 10 (very best care)

Φ The numbers of errors and numbers of adverse events are per 100 patients (patients could experience more than one error and more than one adverse event)

Shaded areas relate to post-intervention epochs

**Table 7.4: Rates (per 100 patients) of errors identified by broad category of error – holistic review**

| | Control Hospitals | | | | $SPI_1$ Hospitals | | | | Effect of $SPI_1$ |
|---|---|---|---|---|---|---|---|---|---|
| | Epoch 1 | | Epoch 2 | | Epoch 1 | | Epoch 2 | | |
| | | | | | | | | | |
| Number of Patients | 243 | | 246 | | 390 | | 381 | | |
| | | | | | | | | | |
| ***Number of Errors**** | 111 | | 104 | | 116 | | 94 | | |
| | | | | | | | | | **Rate Ratio (99% CI)** |
| 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 |
| | | | | | | | | | |
| | **Rate** | **SE** | **Rate** | **SE** | **Rate** | **SE** | **Rate** | **SE** | |
| ***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 |

*Errors can be of multiple categories

Shaded areas relate to post-intervention epochs

**Table 7.5: Rates per 100 patients of errors identified by broad category of error (Standard errors are in parenthesis) $SPI_2$**

| | Control Hospitals | | | $SPI_2$ Hospitals | | | Rate Ratio |
|---|---|---|---|---|---|---|---|
| | **Epoch 1** | **Epoch 2** | **Epoch 3** | **Epoch 1** | **Epoch 2** | **Epoch 3** | **(99% CIs)*** |
| Number of patients | 126 | 126 | 114 | 117 | 120 | 122 | |
| | | | | | | | |
| Number of errors | 67 | 50 | 36 | 44 | 54 | 47 | |
| | | | | | | | |
| Diagnosis/ Assessment/ Admission Error | 63.49 (7.18) | 42.86 (6.00) | 36.84 (6.74) | 44.44 (6.91) | 55.00 (7.28) | 46.72 (6.60) | 1.34 (0.72, 2.51) |
| Hospital-acquired infection | 0 | 0 | 0.87 (0.87) | 0 | 0 | 0 | Not estimable |
| Technical/ management | 10.32 (2.72) | 9.52 (2.63) | 9.65 (2.78) | 4.27 (1.88) | 8.33 (2.53) | 5.74 (2.11) | 0.94 (0.21, 4.28) |
| Medication /Maintenance/ Follow-up | 24.60 (4.46) | 16.67 (3.52) | 8.77 (2.66) | 22.22 (4.22) | 16.67 (3.80) | 17.21 (3.43) | 2.13 (0.69, 6.53) |
| Clinical reasoning | 36.50 (4.30) | 27.78 (4.00) | 20.18 (3.78) | 24.79 (4.01) | 29.17 (4.17) | 27.87 (4.08) | 1.65 (0.72, 3.77) |
| Discharge information | 12.70 (2.98) | 14.29 (3.13) | 9.65 (2.78) | 11.11 (2.92) | 16.67 (3.42) | 13.93 (3.15) | 1.43 (0.44, 4.68) |

Errors can be of multiple categories

* Differences in changes are estimated from a mixed effects model (see methods for details) and represent a difference in change between Epoch 3 and Epochs 1 and 2.

Shaded areas relate to post-intervention epochs

## 7.4 Discussion

This sub-study sought to measure, by using a semi-structured holistic review, the quality of care and the error rates of the case-notes collected in Chapter 7. Overall quality of care ratings remained stable over time in intervention and control hospitals in both evaluations. Error rates were lower in ***SPI₁*** hospitals (one error for every three patients) compared to controls (one error for every two patients), and in both arms the number of errors decreased in Epoch 2. Whilst in ***SPI₂*** evaluation the intervention hospitals had a similar error rate as control hospitals (one error for every 2.5 patients) before the **SPI** was implemented. But after **SPI** implementation the error rate remained stable in the ***SPI₂*** hospitals, but decreased in control hospitals. In both evaluations the most frequent categories of error related to *"diagnosis, assessment or admission"* or error relating to *"poor clinical reasoning"*, possibly reflecting the cognitive difficulty of these tasks (Croskerry, 2009).

Moreover, for both the quality of care measures and error rates there was no difference between control and intervention hospitals for both the ***SPI₁*** and ***SPI₂*** evaluations. These results concur with the explicit review in that there was no added benefit of the **SPI**.

In previous studies, using this method of assessment (Lilford et al., 2007), reliability was low between the reviewers for both quality of care measures and error rates. This low reliability could be due to differences in clinical experience - the principal reviewer was a senior consultant general physician with a particular interest in biochemistry, whilst the second reviewer was a soon to be consultant respiratory physician. Elstad et al., (2010) have

recently concluded that "*physicians overtime gain complex social, behavioural and intuitive wisdom as well as the ability to compare the present day patient against similar past patients*".

It is possible that reliability of the reviewers in a holistic review could be improved through further training, but perhaps it should be accepted that low reliability is inherent in this method, as each reviewer through their own personal experience, heuristics and biases will formulate an "internal" proforma of good care – the type of variation that can provide richness to a review but alas is lacking here.

A weakness of this sub-study was the failure to record the date of review, as it plausible that semi-structured holistic review was also susceptible to learning/fatigue effect observed in the explicit review. In the explicit review, this effect only became obvious during the analysis stage at which point it was too late to include this important piece of information in this sub-study.

Another flaw of this sub-study, as with the explicit review, was channelling such a large review through an individual as a reviewer is prone to fatigue. If this study was to be repeated, I would ensure that there was a panel of reviewers with similar clinical experience.

Furthermore this type of review is expensive, and in light of these results there seems little added value in conducting a holistic review to illicit quality of care ratings and error rates over an implicit review.

# CHAPTER 8: ERROR RATES AND QUALITY OF PERI-OPERATIVE CARE

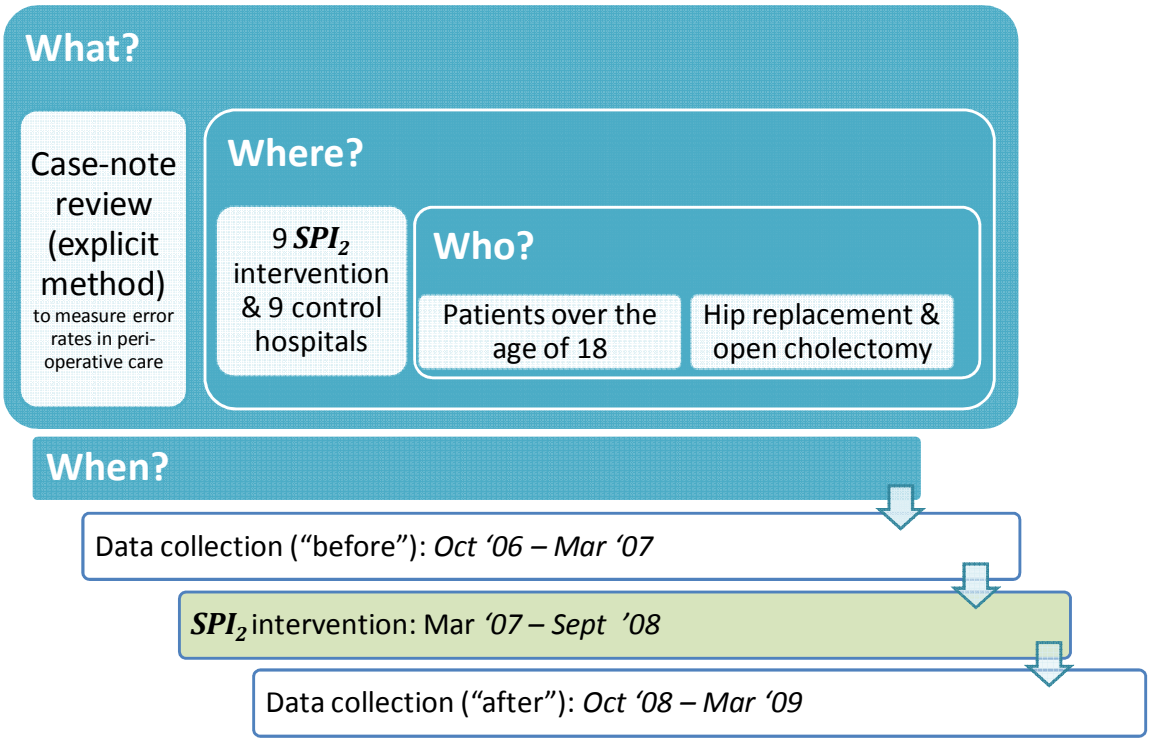
## 8.1 Introduction

A specific target of the **SPI** was to improve peri-operative care (see 3.6.2.5), and this sub-study seeks to measure adherence to pre defined standards of care by using explicit case-note review methodology.

## 8.2 Methods

This sub-study is summarised in Figure 8.1.

**Figure 8.1: An outline of the case-note review of surgical patients**



### 8.2.1 Case-note selection

Patients over the age of 18 years, undergoing major surgical operations of two types (total hip replacement and open colectomy see (Figure 8.1) were selected for the following reasons:

- improving peri-operative care was a specific $\boldsymbol{SPI_2}$ target;
- specific guidelines apply to this group of patients;
- it was believed that compliance with the guidelines was poor.

A set of explicit criteria for peri-operative care was developed using clinical guidelines from IHI (Institute for Healthcare Improvement, 2007) British Orthopaedic Association (British Orthopaedic Association, 2006) and the National Institute for Health and Clinical Excellence (NICE) (National Institute for Health and Clinical Excellence, 2008; National Institute for Health and Clinical Excellence, 2010) and by eliciting expert opinion. The areas of review are as follows:

- administration of prophylactic antibiotics prior to inclusion;
- the use of prophylactic Deep Vein Thrombosis (DVT) treatment (unless contraindicated), which included pharmacological intervention (unfractionated or low molecular weight heparins) and/or mechanical interventions, such as anti-thromboembolism stockings, foot pumps and sequential compression devices;
- intra-operative temperature monitoring (on at least one occasion);

- the use of “advanced methods” of pain control (epidural anaesthesia and/or patient controlled analgesia) for post-operative pain control. Types of anaesthesia administered were also investigated, as there is evidence to suggest that using neuraxial blocks (spinal and epidural) with sedation only or in combination with a general anaesthetic helps with early post-operative pain control and recovery. Likewise there is an evidence base to support the use of patient controlled analgesia (PCA). The quality criterion was that on least one of the modalities (neuraxial block or PCA) should be used.

Within the **SPI** intervention, the IHI advocate the removal of hair by clipping (not shaving); as this standard is not routinely recorded this was not included as a process measure for the evaluation.

### 8.2.2 Case-note assembly

Case-notes were selected from nine control and nine $\boldsymbol{SPI_2}$ hospitals. In this case there was a single pre-intervention epoch (corresponding to Epoch 2, i.e. October 2006 to March 2007) for comparison with the post-intervention epoch (corresponding to Epoch 3, i.e. October 2008 to March 2009). The intention was to analyse ten case-notes from each epoch (five of each surgical operation type) to yield a total sample of 360. To control for seasonal effects, the case-notes were spread across each time period (approximately two per month). The anonymisation procedures used in the sub-study dealing with the management of the acutely sick respiratory patients was followed (see 6.2.2).

All case-notes, were reviewed by a single medically trained reviewer over a period from November 2009 to January 2010. The first 20 cases were read jointly by another senior team member, and each one was discussed for training purposes. The notes were partially scrambled over epochs to assess, and if necessary control for, learning/fatigue effects. Inter-rater agreement was measured using 27 case-notes reviewed by a second reviewer, a surgical trainee.

### 8.2.3 Sample size calculation

Sample size calculation was performed after analysing results for 42 case-notes. There was high compliance (>90%) with the venous thrombo-prophylaxis and antibiotic criteria, such that there was little headroom for post intervention improvement. Thus the sample size calculation was based on intra-operative temperature monitoring, where compliance was about 40% at baseline, (i.e. there was plenty of headroom for improvement in response to **SPI**). Assuming that control hospitals experience an improvement from 40% to 50% compliance over the study period, the sample (n=360) was sufficient to detect an additional 25% to 30% improvement in association with **SPI** at 80% power, see Figure 8.2 and Table 8.1.

**Figure 8.2: Sample sizes for 80% power (at 5% significance)**

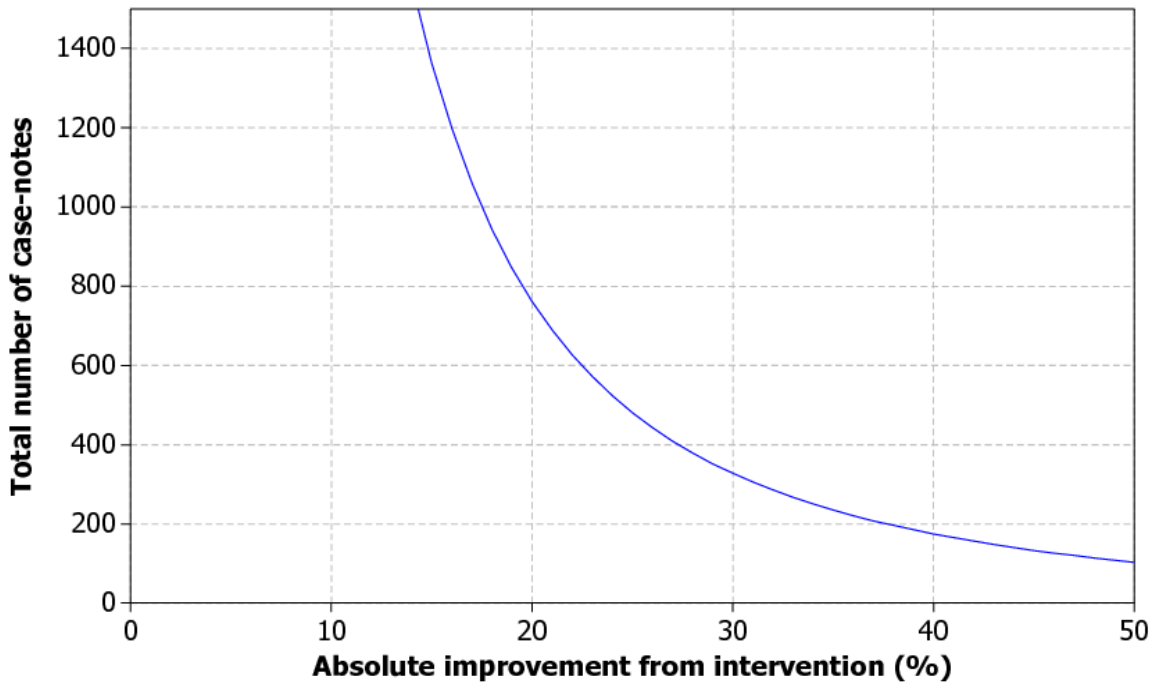



**Table 8.1: Sample sizes for 80% power (at 5% significance)**

| Effect Size (%) | Total number of cases needed for 80% power |
|---|---|
| 15 | 1,364 |
| 20 | 764 |
| 22.5 | 600 |
| 25 | 484 |
| 30 | 328 |
| 35 | 236 |

## 8.3 Results

### 8.3.1 Sample, reviewer reliability and headline message

There was a shortfall in the target number of 360 case-notes, and were only able to retrieve 242 notes; 127 were from admissions for total hip replacements and 115 from admissions for open colectomies. A second reviewer examined 27 case-notes. Percentage agreement and Kappa statistics are given in Table 8.2. These figures indicate low agreement on whether the temperature had been monitored (59%). For all other items the reviewers agreed on at least 85% of the cases.

No significant **SPI** effects were observed for any of the four clinical standards examined and the before/after comparison if anything, leaned towards the control hospitals. The hospitals were similar at baseline except with respect to intra-operative temperature monitoring where controls had more headroom for improvement. The results relating to the individual criteria are given in Table 8.3 and the outcomes of the mixed effects logistic regressions are given in Table 8.4.

### 8.3.2 Pain relief

Hospital staff identified contraindications to either epidural or self-administered analgesia in 15 of 242 cases. The existence of the contraindication was confirmed by the reviewers in all of these 15 cases with an additional contraindication in a patient identified by one of the reviewers. Two-hundred and twenty-six patients were thus eligible for modern analgesic

methods and 199 (88%) received such care. There was little headroom for improvement and there were no differences between control and $\boldsymbol{SPI_2}$ hospitals at either baseline or over time.

### 8.3.3 Prophylactic antibiotics

These were given in 235 of 242 cases (97%). While the breakdown across arms and epochs is summarised in Table 8.4, the full logistic regression analysis was not feasible because of the 100% compliance in the control hospitals at Epoch 2.

### 8.3.4 Temperature monitoring

There was marked but non-significant increase in compliance over epochs in both control and $\boldsymbol{SPI_2}$ hospitals with little difference in rate of improvement (OR 1.8; 0.4-7.6). There is evidence of heterogeneity between hospitals.

### 8.3.5 DVT prophylaxis

Anticoagulation prophylaxis was given in 239 of the 242 cases (99%). Two of these 239 were contra-indicated for prophylaxis. It was correctly withheld in one further contra-indicated case, and in two cases where no contra-indications were recorded.

**Table 8.2: Reviewer agreement in the peri-operative case-note**

| | Appropriate Pain Relief | Prophylactic Antibiotics | Temperature Monitored | DVT prophylaxis |
|---|---|---|---|---|
| % Agreement | 85% | 93% | 59% | 96% |
| Kappa* | 0.46 | - | 0.24 | - |

*Blank entries for Kappa indicate that one reviewer put all cases in the same category

**Table 8.3: Rates of compliance with peri-operative care standards**

| | Control Hospitals (N=9) | | | | *SPI*$_2$ Hospitals (N=9) | | | |
|---|---|---|---|---|---|---|---|---|
| | Pre-intervention | | Post-intervention | | Pre-intervention | | Post-intervention | |
| Number of Patients | 51 | | 43 | | 79 | | 69 | |
| | | | | | | | | |
| | % | SE | % | SE | % | SE | % | SE |
| Advanced method of pain relief* | 94.0 | 3.4 | 94.9 | 3.6 | 85.3 | 4.1 | 82.5 | 4.8 |
| Peri-operative antibiotic given[†] | 94.1 | 3.3 | 100.0 | - | 97.5 | 1.8 | 97.1 | 2.0 |
| Temperature Monitored[‡] | 16.0 | 5.2 | 30.2 | 7.1 | 29.1 | 5.1 | 41.2 | 6.0 |
| Appropriate DVT prophylaxis[†Φ] | 100.0 | - | 100.0 | - | 98.7 | 1.3 | 100.0 | - |

* Hospital staff identified 15 cases with contra-indications to this standard, all of which were corroborated by the reviewers. The data relates to the 227 eligible patients

† Logistic regression impossible because 100% in one cell

‡ Evidence of heterogeneity between hospitals at baseline

Φ Three cases had contra-indications yielding a denominator of 238. It was withheld in only two cases where no contra-indications were present but wrongly administered in two cases where there was a contra-indication

Shaded areas relate to post-intervention epochs

**Table 8.4: Peri-operative review: changes in the level of compliance between *SPI2* and control hospitals and the effect of $SPI_2$**

| | Baseline Comparisons | | Changes in Controls | | Effect of $SPI_2$ | |
|---|---|---|---|---|---|---|
| | *(SPI/Control)* | | *(Epoch 2/Epoch 1)* | | | |
| | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** | **OR (99% CI)** | **p-value** |
| Advanced method of pain relief | 0.3 (0.03, 2.6) | 0.148 | 1.0 (0.1, 17.2) | 0.978 | 0.6 (0.03, 18.4) | 0.820 |
| Peri-operative antibiotic given | 0.8 (0.06, 11.5) | 0.862 | - | - | - | - |
| Temperature monitored* | 1.8 (0.5, 6.5) | 0.227 | 1.8 (0.4, 7.6) | 0.279 | 0.9 (0.1, 5.2) | |

*Temperature monitoring is subject to significant (p=0.01) variation between hospitals within the arms of the study.

## 8.4 Discussion

The retrospective explicit case-note review of peri-operative care demonstrated considerable compliance to the measures of interest with very little or no headroom for improvement. Eighty-eight percent of patients eligible for analgesics received them; 97% of patient received appropriate prophylactic antibiotics and 99% of patients received DVT prophylaxis.

There was good agreement on all items between reviewers aside from intra-operative temperature monitoring. Hospitals were similar at baseline apart from temperature monitoring. There was no observed additive benefit of the **SPI**.

When these results have been shared with others there is a disbelief that these results are correct particularly the levels of compliance for DVT prophylaxis. However, it is supported by the literature – in 2005 two surgical mortality audits (covering the period 1994 to 2002 and 2002 to 2004) reported the increasing use of DVT prophylaxis and a decreasing number of DVT adverse events (Semmens et al., 2005; Thompson et al., 2005). This result is further strengthened by a retrospective case-note review to measure the prevalence of DVT prophylaxis, which measured 100% compliance within orthopaedics in eight English hospitals (Campbell et al., 2001).

Another reason why the reader maybe sceptical of the results is that if compliance for prophylactic antibiotics is so high why has there not been an equally dramatic decrease in

surgical site infection a rates (Health Protection Agency, 2009). This could be due to the fact that the peri-operative bundle alone is not sufficient to reduce surgical site infections. There is evidence by Stulberg and colleagues (Stulberg et al., 2010) that compliance to surgical infection control processes such as the peri-operative care bundle are not associated with a significantly lower probability of infection.

# CHAPTER 9: INDIRECT MEASURE OF HAND HYGIENE

## 9.1 Introduction

Improvement in hand hygiene was a specific aim the **SPI** intervention. Within the UK there has also been a national initiative to improve hand hygiene amongst acute hospital employees - the "*clean**your**hands*" campaign (National Patient Safety Agency, 2010). This initiative consisted of actions to make alcohol hand rub (AHR) available at the bedside, monthly updated posters on wards and a patient empowerment component to encourage patients to ask staff to clean their hands. The campaign was rolled out in England and Wales between December 2004 to June 2005 and continues to date. Since hand hygiene was also a **SPI** target, the hypothesis that **SPI** would have an additive effect was tested.

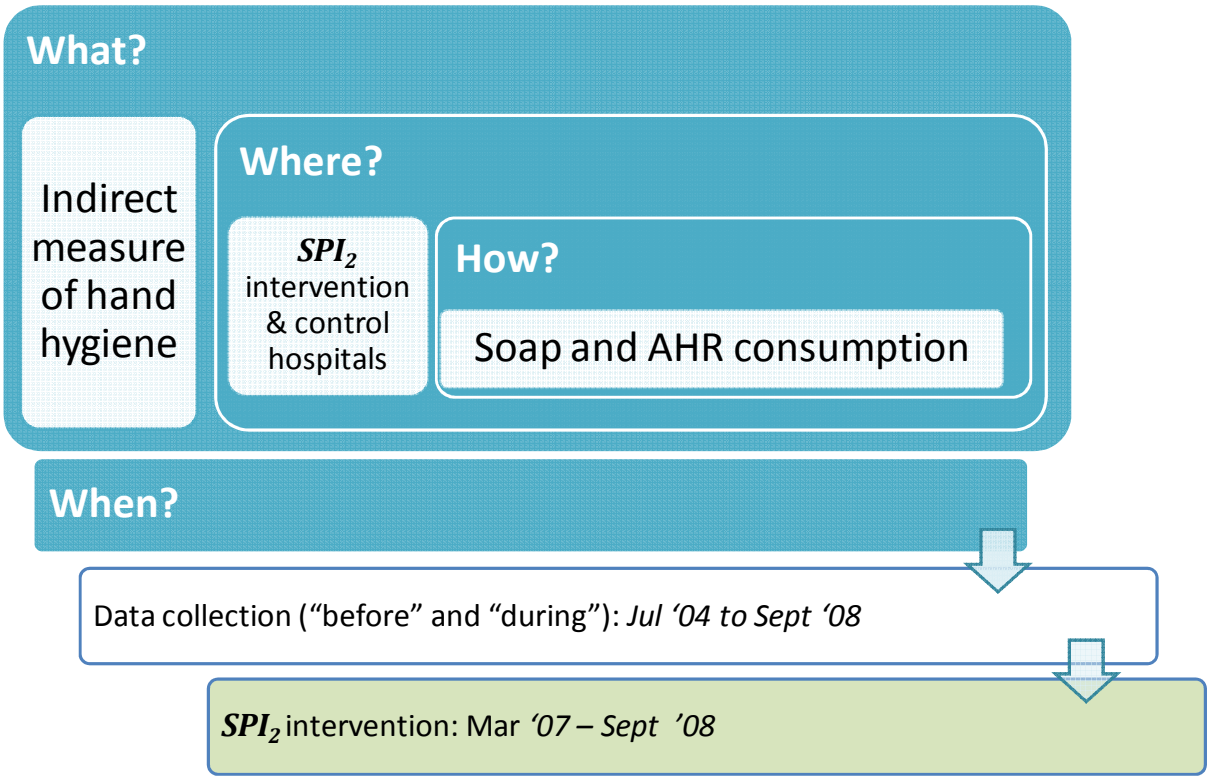
## 9.2 Methods

This sub-study was conducted in the SPI2 evaluation only and is outlined in Figure 9.1.

The success of this campaign was measured by the National Observational Study to Evaluate the "*clean**your**hands*" campaign (NOSEC) (Stone et al., 2007). As part of their study, monthly data from NHS Logistics for soap and AHR consumption (litres) was collected as an indirect measure of hand hygiene compliance. Data were available on a monthly basis for the period July 2004 to September 2008. This spanned a "before" period (July 2004 to February 2007) and a period concurrent with the intervention (March 2007 to September 2008). To adjust

for potential variations in consumption due to hospital size, these data, which were available at hospital level and were expressed as a rate (in litres) per 1,000 bed occupied days. Bed occupancy days were based on yearly averages spanning financial years (Department of Health, 2010).

**Figure 9.1: An outline of the indirect measure of hand hygiene sub-study**




### 9.2.1 Statistical methods

Population averaged (marginal) models were used to assess the effects of the intervention on soap and AHR consumption. To allow for decays in correlations (within hospitals) over time, an auto-regressive (AR 3) correlation structure was included. Results are based on the identity scale as this allows estimation of difference in change. Covariates within the models

included an indicator variable denoting intervention or control hospital and time as a continuous variable (from one to maximum number of temporal observations available). The effect of time was modelled as a polynomial function (cubic) as there was an indication that change in rates were non-linear. Finally, a fixed effect interaction between time and intervention allowed assessment of whether the change in rates of infection differed between control and $\boldsymbol{SPI_2}$ hospitals.

Both models were fitted in STATA using the GEE population averaged class of models. For the before and after comparisons, estimates of differences in differences (as estimated by the GEE models) are presented along with 99% confidence intervals. For the temporal models, smoothed estimates of outcomes over the study period are presented in graphical format, along with p-values for tests of significant differences in changes between control and $\boldsymbol{SPI_2}$ hospitals. Models were weighted with a suitably appropriate denominator (either number of events or standard deviation of outcome for summary data).

## 9.3 Results

### 9.3.1 Data available

Data on soap and AHR (in litres) were available for nine and eight of the control hospitals and for seven and six of the $\boldsymbol{SPI_2}$ hospitals respectively.

### 9.3.2 Soap and alcohol hand rub consumption

The median rate of soap consumption over all hospitals and all time periods was 50 litres per 1,000 bed days (IQR: 32, 71) and the median rate of AHR consumption was 44 litres per 1,000 bed days (IQR: 29, 61). Averaging over all time periods (July 2004 to September 2008) the median rate of soap and AHR consumption was higher in the $\boldsymbol{SPI_2}$ hospitals compared to the control hospitals: the median rate of soap consumption in the $\boldsymbol{SPI_2}$ hospitals was 53 litres (IQR: 30, 79) compared to 46 litres (IQR: 34, 65) in the control hospitals; and the median rate of AHR consumption was 49 litres (IQR: 31, 79) compared to 43 in the control hospitals (IQR: 34 ,65).

Rates of both soap and AHR consumption increased in both control and $\boldsymbol{SPI_2}$ hospitals over the study period (Table 9.1). For example, in the control hospitals the median rate of soap consumption increased from 43 litres (IQR: 32, 54) in the period before the intervention to 63 litres (IQR: 35, 86) in the period during the intervention; and in the $\boldsymbol{SPI_2}$ hospitals this rate similarly increased from 49 litres (IQR: 30, 64) to 71 litres (IQR: 5, 102). Smoothed estimates of rates of increase of consumption of both products, as estimated by the GEE population averaged model, are presented in Figure 9.2 and Figure 9.3.

The rate of increase in rates of consumption of both soap and AHR (i.e. the difference of the differences) were similar between control and $\boldsymbol{SPI_2}$ hospitals and were not significant (p=0.760 and p=0.889 respectively, Appendix D, Table A1), reflecting the fact that rates of consumption of both products was higher in the $\boldsymbol{SPI_2}$ hospitals through-out the study, and not only after the intervention phase.

**Table 9.1: Soap and AHR consumption - median and inter-quartile ranges for control and intervention hospitals, pre- and post-intervention period (total hospital consumption rates per 1,000 bed days)**

| | Control Hospitals | | $SPI_2$ Hospitals | |
|---|---|---|---|---|
| | *Pre-intervention (N=9) | Post-intervention (N=8) | Pre-intervention (N=7) | Post-intervention (N=6) |
| Soap | 43 (33,54) | 63 (35,86) | 49 (30,64) | 75 (5,102) |
| AHR | 34 (12,45) | 56 (45,67) | 39 (28,74) | 60 (42,96) |

*Before period is July 2004 to February 2007 and after (during) period is March 2007 to September 2008; units are litres per 1,000 bed days

Shaded areas relate to the post-intervention epoch

**Figure 9.2: Rate of soap consumption per 1,000 bed days over time in control and $SPI_2$ hospitals**

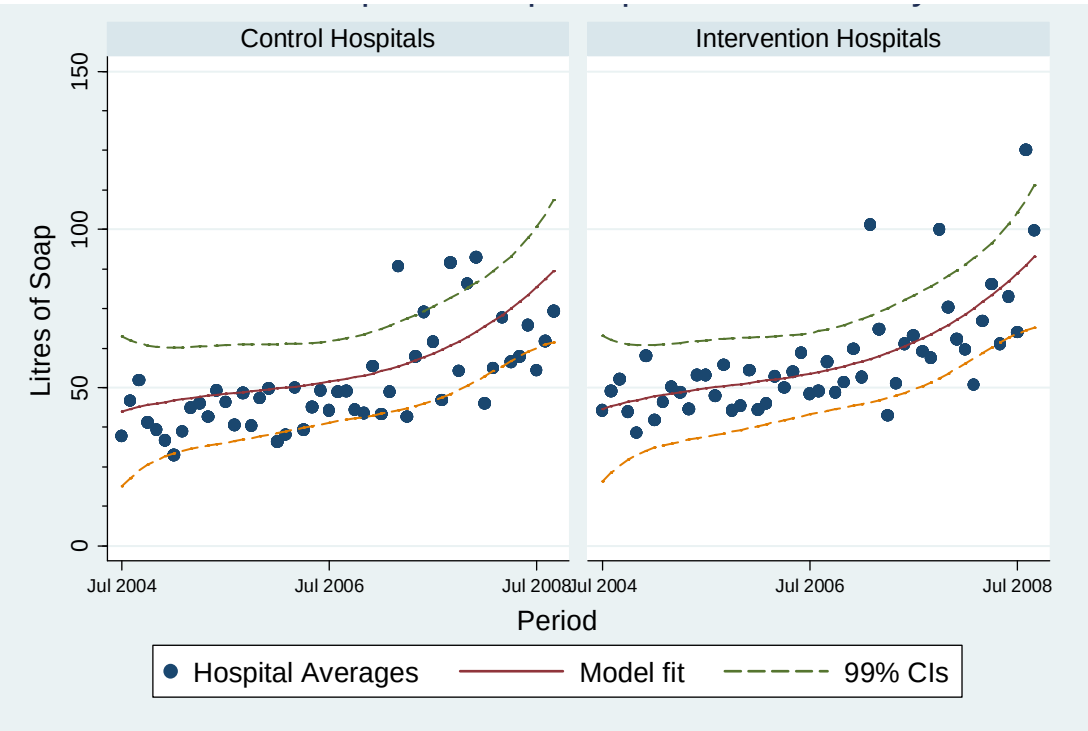
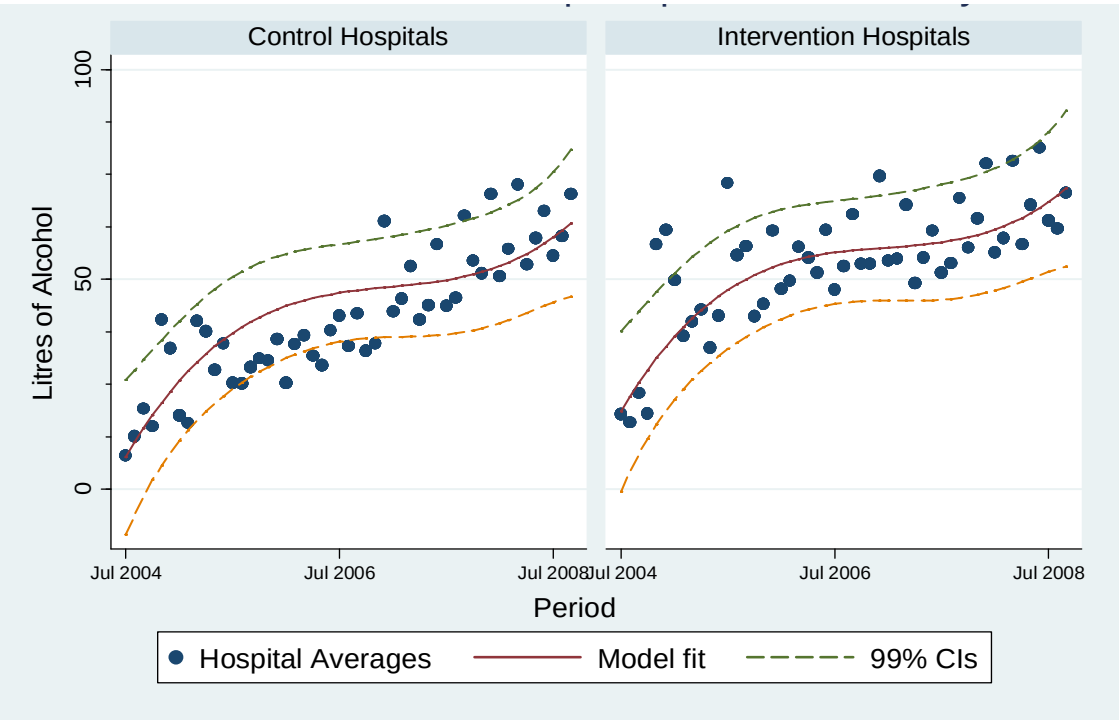


**Figure 9.3: Rate of alcohol hand rub (AHR) consumption per 1,000 bed days over time in control and *SPI2* hospitals**




## 9.4 Discussion

The use of hand hygiene consumables was higher in the ***SPI2*** hospitals compared to controls and consumption of these items has increased in all study hospitals overtime. However, there is no difference in the rate of change between the intervention and control hospitals and no **SPI** effect is observed.

If the premise is accepted that the measurement of soap and AHR can be used as an indirect measure of hand hygiene then perhaps this finding this should not come as a surprise. At the time of the **SPI** there was a large-scale resource intensive drive to improve hand hygiene - the "*clean**your**hands*" campaign (National Patient Safety Agency, 2010) and in addition the

Health Protection Agency (2010a), the Department of Health (2006) and WHO (2010) have all released guidance on this matter. The issue of hand hygiene has also entered the mainstream - posters to wash hands are commonly seen and AHR is now readily available to the general public.

# CHAPTER 10: ADVERSE EVENTS DETECTED IN ACUTE MEDICAL CASE-NOTES

## 10.1 Introduction

***SPI₁*** aimed to make a 50% reduction in the total number of adverse events (Health Foundation, 2006; Shirley, 2008), whilst, in ***SPI₂*** the aim was tempered to a 30% reduction. This sub-study aimed to measure the adverse event rate amongst the patients with acute respiratory disease on medical wards.

## 10.2 Methods

The incidence of patient harm caused by medication was measured as part of the explicit review (Chapter 7). Whilst the holistic review also measured adverse events both overall and by degree of preventability (Chapter 8); and results are presented in the chapter as total adverse events per 100 patients. This sub-study was conducted in both the ***SPI₁*** and **SPI₂** evaluations and utilises the case-notes collected in Chapter 7 (see Figure 6.1).

The results are presented as average numbers of adverse events per 100 patients. Average ratings and average numbers of adverse events were calculated for both control and intervention groups, for in all study epochs (with standard errors). Adverse events were further classified by broad categories (Table 7.1), and categorised into four levels of preventability. (Definitely preventable; preventable on balance of probabilities; not preventable on the balance of probabilities; and definitely not preventable).

### 10.2.1 Additional review of deaths within the $SPI_2$ evaluation

Within the $SPI_2$ evaluation, as a further quality control procedure, each death was re-analysed by a second reviewer ("blinded" to epoch and group) who had been trained in anaesthesia and in public health and who had experience as a reviewer of deaths for the National Enquiry into peri-operative deaths. This was conducted after the completion of the data collection.

### 10.2.2 Statistical methods

No formal adjustments were made for multiple comparisons, although 99% confidence intervals are quoted in all cases. Inter-observer reliability between reviewers was assessed by using the $\kappa$ statistic.

## 10.3 Results

### 10.3.1 Findings of the $SPI_1$ evaluation

The holistic review estimated an adverse event rate of about four per 100 patients treated (there were 56 adverse events overall), which is comparable to the published literature (Table 7.2) (Brennan et al., 1991; Leape et al., 1991; Vincent et al., 2001; Wilson et al., 1995). 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 ***SPI$_1$*** hospitals compared with the control hospitals in Epoch 1, but the reverse was the case for Epoch 2 (Table 7.2). The number of adverse events per 100 patients decreased during Epoch 2 in the ***SPI$_1$*** hospitals, while the number of adverse events increased during Epoch 2 in the control hospitals (Table 7.2). However, once again differences in changes were not significant. A trend in favour of the ***SPI$_1$*** hospitals was observed for five of the six categories of adverse events (Table 10.1): 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 (Table 10.2). At around 1.3% (16/1260), this is a somewhat lower rate of adverse events than reported for hospital inpatients elsewhere (Vincent et al., 2001). Four patients died where there was more than 50% probability that death resulted from an adverse event. 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).

**Table 10.1: Rates per 100 patients of adverse events identified by broad category of adverse events**

| | Control Hospitals (N=18) | | | | *SPI*$_1$ Hospitals (N=4) | | | | Difference in change |
|---|---|---|---|---|---|---|---|---|---|
| | Epoch 1 | | Epoch 2 | | Epoch 1 | | Epoch 2 | | (99% CIs) |
| | | | | | | | | | |
| Number of Patients | 243 | | 246 | | 390 | | 381 | | |
| | | | | | | | | | |
| Number 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
Shaded areas relate to post-intervention epochs

**Table 10.2: Preventable adverse events identified as being strongly* or certainly preventable out of the 1260 case notes reviewed in the *SPI1* holistic review**

| | Control Hospitals | *SPI1* Hospitals |
|---|---|---|
| **Epoch 1** | **1. Given oxygen and became unrousable from carbon dioxide ($CO_2$) retention requiring ICU 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. Shaded cases relate to patients who died

### 10.3.2 Findings of the $SPI_2$ evaluation

Over all hospitals and all epochs, the main reviewer identified 22 adverse events among the 725 case-notes: the average number of adverse events observed was 3.03 per 100 patients. The inter-rater agreement for identifying patients who had experienced an adverse event was low ($\kappa = 0$)

In the control hospitals the average number of adverse events per 100 patients decreased over the three epochs from 4.76 (se 2.21) adverse events per 100 patients in the first epoch, to 3.51 (se 1.73) in the third epoch. In contrast, in the $SPI_2$ hospitals, the average number of adverse events per 100 patients increased between the first and second epoch from 0.85 (se 0.85) to 5.00 (se 1.99); and decreased to zero in the third epoch. Again differences in changes in numbers of adverse events across control and $SPI_2$ hospitals were not significant (Rate Ratio=1.47; 0.74 - 2.90). Classifications by type of adverse event are presented in Table 10.3. Small numbers of identified adverse events preclude informative comparisons.

Three medication related adverse events were found on holistic review. At around 0.004% (3/725), this is also a somewhat lower rate than reported elsewhere (Vincent et al., 2001).

#### *10.3.2.1 Additional review of deaths within the $SPI_2$ evaluation*

The principal reviewer identified strong or certain evidence of preventability in four of the 22 adverse events (i.e. 0.5% of cases overall). These four adverse events occurred in the pre-intervention epochs described in Table 10.2

The first reviewer did not find over 50% probability of death in any of the 91 deaths included among holistic reviews. However, the second reviewer found two further deaths that were preventable on the balance of probabilities (both among control hospitals) in Epoch 3. One was due to brachycardia in a patient with hypokalaemica, and another due to delay in diagnosis of femoral artery thrombosis. This reviewer also found three deaths with strong evidence of probability in earlier epochs. A breakdown of deaths by level of preventability and reviewer is given in Table 10.4. Due to such small numbers of deaths being assessed as preventable, these percentages were not analysed between control and $SPI_2$ hospitals and they serve to shed light on mortality estimates. However, breakdown of deaths by level of preventability and reviewer is given in Table 10.4.

**Table 10.3: Rates (per 100 patients) of adverse events among patients admitted with acute respiratory disease**

| | Control Hospitals | | | *SPI*$_2$ Hospitals | | | Difference in Change |
|---|---|---|---|---|---|---|---|
| | **Epoch 1** | **Epoch 2** | **Epoch 3** | **Epoch 1** | **Epoch 2** | **Epoch 3** | **(99% CIs)*** |
| Number of patients | 126 | 126 | 114 | 117 | 120 | 122 | |
| | | | | | | | |
| ***Number of errors*** | 6 | 5 | 4 | 1 | 6 | 0 | |
| | | | | | | | |
| Diagnosis/ Assessment/ Admission Error | 3.97 (1.75) | 1.59 (1.12) | 2.63 (1.95) | 0.85 (0.85) | 3.33 (1.65) | 0 | -1.98 (-8.18, 4.23) |
| Hospital-acquired infection | 2.38 (1.36) | 1.59 (1.12) | 1.75 (1.24) | 0 | 2.50 (1.43) | 0 | -1.03 (-5.78, 3.74) |
| Technical/management | 0.79 (0.79) | 1.59 (1.12) | 0.88 (0.88) | 0 | 0.83 (0.83) | 0 | -0.10 (-3.48, 3.28) |
| Medication/Maintenance/ Follow-up | 0 | 0 | 0 | 0.85 (0.85) | 1.67 (1.17) | 0 | -1.27 (-3.88, 1.33) |
| Clinical reasoning | 0 | 0 | 0 | 0.85 (0.85) | 0 | 0 | -0.42 (-1.94, 1.09) |
| Discharge information | 0.79 (0.79) | 0 | 0 | 0.85 (0.85) | 0 | 0 | -0.02 (-2.17, 2.11) |

Errors can be of multiple categories

* Differences in changes are estimated from a mixed effects model (see methods for details) and represent a difference in change between Epoch 3 and Epochs 1 and 2.

Shaded areas relate to post-intervention epochs

**Table 10.4: Preventable deaths in acute medical wards across the study epochs**

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Number of deaths within holistic review** | **Preventable deaths: ≥50%** 1st reviewer** | **Preventable deaths: ≥50% 2nd reviewer** | **Preventable deaths: <50%* 1st reviewer** | **Preventable deaths: <50% 2nd reviewer** | **Number of deaths within holistic review** | **Preventable deaths: ≥50% 1st reviewer** | **Preventable deaths: ≥50% 2nd reviewer** | **Preventable deaths: <50% 1st reviewer** | **Preventable deaths: <50% 2nd reviewer** |
| **1** | 17 | 0 | 1† | 1 | 0 | 9 | 0 | 0 | 0 | 0 |
| **2** | 24 | 0 | 1† | 2 | 0 | 11 | 0 | 1† | 0 | 0 |
| **3** | 23 | 0 | 2 | 2‡ | 1‡ | 7 | 0 | 0 | 0 | 0 |

*Preventable deaths <50%: substandard practice was present that could have led to death but the probability that it did so was less than 50%;

**Preventable deaths ≥50%: substandard practice led to death on the balance of probabilities.

†These deaths (both associated with $CO_2$ retention in patients denied non-invasive ventilation – one of whom was given 60% oxygen) were not detected by the reviewer ***SPI1*** evaluation (see Table 10.2)

‡ The reviewers identified different cases, with no overlap.

Shaded area relates to post-intervention epochs

## 10.4 Discussion

For this sub-study agreement between the reviewers was again low, and possible reasons and methods for improvement have been discussed in Chapter 8.

Within the ***SPI$_1$*** evaluation there were a total number of 56 adverse events that equated to 4 adverse events per 100 patients. Whilst with the ***SPI$_2$*** evaluation there were 22 adverse events which is equivalent to 3.03 adverse events per 100 patients. These findings are comparable to the Harvard Medical Practice study (Brennan et al., 1991; Leape et al., 1991) but lower than a previous UK adverse event retrospective review, which reported an overall adverse event rate of 10.8% and adverse event rate within general medicine of 8.8% (Vincent et al., 2001). No difference in change was significant in either evaluation and hence no **SPI** effect was observed.

For the ***SPI$_1$*** evaluation 28% (16/56) of the adverse events were preventable and in the ***SPI$_2$*** evaluation this was 18% (4/22). In both evaluations the adverse event rate is much lower than previously reported in the UK (47% were preventable in total, whilst 75% were preventable in general medicine (Vincent et al., 2001). It is tempting to say that through improving quality of care observed in Chapter 7 preventable adverse events have decreased since the publication of earlier reports and during the period of both phases of the **SPI**. However, I am reticent in making such conclusions - the numbers of adverse events are small and the review was conducted for one episode of care (previous studies have reviewed the

whole patient case-note) such that it is unknown if further adverse events occurred post discharge and during subsequent readmissions.

A novel methodological approach conducted in this sub-study was the review of preventable mortality; although it does not provide any meaningful information on preventability it does provide an insight on how quality of care could be measured in future.

# CHAPTER 11: RATES OF MORTALITY AMONG ACUTE MEDICAL CARE PATIENTS

## 11.1 Introduction

Mortality rates across pre- and post-intervention epochs among patients whose case-notes were selected for review in Chapter 7 were compared. This was because it was feasible and because, arguably, a higher signal to noise ratio would be expected among this group, which not only was especially well placed to benefit from specific **SPI** interventions, but also tends to have high mortality.

## 11.2 Methods

The case-notes were collected as part of the sub-study described in Chapter 7 and this sub-study utilises the same statistical methods (see 6.2.6). This sub-study was undertaken in both the $SPI_1$ and $SPI_2$ evaluation.

## 11.3 Results

### 11.3.1 Mortality within the $SPI_1$ evaluation

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** is not significant. The baseline comparisons showed no significant differences between control and $\boldsymbol{SPI_1}$ hospitals; neither is there significant evidence of temporal change in the control hospitals (Table 11.1 and Table 11.2).

**Table 11.1: Mortality rates**

| | Control Hospitals (N=18) | | $SPI_1$ Hospitals (N=4) | |
|---|---|---|---|---|
| | **Epoch 1** | **Epoch 2** | **Epoch 1** | **Epoch 2** |
| | | | | |
| Number 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 |

Shaded areas relate to post-intervention epochs

**Table 11.2: Analysis of mortality rates**

| | Baseline Comparisons | | Changes in Controls | | Effect of $SPI_1$ | |
|---|---|---|---|---|---|---|
| | **Odds Ratio (99% CI)** | **p-value** | **Odds Ratio (99% CI)** | **p-value** | **Odds Ratio (99% CI)** | **p-value** |
| Mortality (adjusted for Age, Sex, Number 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 |

### 11.3.2 Mortality within the $SPI_2$ evaluation

Crude mortality was higher in the control hospitals than in the $SPI_2$ hospitals (OR 0.7; 0.2-2.1) (Table 11.3), but neither this, nor any other effect – including that of the **SPI** – was significant at the pre-determined 1% level after adjustment for age of patient (OR 0.3; 0.068-1.4) (although the result was just significant [p=0.043] at the 5% level). Sex and number of co-morbidities were also included as patient-level covariates, though only age was significant (p<0.001). The mortality rate increased by 10.3% (CI 6.8%-15.1%) per year of patient age.

**Table 11.3: Mortality among acute medical care patients whose case-notes were reviewed**

| | Control Hospitals (N=9) | | | $SPI_2$ Hospitals (N=9) | | |
|---|---|---|---|---|---|---|
| | **Epoch 1** | **Epoch 2** | **Epoch 3** | **Epoch 1** | **Epoch 2** | **Epoch 3** |
| | | | | | | |
| Number of Patients | 120 | 123 | 112 | 116 | 117 | 114 |
| Deaths | 18 | 24 | 24 | 9 | 15 | 7 |
| % Mortality (SE) | 15.0 (3.3) | 19.5 (3.6) | 21.4 (3.9) | 7.8 (2.5) | 12.8 (3.1) | 6.1 (2.3) |
| Age: Mean (SD) | 77.6 (7.7) | 81.1 (7.9) | 79.6 (8.0) | 77.7 (7.6) | 78.1 (7.1) | 80.6 (7.8) |
| % Female | 63.3 | 53.7 | 53.6 | 53.4 | 50.4 | 52.6 |
| Co-morbidities: Mean | 2.9 | 3.1 | 2.6 | 2.8 | 3.0 | 2.9 |

Shaded areas relate to post-intervention epochs

**Table 11.4: The effect of $SPI_2$ on the mortality among acute medical care patients**

| | Baseline Comparisons | | Changes in Controls | | Effect of $SPI_2$ | |
|---|---|---|---|---|---|---|
| | | | | | | |
| | **Odds Ratio (99% CI)** | **p-value** | **Odds Ratio (99% CI)** | **p-value** | **Odds Ratio (99% CI)** | **p-value** |
| Mortality (adjusted for Age, Sex, Number of co-morbidities) | 0.7 (0.2, 2.1) | 0.391 | 1.4 (0.6, 3.1) | 0.320 | 0.3 (0.08, 1.4) | 0.043 |

## 11.4 Discussion

Age, as can be expected is a significant predictor of death with the odds of dying increasing by 8% and 10% per year of patient age in the first and second evaluations respectively. Again there is no **SPI** effect for either the $SPI_1$ or the $SPI_2$ evaluations, however, it should be noted that if significance was set at the 5% level then the result for the $SPI_2$ evaluation would just be significant (p = 0.043). Too much emphasis should not be put on this as it is inevitable that a study carrying out multiple tests will throw up a significant value at this level (Bland and Altman, 1995), hence the reason for raising the threshold of significance to the 1% level. In addition this finding is not supported by other items measured in case-note review (Chapters 7, 8 and 11), which have all measured no **SPI** effect.

# CHAPTER 12: INTENSIVE CARE UNIT (ICU): MORTALITY, MORBIDITY AND LENGTH OF STAY

## 12.1 Introduction

Data from the Case Mix Programme (CMP) was accessed (Harrison et al., 2004) so that the effectiveness of the **SPI** critical care bundles (see 3.6.2.6) could be measured. The CMP is a comparative audit run by the Intensive Care National Audit and Research Centre (ICNARC). This programme collects patient outcomes from adult, general critical care units (intensive care and combined intensive care/high dependency units) covering England, Wales and Northern Ireland. Critical care units volunteer to join and collect standardised datasets (case mix, patient outcome and activity data) on patients admitted to their unit. These data are submitted to ICNARC for validation and analyses.

## 12.2 Methods

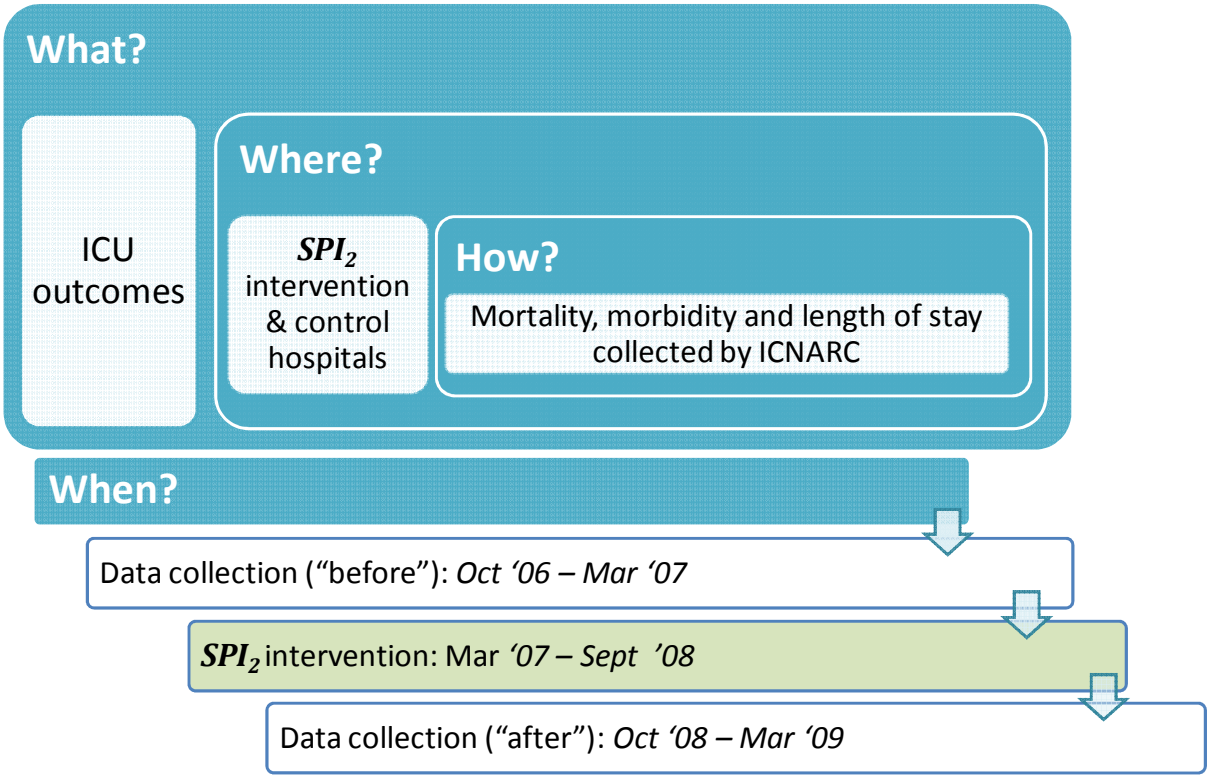
This sub-study was undertaken in the ***SPI$_2$*** evaluation only and is summarised in Figure 12.1.

To provide information relevant to the effectiveness of the critical care bundles data from the Case Mix Programme (CMP) (Harrison et al., 2004) was accessed. The CMP is a comparative audit run by the Intensive Care National Audit and Research Centre (ICNARC). This programme collects patient outcomes from adult, general critical care units (intensive care and combined intensive care/high dependency units) covering England, Wales and

Northern Ireland. Critical care units volunteer to join and collect standardised datasets (case mix, patient outcome and activity data) on patients admitted to their unit. These data are submitted to ICNARC for validation and analyses.

Data for the ICUs for all the study hospitals were available on a monthly basis for six months prior to the **SPI** (from October 2006 to March 2007) and for six months after the intervention (from October 2008 to March 2009). Mortality data were available on the observed numbers of deaths and the risk-adjusted number of deaths, both of which were used to calculate observed to expected mortality ratios. Information was also available on the mean length of stay (LOS) in the unit, along with standard deviation. Finally data were available on the mean risk prediction scores: the APACHE II score (Harrison et al., 2006) and the ICNARC score (Harrison et al., 2007), for patients admitted directly from a ward (along with standard deviation).

**Figure 12.1: An outline of the ICU outcomes sub-study**




### 12.2.1 Statistical methods

For data on intensive care outcomes a mixed modelling population averaged approach was again used to provide information relevant to the effects of the intervention. However, since these data were only available for a single six month period prior to the intervention and for a single six-month period after the intervention (continuous time series data throughout the study period were not available), these data were modelled using a simple difference of difference model (i.e. not including time as a continuous variable and not including an auto-regressive component). Covariates within the model included an indicator variable denoting control or $SPI_2$ hospital and an indicator variable, denoting before or after the intervention.

Correlations within hospitals were incorporated using an exchangeable correlation structure. Adjustment was made for the morbidity covariates, mean APACHE II score and mean ICNARC physiology score. Finally, a fixed effect interaction between intervention and before/after period allowed assessment of whether the change in outcomes between the before and after the intervention period differed between control and $\boldsymbol{SPI_2}$ hospitals.

All models were fitted in STATA using the GEE population averaged class of models. For the before and after comparisons, estimates of differences in differences (as estimated by the GEE models) are presented along with 99% confidence intervals. Full results from fitted GEE models are provided in Appendix D.

## 12.3 Results

### 12.3.1 Data available

Data on mortality, length of stay and several other outcome measures for intensive care units were available for 16 hospitals, eight of which were control hospitals and eight of which were $\boldsymbol{SPI_2}$ hospitals. Data were supplied to ICNARC by seven control and seven $\boldsymbol{SPI_2}$ hospitals for the pre-intervention period (Epoch 1) and for six control hospital and eight $\boldsymbol{SPI_2}$ hospitals post-intervention period (Epoch 2) (there were some hospitals which did not provide data for both periods).

### 12.3.2 Observed to expected mortality

The median observed to expected mortality ratio over all hospitals and all time periods was 1.06 (IQR: 0.93, 1.28). Averaging over all time periods (July 2004 to September 2008) this ratio was lower in the $SPI_2$ hospitals compared to the control hospitals: the median observed to expected mortality ratio in the $SPI_2$ hospitals was 0.98 (IQR: 0.90, 1.15) compared to 1.18 (IQR: 1.01, 1.32) in the control hospitals.

The rate of observed to expected mortality increased in the control hospitals over the study period (Table 12.1). For example, in the control hospitals before the intervention period, the median observed to expected mortality ratio was 1.14 (IQR: 0.99, 1.32) and this rate increased to 1.24 (IQR: 1.02, 1.33) in the six months after the intervention. Whereas, in the $SPI_2$ hospitals, the observed to expected mortality ratio decreased over the two periods: during the first six month period the observed to expected mortality ratio was 1.04 (IQR: 0.90, 1.15) and during the last six month period this decreased to 0.97 (IQR: 0.90, 1.15). At the end of the follow-up period (March 2008), the rate of observed to expected mortality was higher in the control hospitals. However, the adjusted difference-in-differences between control and $SPI_2$ hospitals after adjustment, was not significant at the 99% level (p=0.25, Table 12.1).

### 12.3.3 Median length of stay

The median length of stay was 125 hours (IQR: 96,153) over all hospitals and all time periods. Averaging over all time periods (July 2004 to September 2008) the median length of

stay was lower in the ***SPI2*** hospitals compared to the control hospitals: the median length of stay was 103 hours in the ***SPI2*** hospitals (IQR: 82,132) compared to 146 hours in the control hospitals (IQR: 123, 183). Based on this, control ICUs may have been dealing with a different case-mix from the ***SPI2*** ICUs.

Length of stay increased in the control hospitals over the study period (Table 3.16): during the pre-intervention period the median length of stay was 144 hours (IQR: 117, 174) and this increased to 147 hours (IQR: 126,185) in the post-intervention period. In the ***SPI2*** hospitals the median length of stay remained similar between the pre- and post-intervention periods: during the pre-intervention period the median length of stay was 102 (IQR: 82,130) and during the post-intervention period the median length of stay was 103 hours (IQR: 81, 137) in the six month period October 2007 to March 2008. Once again, differences in the rate of changes in length of stay were not significant (p=0.60, Table 12.1).

### 12.3.4 APACHE II and ICNARC risk prediction scores

Over all time periods and over all hospitals the median APACHE score was 20.0 (IQR: 17.8, 21.8) and the median ICNARC score was 22.1 (IQR: 19.5, 22.1). These scores were similar between control and ***SPI2*** hospitals and were similar between pre- and post-intervention periods (Table 12.1). Tests for differences in differences were not significant (p=0.45 and p=0.16, Table 12.1).

**Table 12.1: Intensive care outcomes– median and inter-quartile ranges for control and $SPI_2$ hospitals, pre- and post-intervention period**

| | Control Hospitals | | $SPI_2$ Hospitals | | Difference-in-difference | |
|---|---|---|---|---|---|---|
| ***Intensive and Critical Care Outcomes**** | **Pre-intervention (N=7)** | **Post-intervention (N=7)** | **Pre-intervention (N=6)** | **Post-intervention (N=8)** | **Change (99% CI)** | **p-value** |
| Adjusted Mortality Ratio | 1.14 (0.99,1.32) | 1.24 (1.02,1.33) | 1.04 (0.90,1.15) | 0.97 (0.90,1.15) | 0.09 (-0.11,0.29) | 0.25 |
| Mean LOS (hours) | 144 (117,174) | 147 (126,185) | 102 (82,130) | 103 (81,137) | 5.86 (-22.78,34.50) | 0.60 |
| Mean APACHE II Score | 20.4 (17.7, 22.6) | 19.0 (17.1, 20.8) | 21.1 (19.1, 23.0) | 20.3 (17.8, 21.8) | -0.83 (-3.63,1.98) | 0.45 |
| Mean ICNARC Score | 22.3 (19.5, 26.3) | 20.7 (18.0, 23.5) | 22.6 (21.2, 25.3) | 22.2 (19.7, 25.1) | -2.26 (-6.39,1.87) | 0.16 |

* Before period is October 2006 to March 2007 and after period is October 2008 to March 2009

LOS: Length of Stay

## 12.4 Discussion

Over the study period observed to expected mortality decreased in the ***SPI₂*** hospitals but increased in control hospitals. This may be explained by the differing LOS between control and intervention hospitals. In the control hospitals the median LOS was 146 hours compared to 103 hours in the ***SPI₂***, which indicates a different case-mix of patients. However, the differences for observed mortality and LOS between control and intervention sites were not significant and no **SPI** effect was observed. There was also no **SPI** effect in the APACHE II and ICNARC risk prediction scores.

This result should not be surprising as concurrent to the ***SPI₂*** intervention the Department of Health were recommending similar interventions to the critical bundles described in section 3.6.2.6 (Department of Health, 2007). It is also possible that discrete areas such as the ICU are readily able to adopt and maintain these types of interventions. In 2006 (Pronovost et al., 2006a) described a large and sustained reduction (66%) in rates of catheter-related bloodstream infections. This was maintained throughout an 18-month post intervention period. A recent follow up to this study has reported this drop has been conserved (Pronovost et al., 2010).

# CHAPTER 13: C. DIFF AND MRSA INFECTION RATES

## 13.1 Introduction

The Health Protection Agency (HPA) collects mandatory HCAI data from all acute hospitals in England and Wales and as several components of the **SPI** intervention (see 3.6.2.4) are related to infection control the numbers of C. diff and MSRA bacteraemia associated diarrhoea in the study hospitals was obtained.

## 13.2 Methods

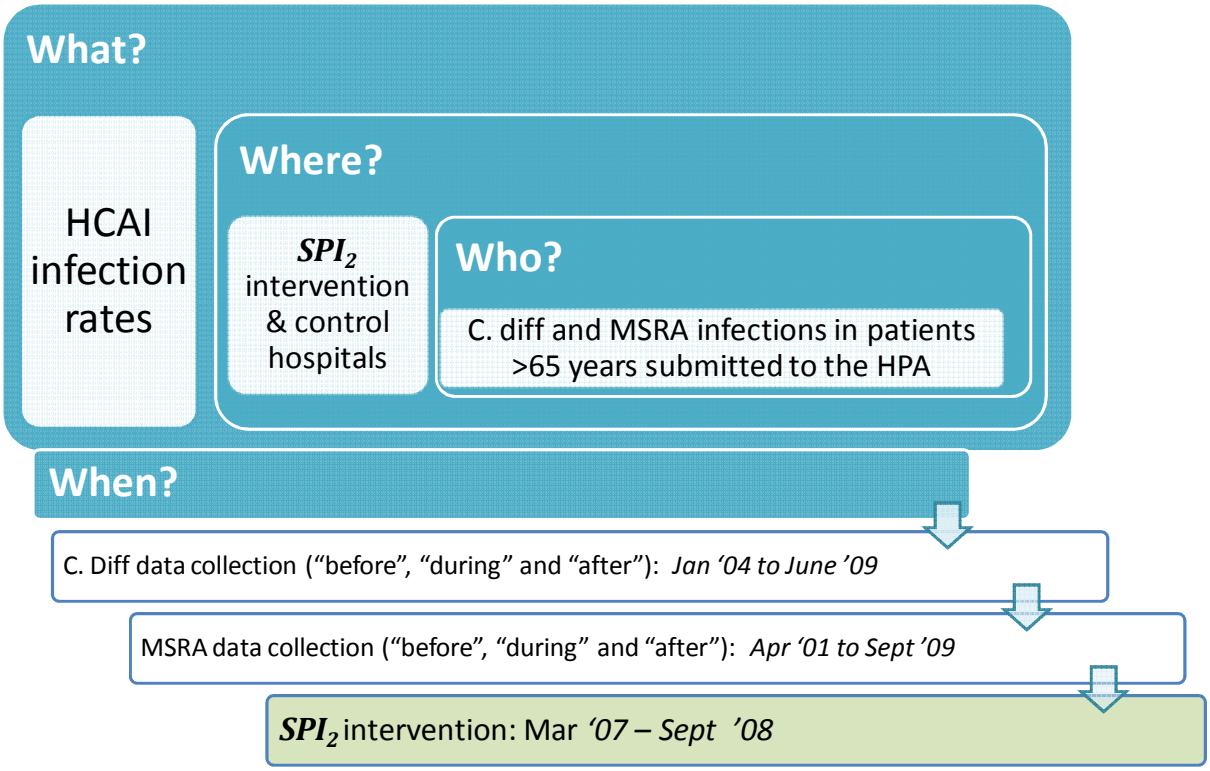
This sub-study was undertaken in the $\boldsymbol{SPI_2}$ evaluation only and is summarised in Figure 13.1.

Several components of the **SPI** intervention are related to infection control. All numbers of C. diff and MSRA bacteraemia associated diarrhoea in the study hospitals, were obtained from the Health Protection Agency (HPA), which collects mandatory HCAI data from all acute tr hospitals in England and Wales.

The C. diff and MRSA data relate to both community and hospital-based infections (i.e. they include cases diagnosed within the first 48 hours of stay) in patients older than 65 years. C. diff data were available quarterly for the period January 2004 to June 2009; and MRSA data were available from April 2001 to September 2009. These data therefore spanned a pre-intervention period (April 2001 or January 2004 to March 2007), a period concurrent with

the intervention (April 2007 to September 2008) and a post-intervention period (October 2008 to June 2009 or September 2009). To adjust for potential variations in numbers of cases due to hospital size, these data were expressed as a rate per 1,000 bed occupancy days for C. diff infections; and as a rate per 100,000 bed occupancy days for the MRSA infections. Bed occupancy days were based on yearly averages spanning financial years (Department of Health, 2010).

**Figure 13.1: An outline of the sub-study to measure HCAI rates**




### 13.2.1 Statistical methods

Population averaged (marginal) models were used to used to assess the effects of the intervention on rates of C. diff and MRSA infections. To allow for decays in correlations

(within hospitals) over time, an auto-regressive (AR 3) correlation structure was included. Model fits were compared between log and identity scales, and results presented here are based on the identity scale (as this allows estimation of difference in change). Covariates within the models included an indicator variable denoting control or $\boldsymbol{SPI_2}$ hospital and time as a continuous variable (from one to maximum number of temporal observations available). The effect of time was modelled as a polynomial function (cubic) as there was an indication that change in rates were non-linear. Finally, a fixed effect interaction between time and intervention allowed assessment of whether the change in rates of infection differed between control and $\boldsymbol{SPI_2}$ hospitals.

Both models were fitted in STATA using the GEE population averaged class of models. For the before and after comparisons, estimates of differences in differences (as estimated by the GEE models) are presented along with 99% confidence intervals. For the temporal models, smoothed estimates of outcomes over the study period are presented in graphical format, along with p-values for tests of significant differences in changes between control and SPI2 hospitals. Full results from fitted GEE models are provided in Appendix D.

## 13.3 Results

### 13.3.1 Data

Data on numbers of C. diff and MRSA cases were available for all 18 hospitals.

### 13.3.2 C. diff

Over all time periods, the median C. diff infection rate was 1.14 cases per 1,000 bed occupied days (IQR: 0.77, 1.64). Averaging over all time periods, the median rate of C. diff infection was similar between the control and $\boldsymbol{SPI_2}$ hospitals: the median C. diff infection rate was 1.15 (IQR: 0.88, 1.55) in the control hospitals and 1.10 (IQR: 0.67, 1.73) in the $\boldsymbol{SPI_2}$ hospitals.

The median C. diff infection rate decreased over the study period in both the control and $\boldsymbol{SPI_2}$ hospitals (Table 13.1). In the control hospitals, in the period before the intervention the median C. diff infection rate was 1.26 (IQR: 0.95, 1.67) and this decreased to 0.77 (IQR: 0.56, 1.02) in the period after the intervention. In the $\boldsymbol{SPI_2}$ hospitals, in the period before the intervention, the median C. diff infection rate was 1.37 (IQR: 0.65, 1.99) and this decreased to 0.66 (IQR: 0.50, 0.88) in the period after the intervention. Differences in changes were not significant between control and $\boldsymbol{SPI_2}$ hospitals (p=0.652, Appendix D, Table A4). Smoothed estimated rates of C. diff infection per 1,000 bed occupied days, by control and $\boldsymbol{SPI_2}$ hospitals, are presented in Figure 13.2.

### 13.3.3 MRSA

Over all time periods, the median MRSA infection rate was 14.75 cases per 100,000 bed occupancies (IQR: 8.93, 21.98). Averaging over all time periods the median rate of MRSA infection was similar between the control and intervention hospitals: the median MRSA

infection rate was 14.87 (IQR: 9.36, 21.63) in the control hospitals and 14.58 (IQR: 8.85, 22.77) in the ***SPI2*** hospitals.

The median MRSA infection rate decreased over the study period in both the control and ***SPI2*** hospitals (Table 13.1). In the control hospitals, in the period before the intervention the median MRSA infection rate was 17.40 (IQR: 12.01, 23.04) and this decreased to 4.31 (IQR: 2.26, 8.18) in the period after the intervention. In the ***SPI2*** hospitals, in the period before the intervention, the median MRSA infection rate was 17.76 (IQR: 11.60, 24.43) and this decreased to 6.77 (IQR: 4.89, 10.65) in the period after the intervention. Differences in changes were not significant between control and ***SPI2*** hospitals (p=0.693, Appendix D, Table A4). Estimated smoothed rates of MSRA infection per 100,000 bed occupied days, by control and ***SPI2*** hospitals, are presented in Figure 13.3.

**Table 13.1: Rates of C.diff (per 1,000 bed days ) and MRSA infections (per 100,000 bed days)**

| | Control hospitals (N=9) | | SPI2 hospitals (N=9) | |
|---|---|---|---|---|
| | **Pre-intervention** | **Post-intervention** | **Pre-intervention** | **Post-intervention** |
| C. diff[†] | 1.26 (0.95,1.67) | 0.77 (0.56,1.02) | 1.37 (0.65,1.99) | 0.66 (0.50,0.88) |
| MRSA[‡] | 17.41 (12.02,23.04) | 4.31 (2.26,8.18) | 17.76 (11.60,24.43) | 6.77 (4.89,10.65) |

[†] Before period is April 2004 to March 2007 and after period is October 2008 to June 2009
[‡] Before period is April 2001 to March 2007 and after period is October 2008 to September 2009
Inter-quartile range is shown in parenthesis

**Figure 13.2: Rate of Clostridium difficile (C. diff) cases per 1,000 bed days in control and $SPI_2$ hospitals**

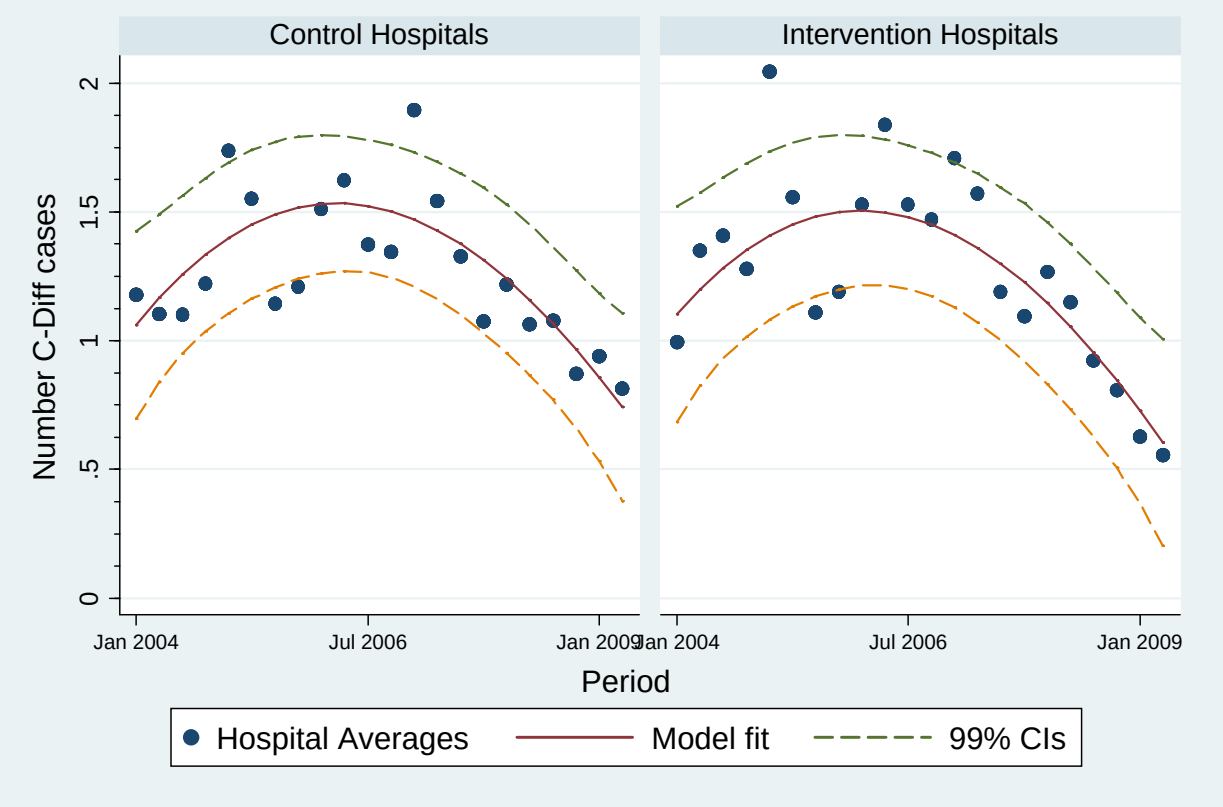
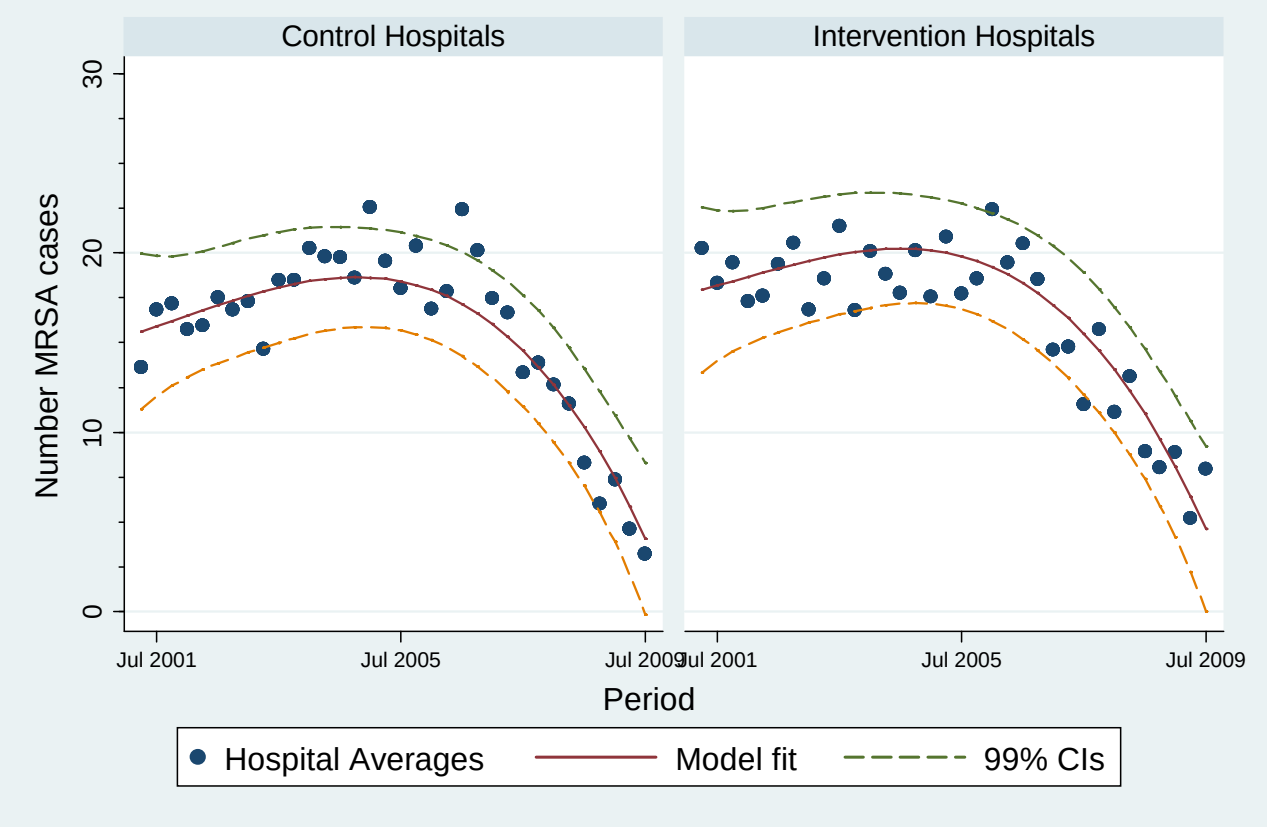


**Figure 13.3: Rate of Methicillin-resistant Staphylococcus aureus (MRSA) cases per 100,000 bed days in control and $SPI_2$ hospitals**



## 13.4 Discussion

Over the study period there is a dramatic fall in C. diff and MRSA infection rates in both control and ***SPI2*** hospitals. In control hospitals median C.diff rates fell from 1.26 to 0.77 per 1000 bed occupied days and MRSA fell from 17.40 to 4.31 per 100, 000 bed occupied days. Whilst in ***SPI2*** hospitals C.diff rates fell from 1.37 to 0.66 per 1000 bed occupied days and MRSA fell from 17.76 to 6.77 per 100, 000 bed occupied days. There was no significant difference between control and intervention hospitals.

The data presented here are cases diagnosed within the first 48 hours of stay, which includes both community and hospital HCAI (the HPA has recently started to separate community and hospital infections but these were not available for the time period of interest for the evaluation) and it is assumed that the fall in infection rates is attributed to the hospital. This assumption is taken as the hospital setting has been the target of the majority infection control measures and, thus, it would be highly improbable that the observed reduction was due to a fall in infection rates within the community setting.

Regardless, no **SPI** effect is observed and this could be due to prevailing interest in HCAI and there prevention. To this end the HPA was established in 2003 (before the **SPI**) and to date continues to report falling C.diff and MSRA rates in England and Wales (Health Protection Agency, 2010b).

# CHAPTER 14: PATIENT SURVEY

## 14.1 Introduction

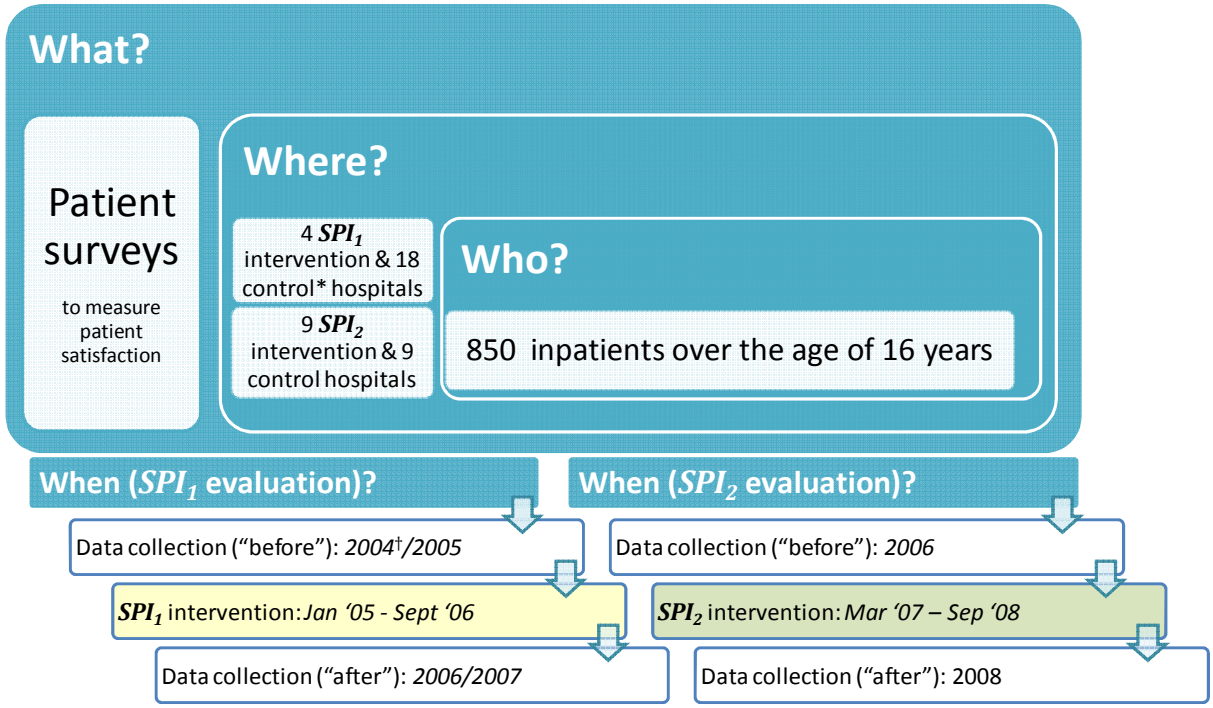
Since quality of care and avoidance of adverse events are important to patients, improvements in practice might plausibly affect patients' views of their care (Cleary and Edgman-Levitan, 1997). Their views were assessed by means of a patient survey.

## 14.2 Methods

This sub study was undertaken in both the $SPI_1$ and $SPI_2$ evaluation and is outlined in Figure 14.1 All English hospitals participate in an annual patient survey, and for purposes of the evaluation the same survey was administered 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 (Care Quality Commission and Picker Institute, 2010). Each hospital identified a random list of 850 eligible patients who had been consecutively discharged in the period June to August. Patients were eligible if they were 16 years or older (in the 2004 this was raised to 18 years as there was a concurrent survey of the experiences of young patient aged 0-17 years of age), had at least one overnight stay, and were not admitted to maternity or psychiatric wards.

Of the fifty core questions five were identified for analysis (Box 14.1): three overall satisfaction scores and two related to cleanliness. The details of these scores can be found in Appendix E.

**Figure 14.1: A summary of the patient survey sub-study**




*The controls for the ***SPI$_1$*** evaluation comprise of the pre-intervention data collected for the ***SPI$_2$*** evaluation. For a full explanation see section 3.2.
† In the English hospitals the sample consisted of patients over the age of 18, as during the same year an additional survey of the experiences of young patients (ages 0 to 17 years) was carried out.

### 14.2.1 Statistical methods used in the SPI$_1$ evaluation

The dates of the surveys were aligned with those of the staff surveys (see 5.2), 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 therefore conducted using a two-way ANOVA (the factors being ***SPI$_1$*** versus control hospital, and survey 1 versus survey 2). Organisation-level scores in the **SPI** arm were formed by averaging all respondents' scores within each hospital.

### 14.2.2 Statistical methods used in the $SPI_2$ evaluation

Data were collected in October to December 2006 (pre-intervention) and October to December 2008 (post-intervention). Methods similar to those for the staff survey were used in the analysis; except that the control variables included were sex, age, length of stay and whether the admission was emergency or elective (control variables were not available for the $SPI_1$ analysis above).

**Box 14.1: Patient survey questions deemed relevant to the SPI**

| Overall, how would you rate the care you received? |
|---|
| 1. How would you rate how well the doctors and nurses worked together? |
| 2. Overall, did you feel you were treated with respect and dignity while you were in the hospital? |
| 3. How would you rate how well the doctors and nurses worked together? |
| 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? |

## 14.3 Results

### 14.3.1 Findings of the patient survey in the $SPI_1$ evaluation

The response rate for the first survey was 54% (1961 of 3624 returned) in the four $SPI_1$ hospitals; for the second it was 53% (1720/3397). In the 18 control hospitals there was a greater drop, from 63% to 53%. Table 14.1 shows the values of the five survey scores in each of the four $SPI_1$ hospitals for the two surveys, along with details of response rates. Table 14.2 shows the changes in both control and $SPI_1$ 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 $\boldsymbol{SPI_1}$ hospitals on any of the scores. One of the survey scores showed a significantly different change between the control and $\boldsymbol{SPI_1}$ hospitals. The rating of cleanliness of toilets and bathrooms decreased in the control hospitals, from 79 to 77 points, whereas this increased in $\boldsymbol{SPI_1}$ 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 $\boldsymbol{SPI_1}$ hospitals were still slightly poorer than the control hospitals in survey 2 despite the change. None of the other four scores showed any significantly different changes between the two groups.

**Table 14.1: Patient survey scores in $SPI_1$ 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% |

Shaded areas relate to post-intervention epoch

**Table 14.2: Patient survey scores in control and $SPI_1$ hospitals at the two periods**

| | Control hospitals | | | $SPI_1$ hospitals | | | Range at baseline | Difference in change (99% CI) | p-value |
|---|---|---|---|---|---|---|---|---|---|
| | 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 |

Shaded areas relate to post-intervention epochs

### 14.3.2 Findings of the patient survey in the $\mathbf{SPI_2}$ evaluation

For the first survey, the overall response rate was 62% (4328 of 7010 valid questionnaires returned) in the nine ***SPI₂*** hospitals; for the second it was slightly lower at 55% (3762/6810). In the nine control hospitals, the response rates were 61% (4262/6791) and 57% (3973/6913) respectively. Table 14.3 shows the changes in both control and ***SPI₂*** hospitals on each of the five scores identified, along with the differences between the groups in these changes and associated 99% confidence intervals. All five scores improved over the study period in both the control and ***SPI₂*** hospitals. None of the five scores showed any significantly different changes between the two groups.

**Table 14.3: Patient survey scores in control and $SPI_2$ hospitals at the two periods**

| | Control hospitals | | | | | $SPI_2$ hospitals | | | | | Range at baseline | Difference in change (99% CI) | p-value |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | N | Survey 1 score (SE) | N | Survey 2 score (SE) | Absolute % change | N | Survey 1 score (SE) | N | Survey 2 score (SE) | Absolute % change | | | |
| Overall, how would you rate the care you received? | 4200 | 82 (0.4) | 3913 | 85 (0.3) | 4 | 4277 | 80 (0.4) | 3705 | 84 (0.3) | 4 | 75-87 | 1 (-1, 3) | 0.292 |
| Overall, did you feel you were treated with respect and dignity while you were in the hospital? | 4111 | 78 (0.4) | 3807 | 82 (0.4) | 4 | 4167 | 76 (0.4) | 3604 | 80 (0.4) | 3 | 65-85 | 0 (-2, 2) | 0.702 |
| How would you rate how well the doctors and nurses worked together? | 4182 | 87 (0.4) | 3878 | 88 (0.4) | 1 | 4220 | 88 (0.4) | 3677 | 89 (0.4) | 1 | 83-91 | 0 (-2, 2) | 0.597 |
| In your opinion, how clean was the hospital room or ward that you were in? | 4113 | 75 (0.4) | 3870 | 77 (0.4) | 2 | 4201 | 77 (0.4) | 3645 | 78 (0.4) | 1 | 70-80 | -1 (-3, 1) | 0.141 |
| How clean were the toilets and bathrooms that you used in hospital? | 4141 | 76 (0.4) | 3877 | 78 (0.4) | 2 | 4220 | 78 (0.4) | 3665 | 79 (0.4) | 1 | 70-82 | -1 (-3, 1) | 0.204 |

Shaded areas relate to post-intervention epoch

## 14.4 Discussion

Within the ***SPI1*** evaluation response rates declined in the control hospitals but remained stable in intervention hospitals. Whilst in the ***SPI2*** the response rates declined in both control and intervention hospitals.

There was one significant finding in the ***SPI1*** evaluation – that of cleanliness of toilets, which decreased in control hospitals but increased in ***SPI1*** hospitals. However, this finding should be treated with caution as the survey in 2004 which was implemented in English hospitals (1 ***SPI1*** hospital and all control hospitals) differed in its inclusion criteria than proceeding surveys [the survey implemented in 2005 (pre-intervention) in non-English ***SPI1*** hospitals and all hospitals post intervention) (see Figure 14.1). The remaining scores remained stable, and there was no significant difference between control and intervention hospitals.

In the ***SPI2*** evaluation no **SPI** effect was observed and patient satisfaction improved in all hospitals in the measures of interest. It is interesting to observe that two of the patient survey questions relate to cleanliness and over the study period there has been an increase in the use of hand hygiene consumables (see Chapter 10) and a decrease in HCAI rates (see Chapter 15).

Unfortunately no direct comparisons can be made with the NHS Inpatient Survey (Care Quality Commission and Picker Institute, 2010) (from which this data originates from) as different methods have been used for the analysis. The NHS Inpatient Survey reports on the

proportions of responses (Garratt, 2009) in each category found in Appendix G, whilst in this evaluation an average score has been calculated. However, the findings of the $\boldsymbol{SPI_2}$ evaluation reflect the NHS Inpatient survey which reports a significant rise in the proportion of patients reporting the highest category of satisfaction in the survey questions deemed relevant to the **SPI** (Garratt, 2009).

# CHAPTER 15: OVERALL DISCUSSION

In this chapter I will provide a recap and discussion of the main findings of the evaluation of **SPI**. I will then go onto conclude this chapter by answering the research question: ***which study design provides a robust method of evaluation of complex patient safety interventions?***

## 15.1 Main findings of the evaluation

Commentaries five years after the publication of two key reports on patient safety in 2000 were characterised by some despair at the apparent lack of progress in the US (Leape and Berwick, 2005). Taken in the round, the quantitative data collected in this evaluation seem to tell the story of an improving NHS. Whilst the staff survey shows little change between epochs, the patient survey shows improvement across all five dimensions pre-specified for the study, suggesting better patient experience. There was even an improvement in medical history-taking. Despite the high mortality rate in the sample of case notes, the adverse event rate was low. There were encouraging trends in the quality of patient care. Firstly, the baseline performance across hospitals was over 90% on many criteria relating to quality, leaving very little headroom for improvement. For example, over 90% of patients with an acute exacerbation of obstructive airways disease received steroids when indicated and the rates of peri-operative prophylaxis against venous thrombosis and wound infection approached 100%. Secondly, where there was scope for improvement, many examples of improved, and none of worsening, practice were found. Both the vigilance of monitoring

vital signs on acute medical wards and the use of severity scoring have seen sharp significant increases and there was a strong upward trend in the incidence of intra-operative temperature monitoring. Rates of hand-washing have increased (if consumption of cleansing materials is accepted as a surrogate) and the incidence of C. diff and MRSA infection has plummeted.

### 15.1.1 Participant experience

The stakeholder interviews and focus groups, conducted in the ***SPI$_1$*** evaluation demonstrated that the **SPI** was greeted enthusiastically at a strategic level in hospitals where it was delivered (see Appendix B). However the ethnography within the same evaluation suggests that staff at the sharp end of medical wards generally had only 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 (see Appendix C). A similar picture emerges from the staff survey. Control and **SPI** hospitals were mostly indistinguishable at baseline and within both evaluations 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. Within the ***SPI$_1$*** evaluation the results favoured the intervention; however, this is reversed in the ***SPI$_2$*** evaluation as the controls improved the most. Taken together, these findings suggest that the impact of **SPI** at medical ward level was at best modest.

### 15.1.2 Controls *versus* SPI

There was one significant finding in the ***SPI$_1$*** evaluation and that was the recording of vital signs at 12 hours. However, on the whole the quantitative data suggests that it is difficult to detect an additive **SPI** effect. Statistically significant observations were made but not between the two groups of hospitals (apart from the one previously mentioned. Quantitative evaluation of response to specific **SPI** targets (items 1a, b, c; 2a, b; 3a; 4 a, b. c, d in Table 3.2) overall 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 periods in both evaluations, but a similar improvement was also observed in the control hospitals. This is likely to be due to contemporaneous policy shifts and other external imperatives encouraging better detection and response to deteriorating patients. The use of the CURB score also improved markedly, but from a very low base.

Prescribing error rates are very sensitive to the methodology (Dean et al., 2005) used to make the measurements. Inter-observer reliability was good. Overlapping observation periods between epochs, enabled the detection of learning and fatigue effects and hence we were able to allow for these in the analysis. Although there was 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 endpoint cannot reliably be used to confirm or refute an **SPI** effect.

One anomaly is the drop in mortality among the acute medical cases in the $SPI_2$ hospitals and an unexplained rise in the control hospitals, such that in the $SPI_2$ evaluation the difference-in-differences would have been just significant if the $p<0.05$ threshold had been selected *a priori*. However, this finding does not align well with either the explicit review of the quality of care or the adverse event tally observed among those same case-notes and it is unlikely that failure to implement **SPI** was causative of an *increase* in mortality among respiratory patients in control hospitals. The analysis of cases of those who had died show no differences in avoidable deaths across the intervention *vs*. control hospitals (see Table 10.4)

Dramatic improvements in use of hand-washing materials and in infection rates produced near mirror image results. The NHS seems responsive to the need to change in certain ways and it is hard to discern any additive effect of the **SPI** initiative. Overall, there is little evidence that good or improved quality and safety in participating NHS hospitals can be reliably attributed to an additive effect of the **SPI**.

### 15.1.3 Interpretation

A large number of different observations have been made. Many of these observations relate to specific **SPI** objectives such as the patient at risk of deterioration, infection control, peri-operative care and intensive care. Statistically significant observations were made but not between the two groups of hospital. This broadly null additive effect of **SPI** (for both $SPI_1$ and $SPI_2$ intervention phases) on patient care should not, however, be translated into a conclusion that there was evidence of no effect. While a null result can never be proven, this

is a greater problem for quality initiatives, where small effect sizes may nevertheless be cost-effective, than it is for studies of clinical effectiveness. It can however be translated, less problematically, into the conclusion that any effect was not large, where large is defined in terms of observed confidence limits. To put this idea in another way, the results are compatible with effects on many end points of a magnitude that lies below the threshold that can be detected statistically in a study of this size.

These 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%" (Health Foundation, 2006; Shirley, 2008). However, the 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 USA (Wachter and Pronovost, 2006).

The findings will come as less of a surprise to observers who start from the premise that it is difficult to achieve improvements in the quality of care and reduce error rates through generic management initiatives, however enthusiastically they are welcomed at strategic level. Creating deep-seated, systemic cultural change in organisations, by means of an external initiative with a modest budget over a limited time scale may be viewed as almost quixotic by more sceptical observers. It was 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, there was a failure to find an intervention effect on more specific targets such as monitoring vital signs or medicines reconciliation.

Lack of a measured additive **SPI** effect might be explained in several ways:

First, improvements may have occurred at a magnitude that eluded statistical detection. For example, the sample was large for the case-note review of acute respiratory patients, with over 1,400 cases it had sufficient statistical power to detect material changes in actions that should affect all patients (e.g. regular monitoring of all vital signs). Power was lower for contingent actions that only applied to smaller sub-groups (i.e. for patients whose condition deteriorated). The English threshold under which an intervention is judged cost-effective is about £30,000 (€33k; $46k) per quality adjusted life year (QALY). The **SPI** would, therefore, need to save fewer than seven lives within $\boldsymbol{SPI_1}$ hospitals with a mean duration of five years to justify the investment of about £775k (€860k; $1.2 million) per $\boldsymbol{SPI_1}$ 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 in each year the signal would be lost in the noise.

Second, there may have been improvements in aspects of safety targeted by individual hospitals. For example, care of cardiovascular disease was identified for special attention in some hospitals. It is *a priori* 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 (Pronovost et al., 2006a), can be found in the literature. The question in this thesis 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.

A third explanation might lie in programme design. It is possible that organisational interventions of this type are simply not highly efficacious and that alternative approaches, such as initiatives focused on professional networks, could be more powerful, as suggested in a study of motivations to change in a maternity context (Wilson et al., 2002).

Fourth, it is possible that the design and implementation of the **SPI** might have not been optimal. Looking back over the evaluations of both programmes, and following many conversations with those responsible for this and other interventions with similar aims, it was suggested that the method by which vertical and horizontal spread of **SPI** might have been achieved was incompletely specified.

While senior stakeholders stressed the bottom-up nature of the intervention (see Appendix B), this was not how it was perceived by most ward staff (see Appendix C). 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 may have identified more precisely how it would work and under what conditions. It is also possible that the **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 studies 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 under-estimated. The intensity of the intervention may not have been sufficient to engender large-scale change (e.g. ***SPI1*** hospitals received £775k [€860k; $1.2 million] spent over 18 months in hospitals with annual budgets of £150 to £300 million [€167m - €334m; $230m – $460m] might simply be too small a "dose", 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 (Bowles, 2003; Shebl et al., 2009). It may also be the case that the impact of measures such as WalkRounds, safety briefings, and SBAR may be too diffuse to have discernible impacts.

A combination of a more explicit programme theory and organisational theory of change might have focused more attention on ensuring clinical engagement, encouraged an earlier recognition that the intervention was broad relative to resource, and identified that effects were likely to be localised in response to "dose" of intervention. In that case, a more focused and less ambitious intervention, and somewhat narrower evaluation, might have ensued.

A fifth explanation for the absence of a measured additional effect of **SPI** might lie in the extent of the policy-level programmes and initiatives that were largely contemporaneous

with the **SPI**. These shared some of its goals, principles and methods, and were targeting several of the same clinical processes as the **SPI**, for example, the "*clean**your**hands*" campaign ran continuously from late 2004/2005 onwards, promoting the same goal of improved hand hygiene as the **SPI**. Similarly, improving recognition and response to deterioration in hospitalised patients (an **SPI** goal) became a focus of policy attention, and guidelines on recognition and response to acutely ill patients were issued by NICE in 2007 (National Institute for Health and Clinical Excellence, 2007). Perhaps most significantly, several initiatives were explicitly modelled upon IHI techniques and principles, which began to have increasing impact on policy-making at around the time that the **SPI** was launched (and it is possible that this was not a coincidence). For example, the Department of Health's Saving Lives programme, beginning in June 2005 with a revised version in 2007 (Department of Health, 2007), included a self-assessment tool for hospitals to assess their managerial and clinical performance, and a set of "High Impact Interventions" that were similar to the IHI bundles and were aimed at several clinical processes also targeted by the **SPI**. In addition, the Health Act 2006 introduced new legislation on mandatory requirements on prevention and control of HCAIs.

It is further relevant that many of these policy initiatives had already been anticipated by significant consensus within professional societies and Medical Colleges about the appropriate measures to be adopted, and thus enjoyed considerable professional legitimacy – a crucial factor in promoting safe and effective practice (Dixon-Woods, 2010).

Moreover, 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 (Brown et al., 2008c); any changes may otherwise be falsely attributed to the intervention.

Finally, it is possible that any additional effects associated with **SPI** may simply not be detected yet. The difference between the control hospitals and the ***SPI2*** hospitals was that the ***SPI2*** hospitals benefited from a specific organisational intervention designed to promote the building of improvement skills into systems of care. Any **SPI** effect may be in the form of "stickiness"; **SPI** hospitals may potentially be better equipped to show sustained improvements after the policy spotlight has moved elsewhere. If, however, no differences can be detected in the longer term, the role of organisational interventions of this type in promoting safety will require further examination.

Patient safety is difficult (Dixon-Woods, 2010) 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 and 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, and improved understanding of how practitioners, middle managers and organisational systems can be better supported in the face of daunting complexity and multiple priorities.

### 15.1.4 Theory building

In the previous section certain ideas that might explain the mostly null comparative results obtained in the evaluation of **SPI** were discussed. These covered the scope of the intervention (the dose may have been too small), the ambitious time scale, and certain features of the intervention, such that it was not fully "owned" by middle grade staff. The observation that the NHS has adopted certain good practices over the same time scale as the initiative suggests a further, rather more radical idea; the originators of **SPI**, along with many opinion formers in management, are working with the wrong theory. The current theory is largely built around the concept of organisations and the pivotal role they are thought to play in "driving up quality". However, when the NHS wishes to change practice, it generally works with professional affiliations such as intensive care societies and Medical Colleges. Research into why evidence based guidelines were adopted or ignored in a maternity care context showed that staff were influenced almost entirely through personal/professional networks and hardly at all via the management route (Wilson et al., 2002). That, is not to say that hospitals do not have an essential role to play, but the idea put forward is that this role is enabling, not generative, in the main. In this respect medical services (and perhaps other highly professionalized groups) may differ from many industries where the hegemony of the organisation can drive change more directly. From the perspective of the evaluation the changes observed across 18 hospitals in our sample are unlikely to have resulted from concerted and simultaneous management action; this might be expected in the **SPI** hospitals, but it is unlikely that this would be mimicked simultaneously in the board rooms of control institutions. The idea put forward here is that

health services may have learned precisely the wrong lesson by adopting certain ideas and mind sets from managers and theorists with an industrial background.

### 15.1.5 Next steps

There are two dangers to be avoided: The first danger is despair and resort to nihilism. The corresponding danger is to privilege positive results over null results. Objective proof shorn 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 enough. The challenge is to build on these observed effects. The staff interviewed theorised about the way forward. They proposed offering more support to the middle layer of management, and engaging clinical leaders at earlier stages and encouraging clinical ownership as a way of securing success in the future. Reducing the number of areas to be tackled, and avoiding areas where there is scientific contestation or dispute about whether something is an important problem were also seen as critical. It was clear that hospitals had learned that addressing issues of legitimacy was a key task. They had identified that introducing initiatives that generated more "paperwork" were unlikely to secure cooperation from 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 (see Appendix C) 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 without disempowering (or perhaps even by making frontline staff feel they were the objects not subjects of inspiration). Although the **SPI** intervention clearly *intended* to achieve this effect, it seems that success overall was limited. In contrast, the Veterans Health Administration "QUERI System" (Quality Enhancement Research Initiative) (Stetlet et al., 2008) 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 QUERI tasks groups; effort, focus and resources are placed in finding out where practice is sub-standard and then tackling the specific causes one by one. There is evidence that such a focused approach spills over into other targets not specifically identified for action, the so-called "halo-effect" (Eccles et al., 1996; Francis and Perlin, 2006). It can be hypothesised that Donabedian's chain (Donabedian, 1980) is a two way street, where efforts to strengthen the front line of practice influence culture as much as culture influences the front line. Far from abandoning the topic of safety improvement or decrying the **SPI** initiative, the results point to promising and reasonable hypotheses about how to catalyse a more truly holistic approach to safety.

## 15.2 Lessons for the study design of future evaluations

In this thesis, I have described the evaluation of a large, highly complex patient safety intervention, which sought to generically strengthen the hospital system whilst improving frontline activities. In comparison to the interventions retrieved from the literature search in Chapter 2, none were as complex as the **SPI** and the evaluation of **SPI** mirrored its complexity. It was a multi-centre, before and after study, using non-randomised controls and mixed-methods. This approach has relative strengths and weaknesses and is discussed below.

### 15.2.1 Strengths and weaknesses

The study was based on a before and after design with contemporaneous controls (Brown et al., 2008b). Such a design is not as robust as a cluster randomised trial. However, despite some notable exceptions (Landon et al., 2007), most quality improvement reports lack contemporaneous controls (see Chapter 2); such a design would evidently have been misleading in this case since the sharp improvement of monitoring of vital signs, the use of a formal scoring system, intra-operative temperature monitoring and a reduction in infection rates in the intervention hospitals could have been incorrectly attributed to the **SPI**.

A limitation of the $SPI_1$ evaluation was that controls were matched with $SPI_2$ hospitals. Hospitals were selected for **SPI** because they were perceived to have contained positive features (see 3.2). Results might have been biased because: (i) **SPI** hospitals might have had less headroom for improvement; and (ii) controls might have had higher than average

performance, particularly since half were also selected as future ***SPI2*** intervention hospitals. Results might have been biased in favour of **SPI** because: (i) intervention hospitals were selected, not chosen at random (the reverse of the possible bias mentioned above); (ii) while both control and intervention hospitals gave consent for the evaluation, this may have had a differentially motivating effect in intervention hospitals (a kind of Hawthorne effect that could not be avoided by randomisation).

In addition, the matching between the ***SPI2*** hospitals and the controls was also far from perfect (the algorithm for selecting controls can be found in section 4.3). At the time of choosing controls, size was based on the number of staff at trust level. When this was revisited after the evaluation with bed numbers at the hospital level (the level of analysis used in this study) there was one discrepancy in a matched pair (pair 9 in Table 4.3), in which the number of beds in the control hospital was double that of the intervention, thus confirming imperfect matching.

The location (urban/rural) is also an important indicator of the quality of care (Keeler et al., 1992) and thus an important consideration when matching between control and intervention hospitals. In this evaluation, a value judgement was made on the location of the hospital but more sophisticated methods do exist. Ideally patients of a hospital would be assigned to their "neighbourhood" (for example, ward, local authority, lower level super output areas) and then the attributes of variables of interest, such as location would be applied. The proportions of these variables then determine the attribute of that hospital i.e. urban or rural.

As well as using a more refined method of matching other attributes should be considered , for example, income deprivation scores (Noble M et al., 2008). The score is effectively a proportion of people in a neighbourhood who live in a household with less than 60% of the national median income and/or are in receipt of one of a number of means tested welfare benefits.

One advantage of contemporaneous controls is that the groups can be compared at baseline. There were differences at baseline for some observations (most notably hospital mortality rate within the $SPI_2$ evaluation) but not for others. Baseline rates on the staff and patient surveys were similar and there is little to distinguish the two groups of hospitals on the explicit reviews in either acute medical or surgical patients; none of the 17 vital sign criteria differed significantly between the two groups of hospitals for example. Thus most of the comparisons that were made were based on end points where no material differences were evident across the groups compared.

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 **SPI** hospitals at baseline; most observations are similar and where statistically significant differences exist, these are of small magnitude. The data do not support the idea that the **SPI** hospitals had such excellent practice at baseline compared to controls that they were jeopardised in the comparison.

Source data for most end points was collected by independent researchers working across the various hospitals and a supply chain of anonymised case-notes was set up for this purpose. However certain data was collected in the participating hospitals (infection rates and data from the ICU) and this could lead to bias in the comparative study if hospital-based observers were motivated to show **SPI** in a good (or bad) light. However, any bias must have affected both sets of hospitals approximately equally since the comparative results are null. Moreover it is unlikely that the observed dramatic reductions in infection rates across all hospitals are the result of "gaming" given the statutory duty to report certain infections when they are identified in the laboratory.

A particular strength of the study arises from possibilities for "triangulation", i.e. some of the observations (qualitative and quantitative) act as a kind of internal control for others. While the funding envelope did not permit qualitative studies to be built into the $SPI_2$ design (as in $SPI_1$), the study did provide the following internal controls:

- findings on use of hand-washing materials and two different types of infection support the hypothesis of general improvement in this area;
- the observation that vital signs were recorded with increasing diligence while use of risk scoring was also used more frequently supports the idea that patients at risk of deterioration are being taken more seriously;
- mortality rates on the acute medical wards could be triangulated, not only by an audit of compliance with process standards, but also by scrutinising each death

in the sample to see if it could have been caused by poor care (only two of the 30 deaths in the post-intervention period were preventable).

Further evidence on this point could have been established by examining the incidence of unsuspected cardiac arrest ("crash calls") but found that this information is not yet collected in a consistent way. The evaluation of ***SPI$_1$*** included qualitative observations which can provide yet a further form of internal control. However the study sponsor felt that theoretical saturation had already been reached in the first evaluation. For example, ethnographic sub-studies within the ***SPI$_1$*** evaluation did indeed confirm that ward staff had taken the importance of close observations of sick patient increasingly to heart.

Within the literature review none of the studies that made before and after comparisons with contemporaneous controls used the "difference-in-difference" method, a method of calculating sample sizes which is based on a clusters propensity to change (see 2.1.3.2). This approach allows sufficient power with a smaller sample and its utilisation within this evaluation is unique. Perhaps with wider dissemination it will encourage future evaluations of complex patient safety interventions to use more robust methods as sample size considerations can be smaller than currently thought.

## 15.3 Conclusion

In this thesis I have described a robust study design to evaluate complex patient safety interventions. The evaluation has taken a mixed method approach, using both quantitative

and qualitative methods. All quantitative observations have been made with contemporaneous controls and uniquely a “difference-in-difference” method was used.

The approach has been vindicated – measures of fidelity, intervening variables, process and outcome, supported by qualitative methods – and has provided a wealth of information of the effectiveness of the intervention and theories as to why the intervention worked the way it did. This evaluation has been commended by Pronovost et al., (2011) as “*a model for the field*” and way in which future evaluations should be conducted.

# REFERENCE LIST


Agostini JV, Zhang Y, Inouye SK. Use of a computer-based reminder to improve sedative-hypnotic prescribing in older hospitalized patients. *Journal of the American Geriatrics Society*. 2007; 55(1):43-8.

Altman D. *Practical statistics for medical research*. London: Chapman and Hall; 1991.

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

Beckett CD and Kipnis G. Collaborative communication: integrating SBAR to improve quality/patient safety outcomes. *Journal for Healthcare Quality*. 2009; 31(5):19-28.

Begaud B, Martin K, Haramburu F, Moore N. Rates of spontaneous reporting of adverse drug reactions in France. *JAMA*. 2002; 288(13):1588.

Benneyan JC, Lloyd RC, Plsek PE. Statistical process control as a tool for research and healthcare improvement. *Qual Saf Health Care*. 2003; 12(6):458-64.

Berwick D and Rothman M. Pursuing perfection: an interview with Don Berwick and Michael Rothman. Interview by Andrea Kabcenell and Jane Roessner. *Jt Comm J Qual Improv*. 2002; 28(5):268-78.

Berwick DM, Calkins DR, McCannon CJ, Hackbarth AD. The 100 000 lives campaign: setting a goal and a deadline for improving health care quality. *JAMA*. 2006; 295(3):324-7.

Bisognano, C., McCannon, J., & Botwinick, L. A campaign for 100,000 lives. Trustee 58, 12-4. 12-9-2005.

Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ*. 1996; 312(7040):1215-8.

Bland JM and Altman DG. Multiple significance tests: the Bonferroni method. *BMJ*. 1995; 310(6973):170.

Bonello RS, Fletcher CE, Becker WK, Clutter KL, Arjes SL, Cook JJ, Petzel RA. An intensive care unit quality improvement collaborative in nine Department of Veterans Affairs hospitals: reducing ventilator-associated pneumonia and catheter-related bloodstream infection rates. *Jt Comm J Qual Patient Saf*. 2008; 34(11):639-45.

Borrill CS, West MA, Shapiro D, Rees A. Team working and effectiveness in health. *British Journal of Healthcare Management*. 2000; 6(8):364-71.

Bowles, J. B. 2003, "An assessment of RPN prioritization in a failure modes effects and criticality analysis", in *Reliability and Maintainability Symposium, 2003. Annual*, 2003 edn, pp. 380-386.

Boxwala AA, Dierks M, Keenan M, Jackson S, Hanscom R, Bates DW, Sato L. Organization and representation of patient safety data: current status and issues around generalizability and scalability. *J Am Med Inform Assoc*. 2004; 11(6):468-78.

Brackbill ML, Kline VT, Sytsma CS, Call JT. Intervention to increase the proportion of acute myocardial infarction or coronary artery bypass graft patients receiving an order for aspirin at hospital discharge. *Journal of Managed Care Pharmacy*. 2010; 16(5):329-36.

Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, Newhouse JP, Weiler PC, Hiatt HH. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *N Engl J Med*. 1991; 324(6):370-6.

Briesacher B, Limcangco R, Simoni-Wastila L, Doshi J, Gurwitz J. Evaluation of nationally mandated drug use reviews to improve patient safety in nursing homes: a natural experiment. *J Am Geriatr Soc*. 2005; 53(6):991-6.

British Orthopaedic Association. Primary Total Hip Replacement: A Guide to Good Practice2006.

British Thoracic Society. BTS Guidelines for the Management of Community Acquired Pneumonia in Adults. *Thorax*. 2001; 56(90004):iv1-64.

British Thoracic Society. BTS Guidelines for the Management of Community Acquired Pneumonia in Adults 2004 Update. 2004. [cited *http://www.brit-thoracic.org.uk/Portals/0/Clinical%20Information/Pneumonia/Guidelines/MACAPrevisedApr04.pdf*

Brook RH, Kamberg CJ, McGlynn EA. Health system reform and quality. *JAMA*. 1996; 276(6):476-80.

Brown C and Lilford R. Evaluating service delivery interventions to enhance patient safety. *BMJ*. 2008; 337(dec17_1):a2764.

Brown C, Hofer T, Johal A, Thomson R, Nicholl J, Franklin BD, Lilford RJ. An epistemology of patient safety research: a framework for study design and interpretation. Part 1. Conceptualising and developing interventions. *Qual Saf Health Care*. 2008a; 17(3):158-62.

Brown C, Hofer T, Johal A, Thomson R, Nicholl J, Franklin BD, Lilford RJ. An epistemology of patient safety research: a framework for study design and interpretation. Part 2. Study design. *Qual Saf Health Care*. 2008b; 17(3):163-9.

Brown C, Hofer T, Johal A, Thomson R, Nicholl J, Franklin BD, Lilford RJ. An epistemology of patient safety research: a framework for study design and interpretation. Part 3. End points and measurement. *Qual Saf Health Care*. 2008c; 17(3):170-7.

Brown C, Hofer T, Johal A, Thomson R, Nicholl J, Franklin BD, Lilford RJ. 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*. 2008d; 17(3):178-81.

Brown, C. A. 2004, *Student medical school choice in the UK,* University of Birmingham.

Bruce J, Russell EM, Mollison J, Krukowski ZH. The measurement and monitoring of surgical adverse events. *Health Technol Assess*. 2001; 5(22):1-194.

Campbell DT and Stanley JC. *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally; 1963.

Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, Tyrer P. Framework for design and evaluation of complex interventions to improve health. *BMJ*. 2000; 321(7262):694-6.

Campbell MK, Thomson S, Ramsay CR, MacLennan GS, Grimshaw JM. Sample size calculator for cluster randomized trials. *Comput Biol Med*. 2004; 34(2):113-25.

Campbell NC, Murray E, Darbyshire J, Emery J, Farmer A, Griffiths F, Guthrie B, Lester H, Wilson P, Kinmonth AL. Designing and evaluating complex interventions to improve health care. *BMJ*. 2007; 334(7591):455-9.

Campbell SE, Walke AE, Grimshaw JM, Campbell MK, Lowe GD, Harper D, Fowkes FG, Petrie JC. The prevalence of prophylaxis for deep vein thrombosis in acute hospital trusts. *Int J Qual Health Care*. 2001; 13(4):309-16.

Caplan RA, Posner KL, Cheney FW. Effect of outcome on physician judgments of appropriateness of care. *JAMA*. 1991; 265(15):1957-60.

Care Quality Commission and Picker Institute. NHS patient surveys. 2010. [cited 2010 Aug 18]. *Available at:*

Chern CH, How CK, Wang LM, Lee CH, Graff L. Decreasing clinically significant adverse events using feedback to emergency physicians of telephone follow-up outcomes. *Annals of Emergency Medicine*. 2005; 45(1):15-23.

Classen DC, Jaser L, Budnitz DS. Adverse drug events among hospitalized Medicare patients: epidemiology and national estimates from a new approach to surveillance. *Jt Comm J Qual Patient Saf*. 2010; 36(1):12-21.

Cleary PD and Edgman-Levitan S. Health care quality. Incorporating consumer perspectives. *JAMA*. 1997; %19;278(19):1608-12.

Cochrane AL. *Effectiveness and efficiency: random refelections on health services*. London: Nuffield Provincial Hospitals Trust; 1979.

Cohen J. A coefficent of agreement for nominal scales. *Educ Psychol Meas*. 1960; 20:37-46.

Cohen MM, Kimmel NL, Benage MK, Cox MJ, Sanders N, Spence D, Chen J. Medication safety program reduces adverse drug events in a community hospital. *Quality & Safety in Health Care*. 2005; 14(3):169-74.

Connelly JB. Evaluating complex public health interventions: theory, methods and scope of realist enquiry. *J Eval Clin Pract*. 2007; 13(6):935-41.

Cook R and Woods D. Operating at the sharp end: the complexity of human error. In: *Human Error in Medicine*, Bogner M, ed. Lillsdale NJ: Lawrence Erlbaum; 1994. pp. 255-310.

Cook TD and Campbell DT. *Quasi-experimentation:design and analysis issues for field settings*. Chicago: RandMcNally; 1979.

Cousins D and Hatoum H. The development of a computerised quality assurance system for clinical pharmacy. *Int J Pharmact Pract*. 1991; 1:86-9.

Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*. 2008; 337:a1655.

Croskerry P. Clinical cognition and diagnostic error: applications of a dual process model of reasoning. *Adv Health Sci Educ Theory Pract*. 2009; 14 Suppl 1:27-35. Epub;%2009 Aug 11.:27-35.

Cullen DJ, Bates DW, Small SD, Cooper JB, Nemeskal AR, Leape LL. The incident reporting system does not detect adverse drug events: a problem for quality improvement. *Jt Comm J Qual Improv*. 1995; 21(10):541-8.

Dean B, Barber N, Schachter M. What is a prescribing error? *Qual Health Care*. 2000; 9(4):232-7.

Dean B, Schachter M, Vincent C, Barber N. Prescribing errors in hospital inpatients: their incidence and clinical significance. *Qual Saf Health Care*. 2002; 11(4):340-4.

Dean FB, Vincent C, Schachter M, Barber N. The incidence of prescribing errors in hospital inpatients: an overview of the research methods. *Drug Saf*. 2005; 28(10):891-900.

Deming WE. *Out of the crisis*. Cambridge, Massachusetts: Massachusetts Institute of Technology; 1986.

Department of Health Expert Group. An organisation with a memory: The Stationery Office; 2000.

Department of Health Expert Group. An organisation with a memory: The Stationery Office; 2000.

Department of Health. Beds open overnight in England. 2010. [cited

Department of Health. Saving Lives: reducing infection, delivering clean and safe care. London: DH Publications; 2007.

Department of Health. Wet, soap, wash, rinse, dry. 2006.

Dixon-Woods M, Suokas A, Pitchforth E, Tarrant C. An ethnographic study of classifying and accounting for risk at the sharp end of medical wards. *Social Science & Medicine*. 2009; 69(3):362-9.

Dixon-Woods M. Why is patient safety so hard? A selective review of ethnographic studies. *Journal of Health Services Research and Policy*. 2010; 15:11-6.

Donabedian A. Explorations in quality assessment and monitoring. In: *The definition of quality and approaches to its assessment*, Griffith JR, ed. Ann Arbor Michigan: Health Administration Press; 1980. pp. 4-163.

Donner A. Some aspects of the design and analysis of cluster randomised trials. *Appl Stat*. 1998; 47:95-113.

Eccles M, Clapp Z, Grimshaw J, Adams PC, Higgins B, Purves I, Russell I. North of England evidence based guidelines development project: methods of guideline development. *BMJ*. 1996; 312(7033):760-2.

Eccles M, Grimshaw J, Campbell M, Ramsay C. Research designs for studies evaluating the effectiveness of change and improvement strategies. *Qual Saf Health Care*. 2003; 12(1):47-52.

Edwards SJ, Braunholtz DA, Lilford RJ, Stevens AJ. Ethical issues in the design and conduct of cluster randomised controlled trials. *BMJ*. 1999; 318(7195):1407-9.

Elstad EA, Lutfey KE, Marceau LD, Campbell SM, von dem KO, McKinlay JB. What do physicians gain (and lose) with experience? Qualitative results from a cross-national study of diabetes. *Soc Sci Med*. 2010; 70(11):1728-36.

Ernst MM, Wooldridge JL, Conway E, Dressman K, Weiland J, Tucker K, Seid M. Using quality improvement science to implement a multidisciplinary behavioral intervention targeting pediatric inpatient airway clearance. *Journal of Pediatric Psychology*. 2010; 35(1):14-24.

Francis J and Perlin JB. Improving performance through knowledge translation in the Veterans Health Administration. *J Contin Educ Health Prof*. 2006; 26(1):63-71.

Frankel A, Grillo SP, Baker EG, Huber CN, Abookire S, Grenham M, Console P, O'Quinn M, Thibault G, Gandhi TK. Patient Safety Leadership WalkRounds at Partners Healthcare: learning from implementation. *Jt Comm J Qual Patient Saf*. 2005; 31(8):423-37.

Franklin BD, O'Grady K, Parr J, Walton I. Using the internet to deliver education on drug safety. *Quality & Safety in Health Care*. 2006; 15(5):329-33.

Freedman L. Bayesian statistical methods. *BMJ*. 1996; 313(7057):569-70.

Gallagher M, Oliver K, Hurwitz M. Improving the use of venous thromboembolism prophylaxis in an Australian teaching hospital. *Qual Saf Health Care*. 2009; 18(5):408-12.

Galpern D, Guerrero A, Tu A, Fahoum B, Wise L. Effectiveness of a central line bundle campaign on line-associated infections in the intensive care unit. *Surgery*. 2008; 144(4):492-5.

Garratt E. Key findings report: inpatient survey results 2008: Picker Institute Europe; 2009.

Giles S, Fletcher M, Baker M, Thomson R. Incident reporting and analysis. In: *Patient safety: research into practice*, Walshe K & Boaden R, eds. Oxford: Oxford University Press; 2005. pp. 108-17.

Glaser B and Thomson R. *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine Publishing Company; 1967.

Greenberg CC, Regenbogen SE, Studdert DM, Lipsitz SR, Rogers SO, Zinner MJ, Gawande AA. Patterns of communication breakdowns resulting in injury to surgical patients. *J Am Coll Surg*. 2007; 204(4):533-40.

Greenfield S, Kaplan SH, Kahn R, Ninomiya J, Griffith JL. Profiling care provided by different groups of physicians: effects of patient case-mix (bias) and physician-level clustering on quality assessment results. *Ann Intern Med*. 2002; 136(2):111-21.

Greengold NL, Shane R, Schneider P, Flynn E, Elashoff J, Hoying CL, Barker K, Bolton LB. The impact of dedicated medication nurses on the medication administration error rate: a randomized controlled trial. *Arch Intern Med*. 2003; 163(19):2359-67.

Groome LJ, Esters FT, Carter SG. Bundling unbundled charges: a case study in an academic obstetrical practice. *Journal of Medical Practice Management*. 2009; 24(6):367-71.

Guba E and Lincoln Y. Competing paradigms in qualitative research. In: *Handbook of qualitative research*, Denzin N & Lincoln Y, eds. Thousand Oaks, CA: Sage; 1994. pp. 105-77.

Guzzo RA, Jette D, Katzell RA. The effects of psychologically based intervention programs on worker productivity: a meta-analysis. *Personel Psychology*. 1985; 38(2):275-92.

Hardeman W, Sutton S, Griffin S, Johnston M, White A, Wareham NJ, Kinmonth AL. A causal modelling approach to the development of theory-based behaviour change programmes for trial evaluation. *Health Educ Res*. 2005; 20(6):676-87.

Harrison DA, Brady AR, Parry GJ, Carpenter JR, Rowan K. Recalibration of risk prediction models in a large multicenter cohort of admissions to adult, general critical care units in the United Kingdom. *Crit Care Med*. 2006; 34(5):1378-88.

Harrison DA, Brady AR, Rowan K. Case mix, outcome and length of stay for admissions to adult, general critical care units in England, Wales and Northern Ireland: the Intensive Care National Audit & Research Centre Case Mix Programme Database. *Crit Care*. 2004; 8(2):R99-111.

Harrison DA, Parry GJ, Carpenter JR, Short A, Rowan K. A new risk prediction model for critical care: the Intensive Care National Audit & Research Centre (ICNARC) model. *Crit Care Med*. 2007; 35(4):1091-8.

Hawe P, Shiell A, Riley T. Complex interventions: how "out of control" can a randomised controlled trial be? *BMJ*. 2004; 328(7455):1561-3.

Hawe P, Shiell A, Riley T. Complex interventions: how "out of control" can a randomised controlled trial be? *BMJ*. 2004; 328(7455):1561-3.

Hayward RA and Hofer TP. Estimating hospital deaths due to medical errors: preventability is in the eye of the reviewer. *JAMA*. 2001; 286(4):415-20.

Healey F, Monro A, Cockram A, Adams V, Heseltine D. Using targeted risk factor reduction to prevent falls in older in-patients: a randomised controlled trial. *Age Ageing*. 2004; 33(4):390-5.

Health Foundation. Description of the SPI intervention. (Personal Communication, 2009 Sep 10).

Health Foundation. Health Foundation. 2010a. [cited 2010 Aug 14a]. *Available at: www.health.org.uk*

Health Foundation. History of the Health Foundation. 2010b. [cited 2010 Jun 25b]. *Available at: http://www.health.org.uk/about_us/how_we_work/history_of_thf.html*

Health Foundation. Safer Patients Initiative trusts explain how they are making hospitals safer for patients. 2006. [cited 2009 Feb 10]. *Available at:*

Health Protection Agency. Fifth report of the mandatory surveillance of surgical site infection in orthopaedic surgery. April 2004 to March 2009. London: Health Protection Agency; 2009.

Health Protection Agency. Handwashing. 2010a. [cited 2010 Aug 15a]. *Available at:*

Health Protection Agency. MRSA and c.difficile infections fall by a third. 2010b. [cited *www.hpa.org.uk/NewsCentre/NationalPressReleases/2010PressReleases/100716MRSAandcdiffdownbyathird*

Healthcare Commission. Acute trusts - the views of staff: Key findings from the 2006 survey of NHS staff. London2007b.

Healthcare Commission. Making sense of your staff survey data. 2006. [cited 2009 Jan 27]. *Available at: http://www.healthcarecommission.org.uk/_db/_documents/Making_sense_of_your_Staff_Survey_Data_200703214701.pdf.*

Healthcare Commission. NHS national staff survey 2004: Summary of key findings. London2005.

Healthcare Commission. NHS National Staff Survey 2006: Key Findings: The Stationery Office; 2007a.

Heinrich HW. *Industrial accident prevention*. New York: McGraw Hill; 1931.

Henderson D, Dempsey C, Larson K, Appleby D. The impact of IMPACT on St. John's Regional Health Center. *Mo Med*. 2003; 100(6):590-2.

Hillman KM, Bristow PJ, Chey T, Daffurn K, Jacques T, Norman SL, Bishop GF, Simmons G. Antecedents to hospital deaths. *Intern Med J*. 2001; 31(6):343-8.

Hixson ED, Nadzam DM, Katzan IL, Furlan AJ. Quality improvement for stroke management at the Cleveland Clinic Health System. *Joint Commission Journal on Quality & Patient Safety*. 2005; 31(8):447-54.

Hutchinson A, Coster JE, Cooper KL, McIntosh A, Walters SJ, Bath PA, Pearson M, Young TA, Rantell K, Campbell MJ, Ratcliffe J. Comparison of case note review methods for evaluating quality and safety in health care. *Health Technol Assess*. 2010; 14(10):iii-x, 1.

Hutchinson TP. Focus on Psychometrics. Kappa muddles together two sources of disagreement: tetrachoric correlation is preferable. *Res Nurs Health*. 1993; 16(4):313-6.

Hutter B. *Regulation and risk: occupational health and safety on the railways*. Oxford University Press; 2001.

Institute for Healthcare Improvement & Health Foundation. Kick off meeting preparation materials. 1-12-2004.

Institute for Healthcare Improvement & Health Foundation. The health foundation's safer patients initiative: kick off meeting. 25-1-2005a. 25-1-2005a.

Institute for Healthcare Improvement & Health Foundation. The health foundation's safer patients initiative: learning session 2. 25-5-2005b.

Institute for Healthcare Improvement & Health Foundation. The health foundation's safer patients initiative: learning session 3. 17-11-2005c.

Institute for Healthcare Improvement & Health Foundation. The health foundation's safer patients initiative: learning session 4. 22-6-2006.

Institute for Healthcare Improvement. Failure modes and effects analysis (FMEA) tool (IHI tool). 2010c. [cited 2010 Jun 25c]. *Available at:*

Institute for Healthcare Improvement. How to guide - prevent surgical site infections2007.

Institute for Healthcare Improvement. Institute for Healthcare Improvement. 2010a. [cited 2010 Aug 14a]. *Available at: www.ihi.org*

Institute for Healthcare Improvement. Move your dot™:measuring, evaluating, and reducing hospital mortality rates (Part 1). Boston: Institute for Healthcare Improvement; 2003.

Institute for Healthcare Improvement. Safety briefings (IHI tool). 2010b. [cited 2010 Jun 25b]. *Available at:*

Iyer SB, Schubert CJ, Schoettker PJ, Reeves SD. Use of quality-improvement methods to improve timeliness of analgesic delivery. *Pediatrics*. 2011; 127(1):e219-e225.

Jeffries HE, Mason W, Brewer M, Oakes KL, Munoz EI, Gornick W, Flowers LD, Mullen JE, Gilliam CH, Fustar S, Thurm CW, Logsdon T, Jarvis WR. Prevention of central venous catheter-associated bloodstream infections in pediatric intensive care units: a performance improvement collaborative. *Infection Control & Hospital Epidemiology*. 2009; 30(7):645-51.

Kahn KL, Rubenstein LV, Draper D, Kosecoff J, Rogers WH, Keeler EB, Brook RH. The effects of the DRG-based prospective payment system on quality of care for hospitalized Medicare patients. An introduction to the series. *JAMA*. 1990; 264(15):1953-5.

Keeler EB, Rubenstein LV, Kahn KL, Draper D, Harrison ER, McGinty MJ, Rogers WH, Brook RH. Hospital characteristics and quality of care. *JAMA*. 1992; 268(13):1709-14.

Kernick D. Assessing complex interventions. Guidelines perpetuate inappropriate methods. *BMJ*. 2008; 337:a2473.

Kernick D. Assessing complex interventions. Guidelines perpetuate inappropriate methods. *BMJ*. 2008; 337:a2473.

Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Ann Fam Med*. 2004; 2(3):204-8.

Kirk SA, Campbell SM, Kennell-Webb S, Reeves D, Roland MO, Marshall MN. Assessing the quality of care of multiple conditions in general practice: practical and methodological problems. *Qual Saf Health Care*. 2003; 12(6):421-7.

Kohn L, Corrigan J, Donaldson M. *To err is human: building a safer health system*. National Academies Press; 2000.

Koplan KE, Regan S, Goldszer RC, Schneider LI, Rigotti NA. A computerized aid to support smoking cessation treatment for hospital patients. *J Gen Intern Med*. 2008; 23(8):1214-7.

Kotagal M, Lee P, Habiyakare C, Dusabe R, Kanama P, Epino HM, Rich ML, Farmer PE. Improving quality in resource poor settings: observational study from rural Rwanda. *BMJ*. 2009; 339:b3488.

Krimsky WS, Mroz IB, McIlwaine JK, Surgenor SD, Christian D, Corwin HL, Houston D, Robison C, Malayaman N. A model for increasing patient safety in the intensive care unit: increasing the implementation rates of proven safety measures. *Quality & Safety in Health Care*. 2009; 18(1):74-80.

Landis J and Koch G. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33:159-74.

Landon BE, Hicks LS, O'Malley AJ, Lieu TA, Keegan T, McNeil BJ, Guadagnoli E. Improving the Management of Chronic Disease at Community Health Centers. *N Engl J Med*. 2007; 356:921-34.

Landon BE, WIlson IB, McInnes K, Landrum MB, Hirschhorn L, Marsden PV, Gustafson D, Cleary PD. Effects of a Quality Improvement Collaborative on the Outcome of Care of Patients with HIV Infection: The EQHIV Study. *Ann Intern Med*. 2004; 140:887-96.

Landrigan CP, Rothschild JM, Cronin JW, Kaushal R, Burdick E, Katz JT, Lilly CM, Stone PH, Lockley SW, Bates DW, Czeisler CA. Effect of reducing interns' work hours on serious medical errors in intensive care units. *N Engl J Med*. 2004; 351(18):1838-48.

Leape LL and Berwick DM. Five years after to err is human: what have we learned? *JAMA*. 2005; 293(19):2384-90.

Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, Hebert L, Newhouse JP, Weiler PC, Hiatt H. The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study II. *N Engl J Med*. 1991; 324(6):377-84.

Levy MM, Dellinger RP, Townsend SR, Linde-Zwirble WT, Marshall JC, Bion J, Schorr C, Artigas A, Ramsay G, Beale R, Parker MM, Gerlach H, Reinhart K, Silva E, Harvey M, Regan S, Angus DC, Surviving SC. The Surviving Sepsis Campaign: results of an international guideline-based performance improvement program targeting severe sepsis. *Critical Care Medicine*. 2010; 38(2):367-74.

Lilford R, Edwards A, Girling A, Hofer T, Di Tanna GL, Petty J, Nicholl J. Inter-rater reliability of case-note audit: a systematic review. *J Health Serv Res Policy*. 2007; 12(3):173-80.

Lilford R, Mohammed MA, Spiegelhalter D, Thomson R. Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma. *Lancet*. 2004; 363(9415):1147-54.

Lilford R, Stirling S, Maillard N. Citation classics in patient safety research: an invitation to contribute to an online bibliography. *Qual Saf Health Care*. 2006; 15(5):311-3.

Lilford RJ and Braunholtz D. For Debate: The statistical basis of public policy: a paradigm shift is overdue. *BMJ*. 1996; 313(7057):603-7.

Lilford RJ, Mohammed MA, Braunholtz D, Hofer TP. The measurement of active errors: methodological issues. *Qual Saf Health Care*. 2003; 12 Suppl 2:ii8-12.:ii8-12.

Lingard L, Regehr G, Espin S, Devito I, Whyte S, Buller D, Sadovy B, Rogers D, Reznick R. Perceptions of operating room tension across professions: building generalizable evidence and educational resources. *Acad Med*. 2005; 80(10 Suppl):S75-S79.

Lipsey MW and Wilson DB. The efficacy of psychological, educational, and behavioral treatment. Confirmation from meta-analysis. *Am Psychol*. 1993; 48(12):1181-209.

Mackenzie M, O'Donnell C, Halliday E, Sridharan S, Platt S. Do health improvement programmes fit with MRC guidance on evaluating complex interventions? *BMJ*. 2010; 340:c185.

Macrae C. Learning from patient safety incidents: Creating participative risk regulation in healthcare. *Health, Risk & Society*. 2008; 10(1):53-67.

Marshall S, Harrison J, Flanagan B. The teaching of a structured tool improves the clarity and content of interprofessional clinical communication. *Qual Saf Health Care*. 2009; 18(2):137-40.

Mason P and Barnes M. Constructing Theories of Change. *Evaluation*. 2007:151-70.

McCulloch P, Kreckler S, New S, Sheena Y, Handa A, Catchpole K. Effect of a "Lean" intervention to improve safety processes and outcomes on a surgical emergency unit. *BMJ*. 2010; 341(7781):1043-6.

McGaughey J, Alderdice F, Fowler R, Kapila A, Mayhew A, Moutray M. Outreach and Early Warning Systems (EWS) for the prevention of intensive care admission and death of critically ill adult patients on general hospital wards. *Cochrane Database Syst Rev*. 2007(3):CD005529.

McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, Kerr EA. The quality of health care delivered to adults in the United States. *N Engl J Med*. 2003; 348(26):2635-45.

McPhail GL, Weiland J, Acton JD, Ednick M, Chima A, VanDyke R, Fenchel MC, Amin RS, Seid M. Improving evidence-based care in cystic fibrosis through quality improvement. *Archives of Pediatrics & Adolescent Medicine*. 2010; 164(10):957-60.

McQuillan P, Pilkington S, Allan A, Taylor B, Short A, Morgan G, Nielsen M, Barrett D, Smith G, Collins CH. Confidential inquiry into quality of care before admission to intensive care. *BMJ*. 1998; 316(7148):1853-8.

Medical Research Council. A framework for devlopment and evaluation of RCT's for complex interventions to improve health: Medical Research Council; 2000.

Medical Research Council. Developing and evaluating complex interventions: new guidance: Medical Research Council; 2008.

Merien AE, van d, V, Mol BW, Houterman S, Oei SG. Multidisciplinary team training in a simulation setting for acute obstetric emergencies: a systematic review. *Obstet Gynecol*. 2010; 115(5):1021-31.

Merle V, Moret L, Pidhorz L, Dujardin F, Gouin F, Josset V, Graveleau S, Petit J, Riou F, Lombrail P, Czernichow P. Does comparison of performance lead to better care? A pilot observational study in patients admitted for hip fracture in three French public hospitals. *Int J Qual Health Care*. 2009; 21(5):321-9.

Michel P, Quenon JL, de Sarasqueta AM, Scemama O. Comparison of three methods for estimating rates of adverse events and rates of preventable adverse events in acute care hospitals. *BMJ*. 2004; 328(7433):199.

Michie S and West M. Managing people and performance: An evidence-based framework applied to health service organisations. *International Journal of Management Reviews*. 2004; 5/6:91-111.

Midlov P, Holmdahl L, Eriksson T, Bergkvist A, Ljungberg B, Widner H, Nerbrand C, Hoglund P. Medication report reduces number of medication errors when elderly patients are discharged from hospital. *Pharm World Sci*. 2008; 30(1):92-8.

Miller MR, Griswold M, Harris JM, Yenokyan G, Huskins WC, Moss M, Rice TB, Ridling D, Campbell D, Margolis P, Muething S, Brilli RJ. Decreasing PICU catheter-associated bloodstream infections: NACHRI's quality transformation efforts. *Pediatrics*. 2010; 125(2):206-13.

Modak I, Sexton JB, Lux TR, Helmreich RL, Thomas EJ. Measuring safety culture in the ambulatory setting: the safety attitudes questionnaire--ambulatory version. *J Gen Intern Med*. 2007; 22(1):1-5.

Mohammed MA, Lilford RJ, Rudge G, Dawson JF, Deeks JJ, Girling A, and Carmalt M. 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; 2008b.

Mohammed MA, Worthington P, Woodall WH. Plotting basic control charts: tutorial notes for healthcare practitioners. *Qual Saf Health Care*. 2008a; 17(2):137-45.

Morrison B and Lilford RJ. Organisational research methods: closing the gap. *Lancet*. 2000; 355(9197):71.

Murff HJ, Patel VL, Hripcsak G, Bates DW. Detecting adverse events for patient safety research: a review of current methodologies. *J Biomed Inform*. 2003; 36(1-2):131-43.

Murray DM and Blitstein JL. Methods to reduce the impact of intraclass correlation in group-randomised trials. *Eval Rev*. 2003; 27:79-103.

Murtagh MJ, Thomson RG, May CR, Rapley T, Heaven BR, Graham RH, Kaner EF, Stobbart L, Eccles MP. Qualitative methods in a randomised controlled trial: the role of an integrated qualitative process evaluation in providing evidence to discontinue the intervention in one arm of a trial of a decision support tool. *Qual Saf Health Care*. 2007; 16(3):224-9.

National Audit Office. A safer place for patients: learning to improve patient safety. London: The Stationery Office; 2005.

National Evaluation of Sure Start. 2010. [cited 2010 Jun 1]. *Available at: http://www.ness.bbk.ac.uk*

National Institute for Health and Clinical Excellence. Clinical Guidelines CG50. Acutely ill patients in hospital. Recognition of and response to acute illness in adults in hospital. London: NICE; 2007.

National Institute for Health and Clinical Excellence. Clinical guideline 65:The management of inadvertent perioperative hypothermia in adults2008.

National Institute for Health and Clinical Excellence. Clinical Guideline 74: Prevention and treatment of surgical site infections2010.

National Patient Safety Agency. Building a memory: preventing harm, reducing risks and improving patient safety. London: National Patient Safety Agency; 2005.

National Patient Safety Agency. cleanyourhand campaign. 2010. [cited *www.npsa.nhs.uk/cleanyourhands*

National Patient Safety Agency. Safer care for the acutely ill patient: learning from serious incidents. London: National Patient Safety Agency; 2007.

Noble M, McLennanD, Wilkinson K, Whitworth A, Barnes H, and Dibbden C. The English Indices of Deprivation 2007. London: Department of Communities and Local Government; 2008.

Nolan K, Schall MW, Erb F, Nolan T. Using a framework for spread: The case of patient access in the Veterans Health Administration. *Jt Comm J Qual Patient Saf*. 2005; 31(6):339-47.

Nolan T, Resar R, Haraden C, and Griffin FA. Improving the reliability of health care. Boston: Institute for Healthcare Improvement; 2004.

Oakley A, Strange V, Bonell C, Allen E, Stephenson J. Process evaluation in randomised controlled trials of complex interventions. *BMJ*. 2006; 332(7538):413-6.

O'Hagan A, Buck CE, Daneshkhah A, Eiser JE, Garthwaite PH, Jenkinson DH, Oakley JE, Rakow T. *Uncertain judgements: eliciting experts' probabilities (statistics in practice)*. Wiley-Blackwell; 2006.

Paige JT, Aaron DL, Yang T, Howell DS, Chauvin SW. Improved operating room teamwork via SAFETY prep: a rural community hospital's experience. *World Journal of Surgery*. 2009; 33(6):1181-7.

Patterson ES, Roth EM, Woods DD, Chow R, Gomes JO. Handoff strategies in settings with high consequences for failure: lessons for health care operations. *Int J Qual Health Care*. 2004; 16(2):125-32.

Pronovost P, Needham D, Berenholtz S, Sinopoli D, Chu H, Cosgrove S, Sexton B, Hyzy R, Welsh R, Roth G, Bander J, Kepros J, Goeschel C. An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med*. 2006a; 355(26):2725-32.

Pronovost PJ, Berenholtz SM, Morlock LL. Is quality of care improving in the UK? *BMJ*. 2011; 342:c6646.

Pronovost PJ, Goeschel CA, Colantuoni E, Watson S, Lubomski LH, Berenholtz SM, Thompson DA, Sinopoli DJ, Cosgrove S, Sexton JB, Marsteller JA, Hyzy RC, Welsh R, Posa P, Schumacher K, Needham D. Sustaining reductions in catheter related bloodstream infections in Michigan intensive care units: observational study. *BMJ*. 2010; 340:c309. doi: 10.1136/bmj.c309.:c309.

Pronovost PJ, Miller MR, Wachter RM. Tracking progress in patient safety: an elusive target. *JAMA*. 2006b; 296(6):696-9.

Quillen K and Murphy K. Quality improvement to decrease specimen mislabeling in transfusion medicine. *Arch Pathol Lab Med*. 2006; 130(8):1196-8.

Rashad MR, Nielsen GA, Nolan K, Schall MW, and Sevin C. A framework for spread: from local improvements to system-wide change. Cambridge, Massachusetts: Institute for Healthcare Improvement; 2006.

Reason J. Human error: models and management. *BMJ*. 2000; 320(7237):768-70.

Resar R, Pronovost P, Haraden C, Simmonds T, Rainey T, Nolan T. Using a bundle approach to improve ventilator care processes and reduce ventilator-associated pneumonia. *Jt Comm J Qual Patient Saf*. 2005; 31(5):243-8.

Resar RK, Rozich JD, Simmonds T, Haraden CR. A trigger tool to identify adverse events in the intensive care unit. *Jt Comm J Qual Patient Saf*. 2006; 32(10):585-90.

Riggio JM, Sorokin R, Moxey ED, Mather P, Gould S, Kane GC. Effectiveness of a clinical-decision-support system in improving compliance with cardiac-care quality measures and supporting resident training. *Academic Medicine*. 2009; 84(12):1719-26.

Robb E, Jarman B, Suntharalingam G, Higgens C, Tennant R, Elcock K. Using care bundles to reduce in-hospital mortality: quantitative survey. *BMJ*. 2010; 340:c1234.

Rogers EM. *Diffusion of innovations*. Free Press; 2003.

Rosen L, Manor O, Engelhard D, Zucker D. In defense of the randomized controlled trial for health promotion research. *Am J Public Health*. 2006; 96(7):1181-6.

Ryckman FC, Schoettker PJ, Hays KR, Connelly BL, Blacklidge RL, Bedinghaus CA, Sorter ML, Friend LC, Kotagal UR. Reducing surgical site infections at a pediatric academic medical center. *Jt Comm J Qual Patient Saf*. 2009; 35(4):192-8.

Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. *N Engl J Med*. 1987; 316(8):450-5.

Saint S, Hofer TP, Rose JS, Kaufman SR, McMahon LF, Jr. Use of critical pathways to improve efficiency: a cautionary tale. *Am J Manag Care*. 2003; 9(11):758-65.

Sandelowski M. Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Res Nurs Health*. 2000; 23(3):246-55.

Schnipper JL, Hamann C, Ndumele CD, Liang CL, Carty MG, Karson AS, Bhan I, Coley CM, Poon E, Turchin A, Labonville SA, Diedrichsen EK, Lipsitz S, Broverman CA, McCarthy P, Gandhi TK. Effect of an electronic medication reconciliation application and process redesign on potential adverse drug events: a cluster-randomized trial. *Archives of Internal Medicine*. 2009; 169(8):771-80.

Schulz KF and Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet*. 2002; 359(9307):696-700.

Semmens JB, Aitken RJ, Sanfilippo FM, Mukhtar SA, Haynes NS, Mountain JA. The Western Australian Audit of Surgical Mortality: advancing surgical accountability. *Med J Aust*. 2005; 183(10):504-8.

Sexton JB, Helmreich RL, Neilands TB, Rowan K, Vella K, Boyden J, Roberts PR, Thomas EJ. The Safety Attitudes Questionnaire: psychometric properties, benchmarking data, and emerging research. *BMC Health Serv Res*. 2006; 6:44.

Shebl NA, Franklin BD, Barber N. Is Failure Mode and Effect Analysis Reliable? *Journal of Patient Safety*. 2009; 5(2).

Shepperd S, Lewin S, Straus S, Clarke M, Eccles MP, Fitzpatrick R, Wong G, Sheikh A. Can We Systematically Review Studies That Evaluate Complex Interventions? *Plos Medicine*. 2009; 6(8).

Shewhart WA. *Statistical method from the viewpoint of quality control*. Washington D.C.: Graduate School of the Department of Agriculture; 1939.

Shiell A, Hawe P, Gold L. Complex interventions or complex systems? Implications for health economic evaluation. *BMJ*. 2008; 336(7656):1281-3.

Shirley PJ. The Safer Patients Initiative: the UK experience of attempting to improve safe clinical care. *Med J Aust*. 2008; 189(7):414.

Shortell SM, Jones RH, Rademaker AW, Gillies RR, Dranove DS, Hughes EF, Budetti PP, Reynolds KS, Huang CF. Assessing the impact of total quality management and organizational culture on multiple outcomes of care for coronary artery bypass graft surgery patients. *Med Care*. 2000; 38(2):207-17.

Smith GB, Prytherch DR, Schmidt P, Featherstone PI, Knight D, Clements G, Mohammed MA. Hospital-wide physiological surveillance-a new approach to the early identification and management of the sick patient. *Resuscitation*. 2006; 71(1):19-28.

Solberg LI, Mosser G, McDonald S. The three faces of performance measurement: improvement, accountability, and research. *Jt Comm J Qual Improv*. 1997; 23(3):135-47.

Speroff T and O'Connor GT. Study designs for PDSA quality improvement research. *Qual Manag Health Care*. 2004; 13(1):17-32.

Stevens DP, Bowen JL, Johnson JK, Woods DM, Provost LP, Holman HR, Sixta CS, Wagner EH. A multi-institutional quality improvement initiative to transform education for chronic illness care in resident continuity practices. *J Gen Intern Med*. 2010; 25 Suppl 4:S574-S580.

Stone S, Slade R, Fuller C, Charlett A, Cookson B, Teare L, Jeanes A, Cooper B, Roberts J, Duckworth G, Hayward A, McAteer J, Michie S. Early communication: does a national campaign to improve hand hygiene in the NHS work? Initial English and Welsh experience from the NOSEC study (National Observational Study to Evaluate the CleanYourHandsCampaign). *J Hosp Infect*. 2007; 66(3):293-6.

Stulberg JJ, Delaney CP, Neuhauser DV, Aron DC, Fu P, Koroukian SM. Adherence to surgical care improvement project measures and the association with postoperative infections. *JAMA*. 2010; 303(24):2479-85.

Suchman MC. Managing Legitimacy - Strategic and Institutional Approaches. *Academy of Management Review*. 1995; 20(3):571-610.

Sullivan H and Stewart M. Who Owns the Theory of Change? *Evaluation*. 2006:179-99.

Taxis K and Barber N. Causes of intravenous medication errors: an ethnographic study. *Qual Saf Health Care*. 2003; 12(5):343-7.

Tennant G. *Six sigma: spc and tqm in manufacturing and services*. Gower Publishing Ltd; 2001.

The Joint Commission. The Joint Commission announces the 2008 National Patient Safety Goals and Requirements. *Jt Comm Perspect*. 2007; 27(7):1, 9-1,22.

The Royal Pharmaceutical Society of Great Britain and British Medical Association. *British National Formulary: v. 53*. Pharmaceutical Press; 2007a.

The Royal Pharmaceutical Society of Great Britain and British Medical Association. *British National Formulary: v. 54*. BMJ Publishing Group; 2007b.

The Royal Pharmaceutical Society of Great Britain and British Medical Association. *British National Formulary: v. 56*. Pharmaceutical Press; 2008.

Thomas EJ and Petersen LA. Measuring errors and adverse events in health care. *J Gen Intern Med*. 2003; 18(1):61-7.

Thomas EJ, Sexton JB, Neilands TB, Frankel A, Helmreich RL. The effect of executive walk rounds on nurse safety climate attitudes: a randomized trial of clinical units[ISRCTN85147255] [corrected].[Erratum appears in BMC Health Serv Res. 2005 Jun 8;5(1):4]. *BMC Health Services Research*. 2005; 5(1):28.

Thompson AM, Ashraf Z, Burton H, Stonebridge PA. Mapping changes in surgical mortality over 9 years by peer review audit. *Br J Surg*. 2005; 92(11):1449-52.

Thomson RG, Eccles MP, Steen IN, Greenaway J, Stobbart L, Murtagh MJ, May CR. A patient decision aid to support shared decision-making on anti-thrombotic treatment of patients with atrial fibrillation: randomised controlled trial. *Qual Saf Health Care*. 2007; 16(3):216-23.

Turner RM, Thompson SG, Spiegelhalter DJ. Prior distributions for the intracluster correlation coefficient, based on multiple previous estimates, and their application in cluster randomized trials. *Clin Trials*. 2005; 2(2):108-18.

Ukoumunne OC, Gulliford MC, Chinn S, Sterne JA, Burney PG. Methods for evaluating area-wide and organisation-based interventions in health and health care: a systematic review. *Health Technol Assess*. 1999; 3(5):iii-92.

Unahalekhaka A, Jamulitrat S, Chongsuvivatwong V, Ovretveit J. Using a collaborative to reduce ventilator-associated pneumonia in Thailand. *Joint Commission Journal on Quality & Patient Safety*. 2007; 33(7):387-94.

van Kasteren ME, Mannien J, Kullberg BJ, de Boer AS, Nagelkerke NJ, Ridderhof M, Wille JC, Gyssens IC. Quality improvement of surgical prophylaxis in Dutch hospitals: evaluation of a multi-site intervention by time series analysis. *J Antimicrob Chemother*. 2005; 56(6):1094-102.

Varkey P, Reller MK, Resar RK. Basics of quality improvement in health care. *Mayo Clinic Proceedings*. 2007; 82(6):735-9.

Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ*. 2001; 322(7285):517-9.

Vincent C. Incident reporting and patient safety. *BMJ*. 2007; 334(7584):51.

Vishwanath A, Brodsky L, Shaha S, Leonard M, Cimino M. Patterns and changes in prescriber attitudes toward PDA prescription-assistive technology. *Int J Med Inform*. 2009; 78(5):330-9.

von PC and Aslaksen A. Improving the quality of palliative care for ambulatory patients with lung cancer. *BMJ*. 2005; 330(7503):1309-13.

Wachter R and Pronovost P. The 100,000 lives campaign: a scientific and policy review. *Joint Commission Journal on Quality and Patient Safety*. 2006; 32(11):621.

Wagner EH, Glasgow RE, Davis C, Bonomi AE, Provost L, McCulloch D, Carver P, Sixta C. Quality improvement in chronic illness care: a collaborative approach. *Jt Comm J Qual Improv*. 2001; 27(2):63-80.

West MA, Guthrie JP, Dawson JF, Borrill CS, Carter MR. Reducing patient mortality in hospitals: The role of human resource management. *Journal of Organizational Behavior*. 2006; 27:983-1002.

Williams E and Talley R. The use of failure mode effect and criticality analysis in a medication error subcommittee. *Hosp Pharm*. 1994; 29(4):331-6, 339.

Wilson B, Thornton JG, Hewison J, Lilford RJ, Watt I, Braunholtz D, Robinson M. The Leeds University maternity audit project. *Int J Qual Health Care*. 2002; 14(3):175-81.

Wilson RM, Runciman WB, Gibberd RW, Harrison BT, Newby L, Hamilton JD. The Quality in Australian Health Care Study. *Med J Aust*. 1995; 163(9):458-71.

Wirtschafter DD, Danielsen BH, Main EK, Korst LM, Gregory KD, Wertz A, Stevenson DK, Gould JB, California Perinatal Quality Care Collaborative. Promoting antenatal steroid use for fetal maturation: results from the California Perinatal Quality Care Collaborative. *Journal of Pediatrics*. 2006; 148(5):606-12.

Woloshynowych M, Neale G, Vincent C. Case record review of adverse events: a new approach. *Qual Saf Health Care*. 2003; 12(6):411-5.

World Health Organization. Clean care is safer care. 2010. [cited 2010 Aug 15]. *Available at: www.who.int/gpsc/en*

Young EM, Commiskey ML, Wilson SJ. Translating evidence into practice to prevent central venous catheter-associated bloodstream infections: a systems-based intervention. *American Journal of Infection Control*. 2006; 34(8):503-6.

# APPENDICES

## Appendix A: Results of the literature review

| Reference | Reviewed from abstract | Single-centre (SC); Multicentre =MC | No contemporaneous controls | | | Contemporaneous controls | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Before & after | Time series | SPC | Before & after randomised controls | Before & after non-randomised controls | Cross-sectional randomised controls | Cross-sectional non-randomised controls |
| Abella BS, Edelson DP, Kim S, Retzer E, Myklebust H, Barry AM, O'Hearn N, Hoek TL, Becker LB. CPR quality improvement during in-hospital cardiac arrest using a real-time audiovisual feedback system. *Resuscitation*. 2007; 73(1):54-61. | | SC | x | | | | | | |
| Aghlmand S, Akbari F, Lameei A, Mohammad K, Small R, Arab M. Developing evidence-based maternity care in Iran: A quality improvement study. *BMC Pregnancy and Childbirth*. 2008; 8. | | SC | x | | | | | | |
| Agostini JV, Zhang Y, Inouye SK. Use of a computer-based reminder to improve sedative-hypnotic prescribing in older hospitalized patients. *Journal of the American Geriatrics Society*. 2007; 55(1):43-8. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Ahmar W, Quarin T, Ajani A, Kennedy M, Grigg L. Improvement in door-to-balloon times in management of acute ST-segment elevation myocardial infarction STEMI through the initiation of 'Code AMI'. *Internal Medicine Journal*. 2008; 38(9):714-8. | | SC | x | | | | | | |
| Aiken LH and Poghosyan L. Evaluation of "magnet journey to nursing excellence program" in Russia and Armenia. *Journal of Nursing Scholarship*. 2009; 41(2):166-74. | | SC | x | | | | | | |
| Allegranzi B, Sax H, Bengaly L, Richet H, Minta DK, Chraiti MN, Sokona FM, Gayet-Ageron A, Bonnabry P, Pittet D, World Health Organization. Successful implementation of the World Health Organization hand hygiene improvement strategy in a referral hospital in Mali, Africa. *Infection Control & Hospital Epidemiology*. 2010; 31(2):133-41. | x | SC | x | | | | | | |
| Andreoli A, Fancott C, Velji K, Baker GR, Solway S, Aimone E, Tardif G. Using SBAR to communicate falls risk and management in inter-professional rehabilitation teams. *Healthc Q*. 2010; 13 Spec No:94-101. | x | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Anwar uH, Saleem AF, Zaidi S, Haider SR. Experience of pediatric rapid response team in a tertiary care hospital in Pakistan. *Indian Journal of Pediatrics*. 2010; 77(3):273-6. | x | SC | x | | | | | | |
| Apisarnthanarak A, Thongphubeth K, Sirinvaravong S, Kitkangvan D, Yuekyen C, Warachan B, Warren DK, Fraser VJ. Effectiveness of multifaceted hospitalwide quality improvement programs featuring an intervention to remove unnecessary urinary catheters at a tertiary care center in Thailand. *Infection Control & Hospital Epidemiology*. 2007; 28(7):791-8. | | SC | x | | | | | | |
| Attena F, Di Palma MA, Esposito S, Galdo V, Gimigliano A, Parmeggiani C, Agozzino E. Quality improvement of medical records in a teaching hospital. *Journal of Preventive Medicine and Hygiene*. 2010; 51(2):53-6 | x | SC | x | | | | | | |
| Balaban RB, Weissman JS, Samuel PA, Woolhandler S. Redefining and redesigning hospital discharge to enhance patient care: a randomized controlled study. *J Gen Intern Med*. 2008; 23(8):1228-33. | | SC | | | | | x | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Balding C. Embedding organisational quality improvement through middle manager ownership. *International Journal of Health Care Quality Assurance*. 2005; 18(4):271-88. | | SC | x | | | | | | |
| Beck CA, Richard H, Tu JV, Pilote L. Administrative Data Feedback for Effective Cardiac Treatment: AFFECT, a cluster randomized trial. *JAMA*. 2005; 294(3):309-17. | | MC | | | | x | | | |
| Beckett CD and Kipnis G. Collaborative communication: integrating SBAR to improve quality/patient safety outcomes. *Journal for Healthcare Quality*. 2009; 31(5):19-28. | x | SC | | | x | | | | |
| Benn J, Burnett S, Parand A, Pinto A, Iskander S, Vincent C. Perceptions of the impact of a large-scale collaborative improvement programme: experience in the UK Safer Patients Initiative. *Journal of Evaluation in Clinical Practice*. 2009; 15(3):524-40. | | MC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Bernhard M, Becker TK, Nowe T, Mohorovicic M, Sikinger M, Brenner T, Richter GM, Radeleff B, Meeder PJ, Buchler MW, Bottiger BW, Martin E, Gries A. Introduction of a treatment algorithm can improve the early management of emergency patients in the resuscitation room. *Resuscitation*. 2007; 73(3):362-73. | | SC | x | | | | | | |
| Berry SA, Doll MC, McKinley KE, Casale AS, Bothe A, Jr. ProvenCare: quality improvement model for designing highly reliable care in cardiac surgery. *Qual Saf Health Care*. 2009; 18(5):360-8. | | MC | x | | | | | | |
| Bertsche T, Bertsche A, Krieg EM, Kunz N, Bergmann K, Hanke G, Hoppe-Tichy T, Ebinger F, Haefeli WE. Prospective pilot intervention study to prevent medication errors in drugs administered to children by mouth or gastric tube: a programme for nurses, physicians and parents. *Quality & Safety in Health Care*. 2010; 19(5):e26. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Bhalla R, Yongue BG, Currie BP, Greenberg MA, Myrie-Weir J, Defino M, Esses D, Menegus MA, McAllen SJ, Monrad ES, Galhotra S, Kalkut G. Improving primary percutaneous coronary intervention performance in an urban minority population using a quality improvement approach. *American Journal of Medical Quality*. 2010; 25(5):370-7. | | SC | x | | | | | | |
| Blankenship JC, Haldis TA, Wood GC, Skelding KA, Scott T, Menapace FJ. Rapid triage and transport of patients with ST-elevation myocardial infarction for percutaneous coronary intervention in a rural health system. *Am J Cardiol*. 2007; 100(6):944-8. | | SC | x | | | | | | |
| Bleakley A, Boyden J, Hobbs A, Walsh L, Allard J. Improving teamwork climate in operating theatres: The shift from multiprofessionalism to interprofessionalism. *Journal of Interprofessional Care*. 2006; 20(5):461-70. | | SC | | | | | x | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Blegen MA, Sehgal NL, Alldredge BK, Gearhart S, Auerbach AA, Wachter RM. Improving safety culture on adult medical units through multidisciplinary teamwork and communication interventions: the TOPS Project. *Qual Saf Health Care*. 2010; 19(4):346-50. | | MC | x | | | | | | |
| Bonello RS, Fletcher CE, Becker WK, Clutter KL, Arjes SL, Cook JJ, Petzel RA. An intensive care unit quality improvement collaborative in nine Department of Veterans Affairs hospitals: reducing ventilator-associated pneumonia and catheter-related bloodstream infection rates. *Jt Comm J Qual Patient Saf*. 2008; 34(11):639-45. | | MC | x | | | | | | |
| Bonner A, MacCulloch P, Gardner T, Chase CW. A student-led demonstration project on fall prevention in a long-term care facility. *Geriatric Nursing*. 2007; 28(5):312-8. | | SC | x | | | | | | |
| Bookbinder M, Blank AE, Arney E, Wollner D, Lesage P, McHugh M, Indelicato RA, Harding S, Barenboim A, Mirozyev T, Portenoy RK. Improving end-of-life care: development and pilot-test of a clinical pathway. *Journal of Pain & Symptom Management*. 2005; 29(6):529-43. | | MC | | | | | x | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Bourgeois FC, Linder J, Johnson SA, Co JPT, Fiskio J, Ferris TG. Impact of a computerized template on antibiotic prescribing for acute respiratory infections in children and adolescents. *Clinical Pediatrics*. 2010; 49(10):976-83. | | MC | | | | | | x | |
| Bowler WA, Bresnahan J, Bradfish A, Fernandez C. An integrated approach to methicillin-resistant Staphylococcus aureus control in a rural, regional-referral healthcare setting. *Infection Control & Hospital Epidemiology*. 2010; 31(3):269-75. | x | MC | x | | | | | | |
| Bowman SM, Sharar SR, Quan L. Impact of a statewide quality improvement initiative in improving the management of pediatric splenic injuries in washington state. *Journal of Trauma-Injury Infection & Critical Care*. 2008; 64(6):1478-83. | | MC | x | | | | | | |
| Boyd S, Aggarwal I, Davey P, Logan M, Nathwani D. Peripheral intravenous catheters: the road to quality improvement and safer patient care. *J Hosp Infect*. 2011; 77(1):37-41. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Brackbill ML, Kline VT, Sytsma CS, Call JT. Intervention to increase the proportion of acute myocardial infarction or coronary artery bypass graft patients receiving an order for aspirin at hospital discharge. *Journal of Managed Care Pharmacy*. 2010; 16(5):329-36. | | SC | x | | | | | | |
| Bradley E, Hartwig KA, Rowe LA, Cherlin EJ, Pashman J, Wong R, Dentry T, Wood WE, Abebe Y. Hospital quality improvement in Ethiopia: a partnership-mentoring model. *Int J Qual Health Care*. 2008; 20(6):392-9. | | MC | x | | | | | | |
| Brand CA, Amatya B, Gordon B, Tosti T, Gorelik A. Redesigning care for chronic conditions: improving hospital-based ambulatory care for people with osteoarthritis of the hip and knee. *Intern Med J*. 2010; 40(6):427-36. | | SC | x | | | | | | |
| Brush JE, Jr., Balakrishnan SA, Brough J, Hartman C, Hines G, Liverman DP, Parker JP, Rich J, Tindall N. Implementation of a continuous quality improvement program for percutaneous coronary intervention and cardiac surgery at a large community hospital. *American Heart Journal*. 2006; 152(2):379-85.) | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Brush JE, Jr., Rensing E, Song F, Cook S, Lynch J, Thacker L, Gurram S, Bonow RO, Brough J, Valentine CM. A statewide collaborative initiative to improve the quality of care for patients with acute myocardial infarction and heart failure. *Circulation*. 2009; 119(12):1609-15. | | MC | | | | | x | | |
| Buckmaster ND, Heazlewood V, Scott IA, Jones M, Haerer W, Hillier K. Using a clinical pathway and education to reduce inappropriate prescribing of enoxaparin in patients with acute coronary syndromes: a controlled study. *Internal Medicine Journal*. 2006; 36(1):12-8. | | SC | | | | | x | | |
| Burkitt KH, Mor MK, Jain R, Kruszewski MS, McCray EE, Moreland ME, Muder RR, Obrosky DS, Sevick MA, Wilson MA, Fine MJ. Toyota production system quality improvement initiative improves perioperative antibiotic therapy. *American Journal of Managed Care*. 2009; 15(9):633-42. | | MC | x | | | | | | |
| Calderon-Margalit R, Mor-Yosef S, Mayer M, Adler B, Shapira SC. An administrative intervention to improve the utilization of laboratory tests within a university hospital. *Int J Qual Health Care*. 2005; 17(3):243-8. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Calvino GS, Schwebel C, Vesin A, Remy J, Dessertaine G, Timsit JF. Interventions to decrease tube, line, and drain removals in intensive care units: the FRATER study. *Intensive Care Medicine*. 2009; 35(10):1772-6. | | SC | x | | | | | | |
| Campbell DA, Jr., Englesbe MJ, Kubus JJ, Phillips LR, Shanley CJ, Velanovich V, Lloyd LR, Hutton MC, Arneson WA, Share DA. Accelerating the pace of surgical quality improvement: the power of hospital collaboration. *Arch Surg*. 2010; 145(10):985-91. | x | MC | | | | | x | | |
| Campino A, Lopez-Herrera MC, Lopez-de-Heredia I, Soler A. Educational strategy to reduce medication errors in a neonatal intensive care unit. *Acta Paediatr*. 2009; 98(5):782-5. | | SC | x | | | | | | |
| Canal DF, Torbeck L, Djuricich AM. Practice-based learning and improvement: a curriculum in continuous quality improvement for surgery residents. *Archives of Surgery*. 482; 142(5):479-82. | | SC | x | | | | | | |
| Cantor MN, Lavarias V, Lam S, Mount L, Laskova V, Nakhamiyayev V, Bier Y, Paiusco D, Antonacci AC. Barriers to implementing a surgical beta-blocker protocol. *Joint Commission Journal on Quality & Patient Safety*. 2005; 31(11):640-8. | | MC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Capan K and Lynch B. A hospital fall assessment and intervention project. *Journal of Clinical Outcomes Management*. 2007; 14(3):155-60. | x | SC | x | | | | | | |
| Carlhed R, Bojestig M, Wallentin L, Lindstrom G, Peterson A, Aberg C, Lindahl B, QUICC study group. Improved adherence to Swedish national guidelines for acute myocardial infarction: the Quality Improvement in Coronary Care (QUICC) study. *American Heart Journal*. 2006; 152(6):1175-81. | | MC | | | | | x | | |
| Carling PC, Parry MF, Bruno-Murtha LA, Dick B. Improving environmental hygiene in 27 intensive care units to decrease multidrug-resistant bacterial transmission. *Critical Care Medicine*. 2010; 38(4):1054-9. | | MC | x | | | | | | |
| Caruso LB, Clough-Gorr KM, Silliman RA. Improving quality of care for urban older people with diabetes mellitus and cardiovascular disease.[Erratum appears in J Am Geriatr Soc. 2007 Oct;55(10):1697]. *Journal of the American Geriatrics Society*. 2007; 55(10):1656-62. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Castellanos-Ortega A, Suberviola B, Garcia-Astudillo LA, Holanda MS, Ortiz F, Llorca J, Delgado-Rodriguez M. Impact of the Surviving Sepsis Campaign protocols on hospital length of stay and mortality in septic shock patients: results of a three-year follow-up quasi-experimental study. *Critical Care Medicine*. 2010; 38(4):1036-43. R; | | SC | x | | | | | | |
| Chalmers JD, Singanayagam A, Akram AR, Choudhury G, Mandal P, Hill AT. Safety and efficacy of CURB65-guided antibiotic therapy in community-acquired pneumonia. *Journal of Antimicrobial Chemotherapy*. 2011; 66(2):416-23. | | SC | x | | | | | | |
| Chen KH, Chen LR, Su S. Applying root cause analysis to improve patient safety: decreasing falls in postpartum women. *Qual Saf Health Care*. 2010; 19(2):138-43. | | MC | | | | | x | | |
| Chern CH, How CK, Wang LM, Lee CH, Graff L. Decreasing clinically significant adverse events using feedback to emergency physicians of telephone follow-up outcomes. *Annals of Emergency Medicine*. 2005; 45(1):15-23. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Chiang WC, Chen WJ, Chen SY, Ko PC, Lin CH, Tsai MS, Chang WT, Chen SC, Tsan CY, Ma MH. Better adherence to the guidelines during cardiopulmonary resuscitation through the provision of audio-prompts. *Resuscitation*. 2005; 64(3):297-301. | | SC | x | | | | | | |
| Choma NN, Huang RL, Dittus RS, Burnham KE, Roumie CL. Quality improvement initiatives improve hypertension care among veterans. *Circulation Cardiovascular Quality & Outcomes*. 2009; 2(4):392-8. | x | MC | x | | | | | | |
| Cohen MM, Kimmel NL, Benage MK, Cox MJ, Sanders N, Spence D, Chen J. Medication safety program reduces adverse drug events in a community hospital. *Quality & Safety in Health Care*. 2005; 14(3):169-74. | | SC | x | | x | | | | |
| Cohn SL, Adekile A, Mahabir V. Improved use of thromboprophylaxis for deep vein thrombosis following an educational intervention. *Journal of Hospital Medicine (Online)*. 2006; 1(6):331-8. | x | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Cooper JB, Blum RH, Carroll JS, Dershwitz M, Feinstein DM, Gaba DM, Morey JC, Singla AK. Differences in safety climate among hospital anesthesia departments and the effect of a realistic simulation-based training program. *Anesthesia & Analgesia*. 2008; 106(2):574-84. | | MC | x | | | | | | |
| Criley JM, Keiner J, Boker JR, Criley SR, Warde CM. Innovative web-based multimedia curriculum improves cardiac examination competency of residents. *Journal of Hospital Medicine (Online)*. 2008; 3(2):124-33. | x | SC | x | | | | | | |
| Cunningham TR, Geller ES, Clarke SW. Impact of electronic prescribing in a hospital setting: a process-focused evaluation. *Int J Med Inform*. 2008; 77(8):546-54. | | SC | | | | | x | | |
| Curran E, Harper P, Loveday H, Gilmour H, Jones S, Benneyan J, Hood J, Pratt R. Results of a multicentre randomised controlled trial of statistical process control charts and structured diagnostic tools to reduce ward-acquired meticillin-resistant Staphylococcus aureus: the CHART Project. *Journal of Hospital Infection*. 2008; 70(2):127-35. | | MC | | | | x | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Curtis JR, Treece PD, Nielsen EL, Downey L, Shannon SE, Braungardt T, Owens D, Steinberg KP, Engelberg RA. Integrating palliative and critical care: evaluation of a quality-improvement intervention. *American Journal of Respiratory & Critical Care Medicine*. 2008; 178(3):269-75. | | SC | x | | | | | | |
| Daudelin DH, Sayah AJ, Kwong M, Restuccia MC, Porcaro WA, Ruthazer R, Goetz JD, Lane WM, Beshansky JR, Selker HP. Improving use of prehospital 12-lead ECG for early identification and treatment of acute coronary syndrome and ST-elevation myocardial infarction. *Circulation Cardiovascular Quality & Outcomes*. 2010; 3(3):316-23. | x | MC | x | | | | | | |
| Dedhia P, Kravet S, Bulger J, Hinson T, Sridharan A, Kolodner K, Wright S, Howell E. A quality improvement intervention to facilitate the transition of older adults from three hospitals back to their homes. *J Am Geriatr Soc*. 2009; 57(9):1540-6. | | MC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Diby M, Romand JA, Frick S, Heidegger CP, Walder B. Reducing pain in patients undergoing cardiac surgery after implementation of a quality improvement postoperative pain treatment program. *J Crit Care*. 2008; 23(3):359-71. | | SC | x | | | | | | |
| Doig GS, Simpson F, Finfer S, Delaney A, Davies AR, Mitchell I, Dobb G. Effect of evidence-based feeding guidelines on mortality of critically ill adults: a cluster randomized controlled trial. *JAMA*. 2008; 300(23):2731-41. | | MC | | | | | | x | |
| Donahoe L, McDonald E, Kho ME, Maclennan M, Stratford PW, Cook DJ. Increasing reliability of APACHE II scores in a medical-surgical intensive care unit: a quality improvement study. *American Journal of Critical Care*. 2009; 18(1):58-64. | | SC | | | | | x | | |
| Donelan K, Mailhot JR, Dutwin D, Barnicle K, Oo SA, Hobrecker K, Percac-Lima S, Chabner BA. Patient perspectives of clinical care and patient navigation in follow-up of abnormal mammography. *Journal of General Internal Medicine*. 2011; 26(2):116-22. | x | SC | | | | | | | x |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| DuPree E, Martin L, Anderson R, Kathuria N, Reich D, Porter C, Chassin MR. Improving patient satisfaction with pain management using Six Sigma tools. *Joint Commission Journal on Quality & Patient Safety*. 2009; 35(7):343-50. | | SC | x | | | | | | |
| Dyer D, Bouman B, Davey M, Ismond KP. An intervention program to reduce falls for adult in-patients following major lower limb amputation. *Healthcare Quarterly*. 2008; 11(3:Spec No.):Spec-21. | x | SC | x | | | | | | |
| Ellrodt G, Glasener R, Cadorette B, Kradel K, Bercury C, Ferrarin A, Jewell D, Frechette C, Seckler P, Reed J, Langou A, Surapaneni N, Multidisciplinary Rounds Team. Multidisciplinary rounds (MDR): an implementation system for sustained improvement in the American Heart Association's Get With The Guidelines program. *Critical Pathways in Cardiology: A Journal of Evidence-Based Medicine*. 2007; 6(3):106-16. | x | SC | x | | | | | | |
| Elnour AA, Ellahham NH, Al Qassas HI. Raising the awareness of inpatient nursing staff about medication errors. *Pharm World Sci*. 2008; 30(2):182-90. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Elwyn G, Owen D, Roberts L, Wareham K, Duane P, Allison M, Sykes A. Influencing referral practice using feedback of adherence to NICE guidelines: a quality improvement report for dyspepsia. *Quality & Safety in Health Care*. 2007; 16(1):67-70. | | MC | x | | | | | | |
| Ernst MM, Wooldridge JL, Conway E, Dressman K, Weiland J, Tucker K, Seid M. Using quality improvement science to implement a multidisciplinary behavioral intervention targeting pediatric inpatient airway clearance. *Journal of Pediatric Psychology*. 2010; 35(1):14-24. | | SC | x | | x | | | | |
| Ertl L and Christ F. Significant improvement of the quality of bystander first aid using an expert system with a mobile multimedia device. *Resuscitation*. 2007; 74(2):286-95. | | SC | | | | | | x | |
| Etchells E, Adhikari NK, Cheung C, Fowler R, Kiss A, Quan S, Sibbald W, Wong B. Real-time clinical alerting: effect of an automated paging system on response time to critical laboratory values--a randomised controlled trial. *Qual Saf Health Care*. 2010; 19(2):99-102. | | SC | | | | | | x | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Evans RS, Johnson KV, Flint VB, Kinder AT, Hawley WL, Lyon CR, Vawdrey DK, Thomsen GE. Unit-wide notification of ventilator disconnections. *AMIA Annual Symposium Proceedings/AMIA Symposium :951, 2005*. 2005:951. | | MC | | | | | x | | |
| Feinberg AN and Soyode O. A 1999 quality improvement initiative-reassessment in 2004. *Clinical Pediatrics*. 2005; 44(6):527-30. | | MC | | | | | x | | |
| Filardo G, Nicewander D, Herrin J, Edwards J, Galimbertti P, Tietze M, McBride S, Gunderson J, Collinsworth A, Haydar Z, Williams J, Ballard DJ. A hospital-randomized controlled trial of a formal quality improvement educational program in rural and small community Texas hospitals: one year results. *Int J Qual Health Care*. 2009; 21(4):225-32. | | MC | | | | x | | | |
| Fonda D, Cook J, Sandler V, Bailey M. Sustained reduction in serious fall-related injuries in older people in hospital. *Medical Journal of Australia*. 2006; 184(8):379-82. | | SC | x | | | | | | |
| Francois P, Vinck D, Labarere J, Reverdy T, Peyrin J-C. Assessment of an intervention to train teaching hospital care providers in quality management. *Qual Saf Health Care*. 2005; 14(4):234-9. | | SC | | | | | | | x |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Franklin BD, O'Grady K, Parr J, Walton I. Using the internet to deliver education on drug safety. *Quality & Safety in Health Care*. 2006; 15(5):329-33. | | SC | x | | | | | | |
| Freund M, Campbell E, Paul C, Sakrouge R, Lecathelinais C, Knight J, Wiggers J, Walsh RA, Jones T, Girgis A, Nagle A. Increasing hospital-wide delivery of smoking cessation care for nicotine-dependent in-patients: a multi-strategic intervention trial. *Addiction*. 2009; 104(5):839-49. | | MC | | | | | x | | |
| Fuhrmann L, Perner A, Klausen TW, Ostergaard D, Lippert A. The effect of multi-professional education on the recognition and outcome of patients at risk on general wards. *Resuscitation*. 2009; 80(12):1357-60. | | SC | x | | | | | | |
| Gallagher M, Oliver K, Hurwitz M. Improving the use of venous thromboembolism prophylaxis in an Australian teaching hospital. *Qual Saf Health Care*. 2009; 18(5):408-12. | | SC | | x | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Galvin JE, Kuntemeier B, Al-Hammadi N, Germino J, Murphy-White M, McGillick J. "Dementia-friendly hospitals: Care not crisis": An educational program designed to improve the care of the hospitalized patient with dementia. *Alzheimer Disease and Associated Disorders*. 2010; 24(4):372-9. | | MC | x | | | | | | |
| Gandara E, Ungar J, Lee J, Chan-Macrae M, O'Malley T, Schnipper JL. Discharge documentation of patients discharged to subacute facilities: a three-year quality improvement process across an integrated health care system. *Joint Commission Journal on Quality & Patient Safety*. 2010; 36(6):243-51. | | MC | x | | | | | | |
| Ginsburg L, Norton PG, Casebeer A, Lewis S. An educational intervention to enhance nurse leaders' perceptions of patient safety culture. *Health Services Research*. 2005; 40(4):997-1020. | | SC | | | | | x | | |
| Giordano R and Stichler JF. Improving Suicide Risk Assessment in the Emergency Department. *Journal of Emergency Nursing*. 2009; 35(1):22-6. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Gleeson SK, O'Mahony SP, Berman L, Carius M, Kulaga ME, Tkacs DJ, Van BP, Wang Y, Fine J. Reducing patient retriage and improving physician attitudes by provision of consensus workup guidelines in the emergency department. *Journal of Clinical Outcomes Management*. 2009; 16(7):309-14. | | SC | x | | | | | | |
| Goren JL, Beck SE, Mills BJ, Shtasel DL, Dufresne RL. Development and delivery of a quality improvement program to reduce antipsychotic polytherapy. *Journal of Managed Care Pharmacy*. 2010; 16(6):393-401. | | MC | x | | | | | | |
| Greenwood BC, Szumita PM, Lowry CM. Pharmacist-driven aminoglycoside quality improvement program. *Journal of Chemotherapy*. 2009; 21(1):42-5. | x | SC | x | | | | | | |
| Groome LJ, Esters FT, Carter SG. Bundling unbundled charges: a case study in an academic obstetrical practice. *Journal of Medical Practice Management*. 2009; 24(6):367-71. | x | SC | | | x | | | | |
| Haller G, Agoritsas T, Luthy C, Piguet V, Griesser A, Perneger T. Collaborative Quality Improvement to Manage Pain in Acute Care Hospitals. *Pain Medicine*. 2011; 12(1):138-47. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Haller G, Garnerin P, Morales MA, Pfister R, Berner M, Irion O, Clergue F, Kern C. Effect of crew resource management training in a multidisciplinary obstetrical setting. *Int J Qual Health Care*. 2008; 20(4):254-63. | | SC | x | | | | | | |
| Halverson AL, Andersson JL, Anderson K, Lombardo J, Park CS, Rademaker AW, Moorman DW. Surgical team training: the Northwestern Memorial Hospital experience. *Archives of Surgery*. 2009; 144(2):107-12. | | SC | x | | | | | | |
| Hansson E, Fridlund B, Hallstrom I. Effects of a quality improvement program in acute care evaluated by patients, nurses, and physicians. *Pain Management Nursing*. 2006; 7(3):93-108. | | SC | x | | | | | | |
| Hatfield E, Quinn C, Orme C, Thompson D, Kurth M. Real-time reporting drives the race toward zero. *Nursing Administration Quarterly*. 2005; 29(1):24-30. | x | SC | x | | | | | | |
| Hermayer KL, Cawley P, Arnold P, Sutton A, Crudup J, Kozlowski L, Hushion TV, Sheakley ML, Epps JA, Weil RP, Carter RE. Impact of improvement efforts on glycemic control and hypoglycemia at a university medical center. *Journal of Hospital Medicine (Online)*. 2009; 4(6):331-9. | x | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Hinchey JA, Shephard T, Tonn ST, Ruthazer R, Hermann RC, Selker HP, Kent DM. The Stroke Practice Improvement Network: a quasiexperimental trial of a multifaceted intervention to improve quality. *Journal of Stroke & Cerebrovascular Diseases*. 2010; 19(2):130-7. | | MC | | | | x | | | |
| Hixson ED, Nadzam DM, Katzan IL, Furlan AJ. Quality improvement for stroke management at the Cleveland Clinic Health System. *Joint Commission Journal on Quality & Patient Safety*. 2005; 31(8):447-54. | | MC | | x | | | | | |
| Holmes-Rovner M, Stommel M, Corser WD, Olomu A, Holtrop JS, Siddiqi A, Dunn SL. Does outpatient telephone coaching add to hospital quality improvement following hospitalization for acute coronary syndrome? *Journal of General Internal Medicine*. 2008; 23(9):1464-70. | | MC | | | | x | | | |
| Howell E, Bessman E, Kravet S, Kolodner K, Marshall R, Wright S. Active bed management by hospitalists and emergency department throughput. *Ann Intern Med*. 2008; 149(11):804-11. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Huang HT, Chuang YH, Chiang KF. Nurses' physical restraint knowledge, attitudes, and practices: the effectiveness of an in-service education program. *Journal of Nursing Research*. 2009; 17(4):241-8. | x | SC | | | | | | | x |
| Iyer SB, Schubert CJ, Schoettker PJ, Reeves SD. Use of quality-improvement methods to improve timeliness of analgesic delivery. *Pediatrics*. 2011; 127(1):e219-e225. | | SC | | | x | | | | |
| Jack BW, Chetty VK, Anthony D, Greenwald JL, Sanchez GM, Johnson AE, Forsythe SR, O'Donnell JK, Paasche-Orlow MK, Manasseh C, Martin S, Culpepper L. A reengineered hospital discharge program to decrease rehospitalization: a randomized trial. *Ann Intern Med*. 2009; 150(3):178-87. | | SC | | | | | | x | |
| Jallo N, Bray K, Padden MP, Levin D. A nurse-driven quality improvement program to improve perinatal outcomes. *J Perinat Neonatal Nurs*. 2009; 23(3):241-50. | x | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Jeffries HE, Mason W, Brewer M, Oakes KL, Munoz EI, Gornick W, Flowers LD, Mullen JE, Gilliam CH, Fustar S, Thurm CW, Logsdon T, Jarvis WR. Prevention of central venous catheter-associated bloodstream infections in pediatric intensive care units: a performance improvement collaborative. *Infection Control & Hospital Epidemiology*. 2009; 30(7):645-51. | | MC | | x | | | | | |
| Johnston SC, Sidney S, Hills NK, Grosvenor D, Klingman JG, Bernstein A, Levin E. Standardized discharge orders after stroke: results of the quality improvement in stroke prevention (QUISP) cluster randomized trial. *Annals of Neurology*. 2010; 67(5):579-89. | | MC | | | | x | | | |
| Jollis JG, Roettig ML, Aluko AO, Anstrom KJ, Applegate RJ, Babb JD, Berger PB, Bohle DJ, Fletcher SM, Garvey JL, Hathaway WR, Hoekstra JW, Kelly RV, Maddox WT, Jr., Shiber JR, Valeri FS, Watling BA, Wilson BH, Granger CB, Reperfusion of Acute Myocardial Infarction in North Carolina Emergency Departments (RACE) Investigators. Implementation of a statewide system for coronary reperfusion for ST-segment elevation myocardial infarction. *JAMA*. 2007; 298(20):2371-80. | | MC | | | | | x | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Jones DA, Mitra B, Barbetti J, Choate K, Leong T, Bellomo R. Increasing the use of an existing medical emergency team in a teaching hospital. *Anaesth Intensive Care*. 2006; 34(6):731-5. | x | SC | x | | | | | | x |
| Kable A, Gibberd R, Spigelman A. Measuring compliance with surgical antibiotic protocols: An intervention. *Clinical Governance*. 2008; 13(4):296-304. | | SC | x | | | | | | |
| Kahan NR, Kahan E, Waitman DA, Kitai E, Chintz DP. The tools of an evidence-based culture: implementing clinical-practice guidelines in an Israeli HMO. *Acad Med*. 2009; 84(9):1217-25. | | MC | | | | x | | | |
| Kalisch BJ, Curley M, Stefanov S. An intervention to enhance nursing staff teamwork and engagement. *Journal of Nursing Administration*. 2007; 37(2):77-84. | x | SC | x | | | | | | |
| Kao LS, Lew DF, Doyle PD, Carrick MM, Jordan VS, Thomas EJ, Lally KP. A tale of 2 hospitals: a staggered cohort study of targeted interventions to improve compliance with antibiotic prophylaxis guidelines. *Surgery*. 2010; 148(2):255-62. | | MC | | | | | x | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Karapinar-Carkit F, Borgsteede SD, Zoer J, Siegert C, van TM, Egberts AC, van den Bemt PM. The effect of the COACH program (Continuity Of Appropriate pharmacotherapy, patient Counselling and information transfer in Healthcare) on readmission rates in a multicultural population of internal medicine patients. *BMC Health Serv Res*. 2010; 10:39. | | SC | x | | | | | | |
| Kashefi C, Messer K, Barden R, Sexton C, Parsons JK. Incidence and prevention of iatrogenic urethral injuries. *Journal of Urology*. 2257; 179(6):2254-7. | | SC | x | | | | | | |
| Kim JC, Askew I, Muvhango L, Dwane N, Abramsky T, Jan S, Ntlemo E, Chege J, Watts C. Comprehensive care and HIV prophylaxis after sexual assault in rural South Africa: the Refentse intervention study. *BMJ*. 2009; 338:b515. | | SC | x | | | | | | |
| Kimmel S, Smith SL, Sabino JN, Gertner E, Dostal J, Greenberg MR. Tobacco screening multicomponent quality improvement network program: beyond education. *Academic Emergency Medicine*. 2009; 16(11):1186-92. | | MC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Klingner J, Moscovice I, Tupper J, Coburn A, Wakefield M. Implementing patient safety initiatives in rural hospitals. *Journal of Rural Health*. 2009; 25(4):352-7. | | SC | x | | | | | | |
| Koehler BE, Richter KM, Youngblood L, Cohen BA, Prengler ID, Cheng D, Masica AL. Reduction of 30-day postdischarge hospital readmission or emergency department (ED) visit rates in high-risk elderly medical patients through delivery of a targeted care bundle. *J Hosp Med*. 2009; 4(4):211-8. | x | SC | | | | | | x | |
| Koplan KE, Regan S, Goldszer RC, Schneider LI, Rigotti NA. A computerized aid to support smoking cessation treatment for hospital patients. *J Gen Intern Med*. 2008; 23(8):1214-7. | | SC | x | | | | | | |
| Kotagal M, Lee P, Habiyakare C, Dusabe R, Kanama P, Epino HM, Rich ML, Farmer PE. Improving quality in resource poor settings: observational study from rural Rwanda. *BMJ*. 2009; 339:b3488. | | SC | | | x | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Krimsky WS, Mroz IB, McIlwaine JK, Surgenor SD, Christian D, Corwin HL, Houston D, Robison C, Malayaman N. A model for increasing patient safety in the intensive care unit: increasing the implementation rates of proven safety measures. *Quality & Safety in Health Care*. 2009; 18(1):74-80. | | SC | | x | | | | | |
| Kritchevsky SB, Braun BI, Bush AJ, Bozikis MR, Kusek L, Burke JP, Wong ES, Jernigan J, Davis CC, Simmons B. The effect of a quality improvement collaborative to improve antimicrobial prophylaxis in surgical patients: a randomized trial. *Ann Intern Med*. 2008; 149(7):472-93. | | MC | | | | x | | | |
| Labresh KA, Fonarow GC, Smith SC, Jr., Bonow RO, Smaha LC, Tyler PA, Hong Y, Albright D, Ellrodt AG, Get With the Guidelines Steering Committee. Improved treatment of hospitalized coronary artery disease patients with the get with the guidelines program. *Critical Pathways in Cardiology: A Journal of Evidence-Based Medicine*. 2007; 6(3):98-105. | x | MC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Labresh KA, Reeves MJ, Frankel MR, Albright D, Schwamm LH. Hospital treatment of patients with ischemic stroke or transient ischemic attack using the "Get With The Guidelines" program. *Archives of Internal Medicine*. 2008; 168(4):411-7. | | MC | x | | | | | | |
| Lakshminarayan K, Borbas C, McLaughlin B, Morris NE, Vazquez G, Luepker RV, Anderson DC. A cluster-randomized trial to improve stroke care in hospitals. *Neurology*. 2010; 74(20):1634-42. | | MC | | | | x | | | |
| Landon BE, Hicks LS, O'Malley AJ, Lieu TA, Keegan T, McNeil BJ, Guadagnoli E. Improving the management of chronic disease at community health centers. *New England Journal of Medicine*. 2007; 356(9):921-34. | | MC | | | | | x | | |
| Larsen GY, Parker HB, Cash J, O'Connell M, Grant MC. Standard drug concentrations and smart-pump technology reduce continuous-medication-infusion errors in pediatric patients. *Pediatrics*. 2005; 116(1):e21-e25. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Leverstein-van Hall MA, Hopmans TE, van der Sprenkel JW, Blok HE, van der Mark WA, Hanlo PW, Bonten MJ. A bundle approach to reduce the incidence of external ventricular and lumbar drain-related infections. *Journal of Neurosurgery*. 2010; 112(2):345-53. | x | SC | x | | | | | | |
| Levy MM, Dellinger RP, Townsend SR, Linde-Zwirble WT, Marshall JC, Bion J, Schorr C, Artigas A, Ramsay G, Beale R, Parker MM, Gerlach H, Reinhart K, Silva E, Harvey M, Regan S, Angus DC, Surviving SC. The Surviving Sepsis Campaign: results of an international guideline-based performance improvement program targeting severe sepsis. *Critical Care Medicine*. 2010; 38(2):367-74. | | MC | | x | | | | | |
| Liebowitz LD and Blunt MC. Modification in prescribing practices for third-generation cephalosporins and ciprofloxacin is associated with a reduction in meticillin-resistant Staphylococcus aureus bacteraemia rate. *J Hosp Infect*. 2008; 69(4):328-36. | | SC | x | | | | | | |
| Lipke VL and Hyott AS. Reducing surgical site infections by bundling multiple risk reduction strategies and active surveillance. *AORN J*. 2010; 92(3):288-96. | | MC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Longhurst CA, Parast L, Sandborg CI, Widen E, Sullivan J, Hahn JS, Dawes CG, Sharek PJ. Decrease in hospital-wide mortality rate after implementation of a commercially sold computerized physician order entry system. *Pediatrics*. 2010; 126(1):14-21. | | SC | x | | | | | | |
| Mahant S, Peterson R, Campbell M, MacGregor DL, Friedman JN. Reducing inappropriate hospital use on a general pediatric inpatient unit. *Pediatrics*. 2008; 121(5):e1068-e1073. | | SC | x | | | | | | |
| Mannien J, van Kasteren ME, Nagelkerke NJ, Gyssens IC, Kullberg BJ, Wille JC, de Boer AS. Effect of optimized antibiotic prophylaxis on the incidence of surgical site infection. *Infect Control Hosp Epidemiol*. 2006; 27(12):1340-6. | x | MC | x | | | | | | |
| Maynard G, Lee J, Phillips G, Fink E, Renvall M. Improved inpatient use of basal insulin, reduced hypoglycemia, and improved glycemic control: effect of structured subcutaneous insulin orders and an insulin management algorithm. *J Hosp Med*. 2009; 4(1):3-15. | x | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| McCahill LE, Ahern JW, Gruppi LA, Limanek J, Dion GA, Sussman JA, McCaffrey CB, Leary DB, Lesage MB, Single RM. Enhancing compliance with Medicare guidelines for surgical infection prevention: experience with a cross-disciplinary quality improvement team. *Archives of Surgery*. 2007; 142(4):355-61. | | SC | x | | | | | | |
| McCormick AJ, Curtis JR, Stowell-Weiss P, Toms C, Engelberg R. Improving social work in intensive care unit palliative care: results of a quality improvement intervention. *J Palliat Med*. 2010; 13(3):297-304. | | SC | x | | | | | | |
| McCoy AB, Waitman LR, Gadd CS, Danciu I, Smith JP, Lewis JB, Schildcrout JS, Peterson JF. A computerized provider order entry intervention for medication safety during acute kidney injury: a quality improvement report. *American Journal of Kidney Diseases*. 2010; 56(5):832-41. | x | SC | x | | | | | | |
| McCulloch P, Kreckler S, New S, Sheena Y, Handa A, Catchpole K. Effect of a "Lean" intervention to improve safety processes and outcomes on a surgical emergency unit. *BMJ*. 2010; 341(7781):1043-6. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| McKenna GS, Karthikesalingam A, Walsh SR, Tang TY, Quick CR. Prevention of venous thromboembolism: improving practice in surgical patients. *International Journal Of Surgery*. 2009; 7(1):50-3. | | SC | x | | | | | | |
| McMullin J, Cook D, Griffith L, McDonald E, Clarke F, Guyatt G, Gibson J, Crowther M. Minimizing errors of omission: behavioural reenforcement of heparin to avert venous emboli: the BEHAVE study. *Critical Care Medicine*. 2006; 34(3):694-9. | | SC | x | | | | | | |
| McPhail GL, Ednick MD, Fenchel MC, VanDyke R, Chima A, Amin RS, Seid M. Improving follow-up in hospitalised children. *Quality & Safety in Health Care*. 2010a; 19(5):e35. | | SC | | | | | | | x |
| McPhail GL, Weiland J, Acton JD, Ednick M, Chima A, VanDyke R, Fenchel MC, Amin RS, Seid M. Improving evidence-based care in cystic fibrosis through quality improvement. *Archives of Pediatrics & Adolescent Medicine*. 2010b; 164(10):957-60. | | SC | | | x | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Mercier CE, Barry SE, Paul K, Delaney TV, Horbar JD, Wasserman RC, Berry P, Shaw JS. Improving newborn preventive services at the birth hospitalization: a collaborative, hospital-based quality-improvement project. *Pediatrics*. 2007; 120(3):481-8. | | MC | x | | | | | | |
| Merle V, Moret L, Pidhorz L, Dujardin F, Gouin F, Josset V, Graveleau S, Petit J, Riou F, Lombrail P, Czernichow P. Does comparison of performance lead to better care? A pilot observational study in patients admitted for hip fracture in three French public hospitals. *Int J Qual Health Care*. 2009; 21(5):321-9. | | SC | x | | | | | | |
| Midlov P, Holmdahl L, Eriksson T, Bergkvist A, Ljungberg B, Widner H, Nerbrand C, Hoglund P. Medication report reduces number of medication errors when elderly patients are discharged from hospital. *Pharm World Sci*. 2008; 30(1):92-8. | | SC | x | | | | | | |
| Miller MR, Griswold M, Harris JM, Yenokyan G, Huskins WC, Moss M, Rice TB, Ridling D, Campbell D, Margolis P, Muething S, Brilli RJ. Decreasing PICU catheter-associated bloodstream infections: NACHRI's quality transformation efforts. *Pediatrics*. 2010; 125(2):206-13. | | MC | | x | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Minton J, Clayton J, Sandoe J, Mc GH, Wilcox M. Improving early management of bloodstream infection: a quality improvement project. *BMJ*. 2008; 336(7641):440-3. | | SC | x | | | | | | |
| Moffett BS, Parham AL, Caudilla CD, Mott AR, Gurwitch KD. Oral anticoagulation in a pediatric hospital: impact of a quality improvement initiative on warfarin management strategies. *Qual Saf Health Care*. 2006; 15(4):240-3. | | SC | x | | | | | | |
| Moscucci M, Rogers EK, Montoye C, Smith DE, Share D, O'Donnell M, Maxwell-Eward A, Meengs WL, De Franco AC, Patel K, McNamara R, McGinnity JG, Jani SM, Khanal S, Eagle KA. Association of a continuous quality improvement initiative with practice and outcome variations of contemporary percutaneous coronary interventions. *Circulation*. 2006; 113(6):814-22. | | MC | | | | | x | | |
| Mudge A, Denaro C, Scott I, Bennett C, Hickey A, Jones MA. The paradox of readmission: effect of a quality improvement program in hospitalized patients with heart failure. *Journal of Hospital Medicine (Online)*. 2010; 5(3):148-53. | x | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Nagle AL, Hensley MJ, Schofield MJ, Koschel AJ. A randomised controlled trial to evaluate the efficacy of a nurse-provided intervention for hospitalised smokers. *Australian & New Zealand Journal of Public Health*. 2005; 29(3):285-91. | | SC | | | | x | | | |
| Nallamothu BK, Krumholz HM, Peterson ED, Pan W, Bradley E, Stern AF, Masoudi FA, Janicke DM, Hernandez AF, Cannon CP, Fonarow GC, American Heart Association Get-With-The-Guidelines Investigators BAat. Door-to-balloon times in hospitals within the get-with-the-guidelines registry after initiation of the door-to-balloon (D2B) Alliance. *American Journal of Cardiology*. 2009; 103(8):1051-5. | | MC | | | | | x | | |
| Naughton BJ, Saltzman S, Ramadan F, Chadha N, Priore R, Mylotte JM. A multifactorial intervention to reduce prevalence of delirium and shorten hospital length of stay. *J Am Geriatr Soc*. 2005; 53(1):18-23. | x | SC | x | | | | | | |
| Naughton BJ, Singh R, Wisniewski AM, Singh G, Anderson DR. Improving quality of NSAID prescribing by internal medicine trainees with an educational intervention. *Teach Learn Med*. 2010; 22(4):287-92. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Nawaz H, Abbas A, Sarfraz A, Slade MD, Calvocoressi L, Wild DM, Tessier-Sherman B. A randomized clinical trial to compare the use of safety net enclosures with standard restraints in agitated hospitalized patients. *J Hosp Med*. 2007; 2(6):385-93. | x | SC | | | | | | x | |
| Needham DM, Korupolu R, Zanni JM, Pradhan P, Colantuoni E, Palmer JB, Brower RG, Fan E. Early physical medicine and rehabilitation for patients with acute respiratory failure: a quality improvement project. *Archives of Physical Medicine & Rehabilitation*. 2010; 91(4):536-42. | | SC | x | | | | | | |
| Nelson NC, Evans RS, Samore MH, Gardner RM. Detection and prevention of medication errors using real-time bedside nurse charting. *J Am Med Inform Assoc*. 2005; 12(4):390-7. | | SC | | | | | x | | |
| Nguyen HB, Corbett SW, Steele R, Banta J, Clark RT, Hayes SR, Edwards J, Cho TW, Wittlake WA. Implementation of a bundle of quality indicators for the early management of severe sepsis and septic shock is associated with decreased mortality. *Critical Care Medicine*. 2007; 35(4):1105-12. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Nielsen PE, Goldman MB, Mann S, Shapiro DE, Marcus RG, Pratt SD, Greenberg P, McNamee P, Salisbury M, Birnbach DJ, Gluck PA, Pearlman MD, King H, Tornberg DN, Sachs BP. Effects of teamwork training on adverse outcomes and process of care in labor and delivery: a randomized controlled trial. *Obstetrics & Gynecology*. 2007; 109(1):48-55. | | MC | | | | | | x | |
| Noschese M, Donihi AC, Koerbel G, Karslioglu E, Dinardo M, Curll M, Korytkowski MT. Effect of a diabetes order set on glycaemic management and control in the hospital. *Quality & Safety in Health Care*. 2008; 17(6):464-8. | | SC | | | | | | | x |
| O'Brien JM, Greenhouse PK, Schafer JJ, Wheeler CA, Titus A, Pontzer RE, O'Neill MM, Wolf D. Implementing and improving the efficiency of a methicillin-resistant Staphylococcus aureus active surveillance program using information technology. *American Journal of Infection Control*. 2008; 36(3 SUPPL.):S62-S66. | | SC | x | | | | | | |
| O'Leary KJ, Wayne DB, Haviley C, Slade ME, Lee J, Williams MV. Improving teamwork: Impact of structured interdisciplinary rounds on a medical teaching unit. *Journal of General Internal Medicine*. 2010; 25(8):826-32. | x | SC | | | | | | | x |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Paige JT, Aaron DL, Yang T, Howell DS, Chauvin SW. Improved operating room teamwork via SAFETY prep: a rural community hospital's experience. *World Journal of Surgery*. 2009; 33(6):1181-7. | | SC | x | | | | | | |
| Paton C, Barnes TR, Cavanagh MR, Taylor D, Lelliott P, POMH-UK project team. High-dose and combination antipsychotic prescribing in acute adult wards in the UK: the challenges posed by p.r.n. prescribing. *British Journal of Psychiatry*. 2008; 192(6):435-9. | | MC | x | | | | | | |
| Junod PN, Piguet V, Bovier PA. Long-term effectiveness of a multifaceted intervention on pain management in a walk-in clinic. *Qjm*. 2007; 100(4):225-32. | | SC | x | | | | | | |
| Phelan MP, Glauser J, Smith E, Martin C, Schrump S, Mahone P, Peacock WF. Improving emergency department door-to-electrocardiogram time in ST segment elevation myocardial infarction. *Critical Pathways in Cardiology: A Journal of Evidence-Based Medicine*. 2009; 8(3):119-21. | x | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Polkinghorne KR, Seneviratne M, Kerr PG. Effect of a vascular access nurse coordinator to reduce central venous catheter use in incident hemodialysis patients: a quality improvement report. *American Journal of Kidney Diseases*. 2009; 53(1):99-106. | x | SC | x | | | | | | |
| Poon EG, Cina JL, Churchill W, Patel N, Featherstone E, Rothschild JM, Keohane CA, Whittemore AD, Bates DW, Gandhi TK. Medication dispensing errors and potential adverse drug events before and after implementing bar code technology in the pharmacy. *Ann Intern Med*. 2006; 145(6):426-34. | | SC | x | | | | | | |
| Poon EG, Keohane CA, Yoon CS, Ditmore M, Bane A, Levtzion-Korach O, Moniz T, Rothschild JM, Kachalia AB, Hayes J, Churchill WW, Lipsitz S, Whittemore AD, Bates DW, Gandhi TK. Effect of bar-code technology on the safety of medication administration. *N Engl J Med*. 2010; 362(18):1698-707. | | SC | | | | | x | | |
| Porat N, Bitan Y, Shefi D, Donchin Y, Rozenbaum H. Use of colour-coded labels for intravenous high-risk medications and lines to improve patient safety. *Qual Saf Health Care*. 2009; 18(6):505-9. | | SC | | | | | | x | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Power M, Wigglesworth N, Donaldson E, Chadwick P, Gillibrand S, Goldmann D. Reducing Clostridium difficile infection in acute care by using an improvement collaborative. *BMJ*. 2010; 341:c3359. | | MC | | | | | x | | |
| Prabhakaran D, Jeemon P, Mohanan PP, Govindan U, Geevar Z, Chaturvedi V, Reddy KS. Management of acute coronary syndromes in secondary care settings in Kerala: impact of a quality improvement programme. *National Medical Journal of India*. 2008; 21(3):107-11. | x | MC | x | | | | | | |
| Pronovost PJ, Berenholtz SM, Goeschel C, Thom I, Watson SR, Holzmueller CG, Lyon JS, Lubomski LH, Thompson DA, Needham D, Hyzy R, Welsh R, Roth G, Bander J, Morlock L, Sexton JB. Improving patient safety in intensive care units in Michigan. *Journal of Critical Care*. 2008; 23(2):207-21. | | MC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Pronovost PJ, Goeschel CA, Colantuoni E, Watson S, Lubomski LH, Berenholtz SM, Thompson DA, Sinopoli DJ, Cosgrove S, Sexton JB, Marsteller JA, Hyzy RC, Welsh R, Posa P, Schumacher K, Needham D. Sustaining reductions in catheter related bloodstream infections in Michigan intensive care units: observational study. *BMJ*. 2010; 340:c309. doi: 10.1136/bmj.c309.:c309. | | MC | | x | | | | | |
| Pulver LK, Tett SE, Coombes J. The Queensland experience of participation in a national drug use evaluation project, Community-Acquired Pneumonia Towards Improving Outcomes Nationally (CAPTION). *BMC Pulmonary Medicine*. 2009; 9:38. | | SC | x | | | | | | |
| Pun BT, Gordon SM, Peterson JF, Shintani AK, Jackson JC, Foss J, Harding SD, Bernard GR, Dittus RS, Ely EW. Large-scale implementation of sedation and delirium monitoring in the intensive care unit: a report from two medical centers. *Crit Care Med*. 2005; 33(6):1199-205. | | MC | x | | | | | | |
| Quillen K and Murphy K. Quality improvement to decrease specimen mislabeling in transfusion medicine. *Arch Pathol Lab Med*. 2006; 130(8):1196-8. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Rana R, Afessa B, Keegan MT, Whalen FX, Jr., Nuttall GA, Evenson LK, Peters SG, Winters JL, Hubmayr RD, Moore SB, Gajic O, Transfusion in the ICU Interest Group. Evidence-based red cell transfusion in the critically ill: quality improvement using computerized physician order entry. *Critical Care Medicine*. 2006; 34(7):1892-7. | | SC | x | | | | | | |
| Rasmussen HH, Kondrup J, Staun M, Ladefoged K, Lindorff K, Jorgensen LM, Jakobsen J, Kristensen H, Wengler A. A method for implementation of nutritional therapy in hospitals. *Clinical Nutrition*. 2006; 25(3):515-23. | | SC | x | | | | | | |
| Ravikumar TS, Sharma C, Marini C, Steele GD, Jr., Ritter G, Barrera R, Kim M, Safyer SM, Vandervoort K, De GM, Baker L, Levi P, Pierdon S, Horgan M, Maynor K, Maloney G, Wojtowicz M, Nelson K. A validated value-based model to improve hospital-wide perioperative outcomes: adaptability to combined medical/surgical inpatient cohorts. *Annals of Surgery*. 496; 252(3):486-96. | | MC | | | | | x | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Rezaee ME, Conley SM, Anderson TA, Brown JR, Yanofsky NN, Niles NW. Primary percutaneous coronary intervention for patients presenting with ST-elevation myocardial infarction: process improvements in rural prehospital care delivered by emergency medical services. *Progress in Cardiovascular Diseases*. 2010; 53(3):210-8. | | SC | x | | | | | | |
| Richardson J, Murray D, House CK, Lowenkopf T. Successful implementation of the National Institutes of Health Stroke Scale on a stroke/neurovascular unit. *Journal of Neuroscience Nursing*. 2006; 38(4:Suppl):Suppl-15. | x | SC | x | | | | | | |
| Riggio JM, Sorokin R, Moxey ED, Mather P, Gould S, Kane GC. Effectiveness of a clinical-decision-support system in improving compliance with cardiac-care quality measures and supporting resident training. *Academic Medicine*. 2009; 84(12):1719-26. | | SC | | x | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Rihal CS, Kamath CC, Holmes DR, Jr., Reller MK, Anderson SS, McMurtry EK, Long KH. Economic and clinical outcomes of a physician-led continuous quality improvement intervention in the delivery of percutaneous coronary intervention. *American Journal of Managed Care*. 2006; 12(8):445-52. | | MC | x | | | | | | |
| Roark SV. Intimate partner violence: screening and intervention in the health care setting. *Journal of Continuing Education in Nursing*. 496; 41(11):490-5. | x | SC | x | | | | | | |
| Robb E, Jarman B, Suntharalingam G, Higgens C, Tennant R, Elcock K. Using care bundles to reduce in-hospital mortality: quantitative survey. *BMJ*. 2010; 340:c1234. | | SC | | x | | | | | |
| Roberts CM, Stone RA, Buckingham RJ, Pursey NA, Harrison BD, Lowe D, Potter JM. A randomised trial of peer review: the UK National Chronic Obstructive Pulmonary Disease Resources and Outcomes Project. *Clinical Medicine*. 2010; 10(3):223-7. | | MC | | | | x | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Romagnuolo J, Flemons WW, Perkins L, Lutz L, Jamieson PC, Hiscock CA, Foley L, Meddings JB. Post-endoscopy checklist reduces length of stay for non-variceal upper gastrointestinal bleeding. *Int J Qual Health Care*. 2005; 17(3):249-54. | | SC | x | | | | | | |
| Rubin FH, Williams JT, Lescisin DA, Mook WJ, Hassan S, Inouye SK. Replicating the Hospital Elder Life Program in a community hospital and demonstrating effectiveness using quality improvement methodology. *Journal of the American Geriatrics Society*. 2006; 54(6):969-74. | | SC | x | | | | | | |
| Rushton CH, Reder E, Hall B, Comello K, Sellers DE, Hutton N. Interdisciplinary interventions to improve pediatric palliative care and reduce health care professional suffering. *Journal of Palliative Medicine*. 2006; 9(4):922-33. | x | SC | x | | | | | | |
| Scales DC, Dainty K, Hales B, Pinto R, Fowler RA, Adhikari NK, Zwarenstein M. A multifaceted intervention for quality improvement in a network of intensive care units: a cluster randomized trial. *JAMA*. 2011; 305(4):363-72. | | MC | | | | x | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Schade CP, Esslinger E, Anderson D, Sun Y, Knowles B. Impact of a national campaign on hospital readmissions in home care patients. *Int J Qual Health Care*. 2009; 21(3):176-82. | | MC | | | | | x | | |
| Schneider J, Moromisato D, Zemetra B, Rizzi-Wagner L, Rivero N, Mason W, Imperial-Perez F, Ross L. Hand hygiene adherence is influenced by the behavior of role models. *Pediatric Critical Care Medicine*. 2009; 10(3):360-3. | x | SC | x | | | | | | |
| Schnipper JL, Hamann C, Ndumele CD, Liang CL, Carty MG, Karson AS, Bhan I, Coley CM, Poon E, Turchin A, Labonville SA, Diedrichsen EK, Lipsitz S, Broverman CA, McCarthy P, Gandhi TK. Effect of an electronic medication reconciliation application and process redesign on potential adverse drug events: a cluster-randomized trial. *Archives of Internal Medicine*. 2009a; 169(8):771-80. | x | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Schnipper JL, Ndumele CD, Liang CL, Pendergrass ML. Effects of a subcutaneous insulin protocol, clinical education, and computerized order set on the quality of inpatient management of hyperglycemia: results of a clinical trial. *Journal of Hospital Medicine (Online)*. 2009b; 4(1):16-27. | | MC | | | | | | x | |
| Schonlau M, Mangione-Smith R, Chan KS, Keesey J, Rosen M, Louis TA, Wu SY, Keeler E. Evaluation of a quality improvement collaborative in asthma care: does it improve processes and outcomes of care? *Ann Fam Med*. 2005; 3(3):200-8. | | MC | | | | | x | | |
| Schwamm LH, Fonarow GC, Reeves MJ, Pan W, Frankel MR, Smith EE, Ellrodt G, Cannon CP, Liang L, Peterson E, Labresh KA. Get With the Guidelines-Stroke is associated with sustained improvement in care for patients hospitalized with acute stroke or transient ischemic attack. *Circulation*. 2009; 119(1):107-15. | | MC | x | | | | | | |
| Scott LD, Hofmeister N, Rogness N, Rogers AE. An interventional approach for patient and nurse safety: a fatigue countermeasures feasibility study. *Nurs Res*. 2010; 59(4):250-8. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Serisier DJ and Bowler SD. Effect of a simple educational intervention on the hospital management of community-acquired pneumonia. *Respirology*. 2007; 12(3):389-93. | | SC | x | | | | | | |
| Sharek PJ, McClead RE, Jr., Taketomo C, Luria JW, Takata GS, Walti B, Tanski M, Nelson C, Logsdon TR, Thurm C, Federico F. An intervention to decrease narcotic-related adverse drug events in children's hospitals. *Pediatrics*. 2008; 122(4):e861-e866. | | MC | x | | | | | | |
| Sharek PJ, Parast LM, Leong K, Coombs J, Earnest K, Sullivan J, Frankel LR, Roth SJ. Effect of a rapid response team on hospital-wide mortality and code rates outside the ICU in a Children's Hospital. *JAMA*. 2007; 298(19):2267-74. | | SC | x | | | | | | |
| Shavit I, Keidan I, Hoffmann Y, Mishuk L, Rubin O, Ziv A, Steiner IP. Enhancing patient safety during pediatric sedation: the impact of simulation-based training of nonanesthesiologists. *Archives of Pediatrics & Adolescent Medicine*. 2007; 161(8):740-3. | | SC | | | | | | | x |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Siminoff LA and Traino HM. Improving donation outcomes: hospital development and the Rapid Assessment of Hospital Procurement Barriers in Donation. *Progress in Transplantation*. 2009; 19(2):180-7. | x | MC | x | | | | | | |
| Simunovic M, Coates A, Goldsmith CH, Thabane L, Reeson D, Smith A, McLeod RS, DeNardi F, Whelan TJ, Levine MN. The cluster-randomized Quality Initiative in Rectal Cancer trial: evaluating a quality-improvement strategy in surgery. *CMAJ Canadian Medical Association Journal*. 2010; 182(12):1301-6. | | MC | | | | | | x | |
| Skledar SJ, McKaveney TP, Sokos DR, Ervin KA, Coldren M, Hynicka L, Lavsa S, Reynolds R, Beauchamp R. Role of student pharmacist interns in hospital-based standing orders pneumococcal vaccination program. *Journal of the American Pharmacists Association: JAPhA*. 2007; 47(3):404-9. | x | SC | x | | | | | | |
| Snijders C, Kollen BJ, van Lingen RA, Fetter WP, Molendijk H. Which aspects of safety culture predict incident reporting behavior in neonatal intensive care units? A multilevel analysis. *Crit Care Med*. 2009; 37(1):61-7. | | MC | | | | | x | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Snow V, Reynolds CE, Bennett L, Weiss KB, Snooks Q, Qaseem A. Closing the gap-cardiovascular risk and primary prevention: results from the American College of Physicians quality improvement program. *Am J Med Qual*. 2010; 25(4):261-7. | | MC | x | | | | | | |
| Stephan F, Sax H, Wachsmuth M, Hoffmeyer P, Clergue F, Pittet D. Reduction of urinary tract infection and antibiotic use after surgery: a controlled, prospective, before-after intervention study. *Clin Infect Dis*. 2006; 42(11):1544-51. | | SC | x | | | | | | |
| Stevens DP, Bowen JL, Johnson JK, Woods DM, Provost LP, Holman HR, Sixta CS, Wagner EH. A multi-institutional quality improvement initiative to transform education for chronic illness care in resident continuity practices. *J Gen Intern Med*. 2010; 25 Suppl 4:S574-S580. | x | MC | | | x | | | | |
| Stommel M, Olomu A, Holmes-Rovner M, Corser W, Gardiner JC. Changes in practice patterns affecting in-hospital and post-discharge survival among ACS patients. *BMC Health Services Research*. 2006; 6. | | MC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Strunk LB, Maison AW, Steinke D. Impact of a pharmacist on medication reconciliation on patient admission to a veterans affairs medical center. *Hospital Pharmacy*. 2008; 43(8):643-9. | x | SC | x | | | | | | |
| Suojaranta-Ylinen RT, Roine RO, Vento AE, Niskanen MM, Salmenpera MT. Improved neurologic outcome after implementing evidence-based guidelines for cardiac surgery. *Journal of Cardiothoracic & Vascular Anesthesia*. 2007; 21(4):529-34. | | SC | x | | | | | | |
| Thomas EJ, Sexton JB, Neilands TB, Frankel A, Helmreich RL. The effect of executive walk rounds on nurse safety climate attitudes: a randomized trial of clinical units[ISRCTN85147255] [corrected].[Erratum appears in BMC Health Serv Res. 2005 Jun 8;5(1):4]. *BMC Health Services Research*. 2005; 5(1):28. | | MC | | | | x | | | |
| Ting HH, Rihal CS, Gersh BJ, Haro LH, Bjerke CM, Lennon RJ, Lim CC, Bresnahan JF, Jaffe AS, Holmes DR, Bell MR. Regional systems of care to optimize timeliness of reperfusion therapy for ST-elevation myocardial infarction: the Mayo Clinic STEMI Protocol. *Circulation*. 2007; 116(7):729-36. | | | | | | | | | x |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Titler MG, Herr K, Brooks JM, Xie XJ, Ardery G, Schilling ML, Marsh JL, Everett LQ, Clarke WR. Translating research into practice intervention improves management of acute pain in older hip fracture patients. *Health Serv Res*. 2009; 44(1):264-87. | | MC | | | | x | | | |
| Tragulpiankit P, Kaojarern S, Pummangura C, Wananukul W, Luscombe D. Pharmacist participation in the reduction of preventable adverse drug reactions in Thai hospitalised patients. *Journal of Applied Therapeutic Research*. 2009; 7(1):3-10. | x | SC | x | | | | | | |
| Tschopp JM, Frey JG, Janssens JP, Burrus C, Garrone S, Pernet R, Imhof K, Besse F, Marty S, Rosset C, Assal JP. Asthma outpatient education by multiple implementation strategy. Outcome of a programme using a personal notebook. *Respir Med*. 2005; 99(3):355-62. | | MC | x | | | | | | |
| Tu JV, Donovan LR, Lee DS, Wang JT, Austin PC, Alter DA, Ko DT. Effectiveness of public report cards for improving the quality of cardiac care: the EFFECT study: a randomized trial. *JAMA*. 2009; 302(21):2330-7. | | MC | | | | x | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Unahalekhaka A, Jamulitrat S, Chongsuvivatwong V, Ovretveit J. Using a collaborative to reduce ventilator-associated pneumonia in Thailand. *Joint Commission Journal on Quality & Patient Safety*. 2007; 33(7):387-94. | | MC | | x | | | | | |
| van d, V and Leong T. Using the medical emergency team to supplement an existing adverse drug event reporting framework. *Journal of Pharmacy Practice and Research*. 2007; 37(3):197-9. | x | SC | x | | | | | | |
| van den Bemt PM, Cusell MB, Overbeeke PW, Trommelen M, van DD, Ophorst WR, Egberts AC. Quality improvement of oral medication administration in patients with enteral feeding tubes. *Qual Saf Health Care*. 2006; 15(1):44-7. | | MC | x | | | | | | |
| Van Eaton EG, McDonough K, Lober WB, Johnson EA, Pellegrini CA, Horvath KD. Safety of using a computerized rounding and sign-out system to reduce resident duty hours. *Academic Medicine*. 2010; 85(7):1189-95. | | MC | | | | x | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| van Gaal BG, Schoonhoven L, Vloet LC, Mintjes JA, Borm GF, Koopmans RT, van AT. The effect of the SAFE or SORRY? programme on patient safety knowledge of nurses in hospitals and nursing homes: a cluster randomised trial. *Int J Nurs Stud*. 2010; 47(9):1117-25. | | MC | | | | x | | | |
| van Hees BC, de RE, Wiltink EH, de Jongh BM, Tersmette M. Optimizing use of ciprofloxacin: a prospective intervention study. *J Antimicrob Chemother*. 2008; 61(1):210-3. | | SC | x | | | | | | |
| van Kasteren ME, Mannien J, Kullberg BJ, de Boer AS, Nagelkerke NJ, Ridderhof M, Wille JC, Gyssens IC. Quality improvement of surgical prophylaxis in Dutch hospitals: evaluation of a multi-site intervention by time series analysis. *J Antimicrob Chemother*. 2005; 56(6):1094-102. | | MC | | x | | | | | |
| Vishwanath A, Brodsky L, Shaha S, Leonard M, Cimino M. Patterns and changes in prescriber attitudes toward PDA prescription-assistive technology. *Int J Med Inform*. 2009; 78(5):330-9. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| von PC and Aslaksen A. Improving the quality of palliative care for ambulatory patients with lung cancer. *BMJ*. 2005; 330(7503):1309-13. | | SC | x | | x | | | | |
| Walsh M, Laptook A, Kazzi SN, Engle WA, Yao Q, Rasmussen M, Buchter S, Heldt G, Rhine W, Higgins R, Poole K. A cluster-randomized trial of benchmarking and multimodal quality improvement to improve rates of survival free of bronchopulmonary dysplasia for infants with birth weights of less than 1250 grams. *Pediatrics*. 2007; 119(5):876-90. | | MC | | | | x | | | |
| Ward KT, Bates-Jensen B, Eslami MS, Whiteman E, Dattoma L, Friedman JL, DeCastro MJ, Moore AA. Addressing delays in medication administration for patients transferred from the hospital to the nursing home: a pilot quality improvement project. *American Journal of Geriatric Pharmacotherapy*. 2008; 6(4):205-11. | | MC | | | | | | | x |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Weingart SN, Price J, Duncombe D, Connor M, Conley K, Conlin GJ, Sullivan AM, Powell M, Ponte PR, Bierer BE. Enhancing safety reporting in adult ambulatory oncology with a clinician champion: a practice innovation. *Journal of Nursing Care Quality*. 2009; 24(3):203-10. | x | SC | | | | | x | | |
| Wesorick DH, Grunawalt J, Kuhn L, Rogers MA, Gianchandani R. Effects of an educational program and a standardized insulin order form on glycemic outcomes in non-critically ill hospitalized patients. *J Hosp Med*. 2010; 5(8):438-45. | x | SC | | | | | x | | |
| Wessels H, de GA, Groenewegen G, Wynia K, de HM, Vos JB, Tjia P, Kruitwagen CL, Teunissen SC, Voest EE. Impact of integration of clinical and outpatient units on cancer patient satisfaction. *Int J Qual Health Care*. 2010; 22(5):358-64. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Williams JB, Delong ER, Peterson ED, Dokholyan RS, Ou FS, Ferguson TB, Jr., Society of Thoracic Surgeons and the National Cardiac Database. Secondary prevention after coronary artery bypass graft surgery: findings of a national randomized controlled trial and sustained society-led incorporation into practice. *Circulation*. 2011; 123(1):39-45. | | MC | | | | x | | | |
| Winans AR, Rudd KM, Triller D. Assessing anticoagulation knowledge in patients new to warfarin therapy. *Annals of Pharmacotherapy*. 2010; 44(7-8):1152-7. | x | SC | | | | | | | x |
| Wingard RL, Pupim LB, Krishnan M, Shintani A, Ikizler TA, Hakim RM. Early intervention improves mortality and hospitalization rates in incident hemodialysis patients: RightStart program. *Clinical Journal of The American Society of Nephrology: CJASN*. 2007; 2(6):1170-5. | | MC | | | | | x | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Wirtschafter DD, Danielsen BH, Main EK, Korst LM, Gregory KD, Wertz A, Stevenson DK, Gould JB, California Perinatal Quality Care Collaborative. Promoting antenatal steroid use for fetal maturation: results from the California Perinatal Quality Care Collaborative. *Journal of Pediatrics*. 2006; 148(5):606-12. | | MC | x | | | | | | |
| Wong R and Bradley EH. Developing patient registration and medical records management system in Ethiopia. *Int J Qual Health Care*. 2009; 21(4):253-8. | | SC | x | | | | | | |
| Woo J, Cheng JOY, Lee J, Lo R, Hui E, Lum CM, Or KH, Yeung F, Wong F, Mak B. Evaluation of a Continuous Quality Improvement Initiative for End-of-Life Care for Older Noncancer Patients. *Journal of the American Medical Directors Association*. 2011; 12(2):105-13. | x | SC | x | | | | | | |
| Wood SK, Campbell AK, Marden JD, Schmidtman L, Blundell GH, Sheerin NJ, Davidson PM. Inpatient care to community care: improving clinical handover in the private mental health setting. *Medical Journal of Australia*. 2009; 190(11:Suppl):Suppl-9. | | MC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Wren SM, Martin M, Yoon JK, Bech F. Postoperative pneumonia-prevention program for the inpatient surgical ward. *Journal of the American College of Surgeons*. 2010; 210(4):491-5. | | SC | x | | | | | | |
| Yilmaz M, Keegan MT, Iscimen R, Afessa B, Buck CF, Hubmayr RD, Gajic O. Toward the prevention of acute lung injury: protocol-guided limitation of large tidal volume ventilation and inappropriate transfusion. *Critical Care Medicine*. 1667; 35(7):1660-6. | | SC | x | | | | | | |
| Young EM, Commiskey ML, Wilson SJ. Translating evidence into practice to prevent central venous catheter-associated bloodstream infections: a systems-based intervention. *American Journal of Infection Control*. 2006; 34(8):503-6. | | SC | | x | | | | | |
| Zafar A, Habib F, Hadwani R, Ejaz M, Khowaja K, Khowaja R, Irfan S. Impact of infection control activities on the rate of needle stick injuries at a tertiary care hospital of Pakistan over a period of six years: An observational study. *BMC Infectious Diseases*. 2009; 9. | | SC | x | | | | | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Zolotor AJ, Randolph GD, Johnson JK, Wegner S, Edwards L, Powell C, Esporas MH. Effectiveness of a practice-based, multimodal quality improvement intervention for gastroenteritis within a Medicaid managed care network. *Pediatrics*. 2007; 120(3):e644-e650. | | MC | | | | | x | | |

# Appendix B: Strategic stakeholder interviews

## B.1 Introduction

The potential importance of the "blunt end" (the source of the resources and constraints that shape the working environment) in influencing the "sharp end" (where practitioners care for patients) is well recognised (Cook and Woods, 1994). In addition, it is important to identify the theory of change, i.e. a theory of how and why an initiative works (Mason and Barnes, 2007; Sullivan and Stewart, 2006) and knowing whether the intervention was implemented as intended (see 4.1.2.1). If strategic stakeholders (people in senior positions) do not understand and share the same theory of change then the implementation of a complex organisational intervention is likely to be suboptimal. How strategic stakeholders understood and responded to the **SPI** as a programme of change was investigated by means of strategic stakeholder interviews.

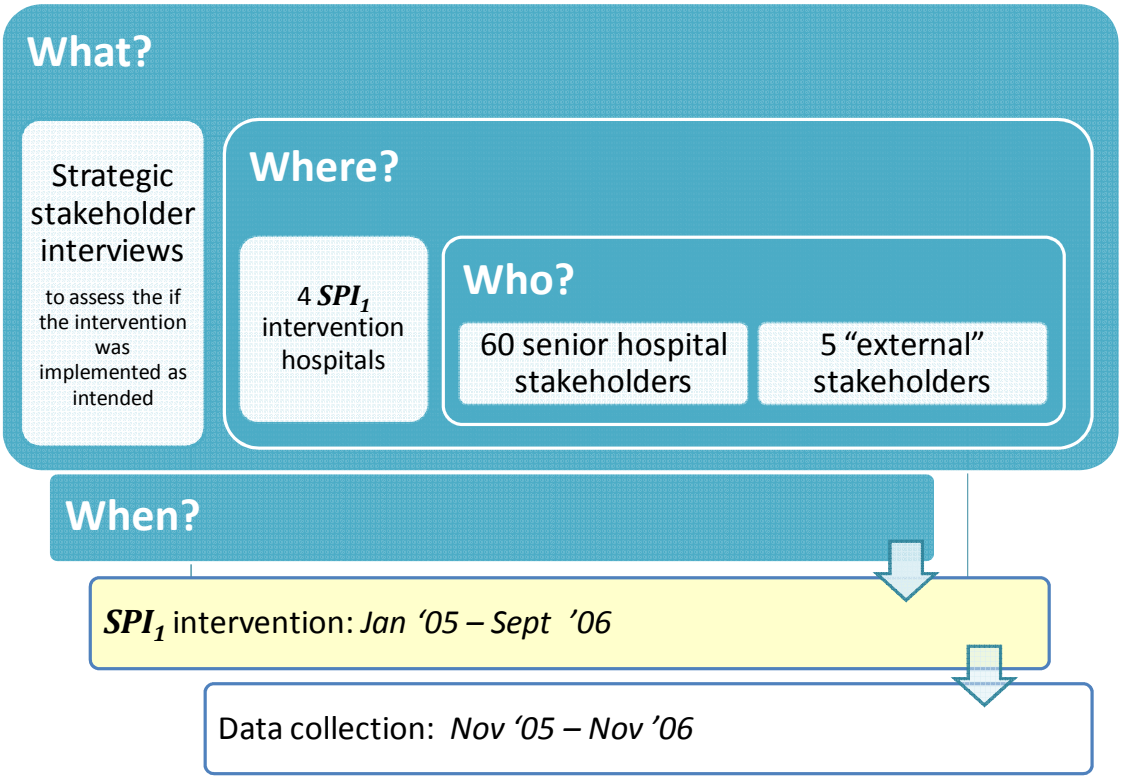
## B.2 Methods

This sub-study was undertaken in ***SPI1*** hospitals only and an overview can be found in Figure B.1.

Participants were recruited from all four ***SPI1*** hospital sites by asking hospital contacts involved in the **SPI** to identify people in particular roles, including chief executive, medical director, senior nurses, senior pharmacists, and others involved at a strategic level with implementation of ***SPI1***. Interviews were also conducted with a small number of "external"

stakeholders at the Health Foundation and IHI who had been involved in commissioning and designing the **SPI** (they were responsible for the commissioning and designing both ***SPI*$_1$** and ***SPI*$_2$** intervention phases).

**Figure B.1: An outline of the strategic stakeholder interview sub-study**




An experienced senior researcher conducted telephone interviews by using a semi-structured prompt guide. The interviews intended to identify:

- what participants saw as the aims of intervention;
- what theory of change participants were working with;
- barriers and facilitators they saw to change;
- participants' views on the acceptability and validity of the **SPI**.

Participants were allowed to describe the **SPI** in their own terms and responses were tape-recorded and transcribed verbatim by an external contractor. The transcripts were subject to two separate analyses. Firstly, the analysis of all transcripts was based on the constant comparative method (Glaser and Thomson, 1967). "Open" codes to describe each unit of meaning were initially generated. Through comparison across transcripts, the open codes were developed into higher-order thematic categories and subcategories to provide frameworks for coding, assisted by NVivo software.

Secondly, the responses given by transcripts by hospital (not commissioning) stakeholders were quantified. These transcripts were read by three independent reviewers who scored each individual's interview (on a 1-10 scale) for level of knowledge displayed and level of enthusiasm exhibited. Inter-observer agreement was tested using the ICC statistic. Pearson's correlation coefficient was used to measure association between enthusiasm and knowledge.

### B.2.1 Ethical considerations

It is good research practice to attribute quotes to the person that made them, for example, "participant 21, medical director, hospital 1", as it enables the reader to make a judgement on if the findings of the research are representative and based on the opinion of more than one person. However, as the **SPI** had a high public profile participants would be easily identifiable, so to preserve anonymity and not to contravene the ethical approval granted for this study, direct quotes have not been attributed.

## B.3 Results

Sixty strategic-level hospital stakeholder interviews were conducted over the period November 2005 to November 2006. Fourteen participants were in very senior strategic positions, including at each hospital the chief executive, the medical director, and the nursing director, and at least one clinical director. Other participants included 36 staff in senior managerial positions at the "front-line", including consultant grade doctors, nurses and pharmacists. The remaining six participants were in mid-level managerial positions (such as theatre managers and patient safety co-ordinators). Although some hospital participants had been directly involved in the competitive bidding process for their trust to participate in the **SPI** programme, most had not. Five "external" stakeholders, all at senior position in their organisation were also interviewed. Despite the interviews being taken over a year they suggested a high degree of congruence between participants.

### B.3.1 Aims of the SPI

All hospital participants spoke of the initiative as aiming to address patient safety problems, particularly by reducing the number of adverse incidents and making the hospital environment safer.

> *It's trying to improve patient care generally, but specifically to try to reduce the number of adverse events. I think with a bit of hard work and persistence to institute it, it could embed quite well in the NHS and it would be a healthy mechanism for maintaining standards and improving standards.* (Consultant)

> *If we can't make patients safe in hospital … then we're wasting our time. … So I strongly recommend it and I think it is transformational.* (Head of Nursing)

Seven participants described **SPI** solely in terms of cultural change and 11 solely in terms of changes to systems, but most (n=42) described **SPI** in terms of changes *both* to culture (such as increasing awareness or changing attitudes) and to systems, practices and processes, particularly in increasing consistency and monitoring and implementing evidence-based practice.

> *it is … about us looking at how we change the culture and the way that people think so that everybody realises this isn't an extra add-on to their day-job which probably at the moment some staff do perceive, but actually this is how we do business here and safety is our first and foremost, you know, ethos.* (Patient Safety Manager)

> *It's trying to look at the high risk areas and actually assessing whether our processes are reliable at the moment and if they're actually putting patients at increased risk, and trying to identify ways to improve those processes*. (Pharmacist)

Fifty-three hospital participants expressed positive attitudes towards the **SPI** during their interview, although 19 of these also made statements that indicated a muted positive or ambivalent attitude. Less positive attitudes included expressions of scepticism (n=10) or feeling apprehensive, overwhelmed or confused by the initiative (n=17). These more negative accounts tended to be descriptions of how the interviewees had felt when the initiative had first been introduced; many of those interviewed later in the study described having developed more positive attitudes over time. Four hospital participants described feeling only negatively about the initiative. Thus, for the most part, strategic stakeholders saw the **SPI** as a legitimate and promising response to the problems of patient safety.

### B.3.2 How will it work?

When asked to describe how the **SPI** would work, five hospital participants explicitly stated that they were unable to give an account of any specific element of the **SPI**, and a further two appeared to have very little knowledge of the initiative. However, most of the accounts from hospital stakeholders appeared to demonstrate a shared understanding of the main elements of the initiative. In many cases, hospital stakeholders' accounts demonstrated a high degree of symmetry with "external" stakeholders, as these extracts illustrate in their emphasis on the need for standardisation of practice as a way of improving safety:

> *[we] believe very strongly that by the application of our model of the rapid cycle and the reliability principles that we're going to have less variability and more reliable processes to ensure that the care is delivered the way that we would want it to be delivered.* (IHI, "external" stakeholder)

> *It's about standardisation. A lot of these things, over the years consultants have developed their own particular practices and of course that's a problem if you've got juniors going from one consultant to another, they get a bit confused in the end and so it doesn't happen. So it's about bringing everyone together and making an agreement about what the process is than then making sure it happens.* (Pharmacist)

An example of the shared understanding among strategic stakeholders could be found in hospital stakeholders' descriptions of the PDSA cycle, which was described as a critical element of the **SPI** in the "external" stakeholder accounts and in the learning sessions. Most (n=41) hospital participants offered a reasonable description of PDSA, with 22 explicitly using the term *PDSA*, although accounts varied considerably in their precision. For many, a key element of their theory of change relating to PDSA could be summarised as "seeing is believing", and an emphasis on how involving staff could lead to ownership and hence to

effective changes in practice. Again, there was considerable symmetry between "external" stakeholders and hospital stakeholders in their accounts.

> *it's based on the IHI's own kind of interpretation of … approaches to quality improvement which are based around what they call these PDSA cycles which is a [way] of getting people into making relatively small tests of change—doing the measurements, introducing the improvements, reviewing and then moving on and building that up over time to give everybody involved a sense of purpose and confidence.* (Health Foundation, "external" stakeholder)

> *so there's a kind of planning stage on how it's gonna work and then they start doing it in a small group of patients first one patient as far as I know and then looked at how it's going and then roll it out to three patients, five patients and then everybody.* (Health Professional in Haematology)

In addition to identifying the principal elements of the design of the **SPI** programme, participants often identified very specific strands or components of the initiative. These included leadership (n=37), interventions relating to general ward (n=35), intensive/critical care (n=34), peri-operative/theatres (n=40), medication/drug errors (n=45) and hospital-acquired infections (n=33). Participants also spoke about how the initiative worked in terms of practices that have been adopted, including SBAR (n=15), early warning and critical response system (n=29), safety briefings (n=15) and multidisciplinary teams (n=12). (See Chapter 2 for details of these elements of the intervention).

### B.3.3 Securing implementation

Participants' accounts highlighted factors that they believed would facilitate the implementation of the initiative. For hospital stakeholders, these included good leadership in the trust (n=43), motivation and commitment of staff (n=41), the existing culture or

structure of the trust (n=42), resources (n=18), taking a bottom-up approach and ensuring local ownership (n=15), good information and communication (n=22), education and training (n=9), sharing of experiences with others involved in **SPI** (n=9), and evidence that the initiative works in practice (n=14). Many (n=42) talked of factors specific to their organisation that were already in place as providing "fertile ground" for **SPI** to embed. Such factors included people being receptive to change, a good track record with clinical and risk management, good communication, high levels of awareness and an open culture.

> *Well I think it is more likely to work here. Because I think we have a culture of openness, freely admitting mistakes, good communication. We're regarded I suspect as a can-do organisation. … I think we're an organisation that readily accepts change.* (Medical Director)

> *I think that it's a very positive and supportive organisation … so it's not seen as something that's being brought in from the top. It's hopefully being seen as a genuine culture change.* (Clinical Director)

However, most (n=57) hospital stakeholders also described barriers to the initiative's successful implementation. For these stakeholders, success was seen as depending to a large extent on how far other stakeholders in the organisation—particularly those at the sharp end—could be mobilised around the initiative. Participants cited difficulties in changing attitudes and culture, and particularly in getting front-line staff to realise that their current practice was not necessarily safe practice; reconciling (new) standardised practices with clinical autonomy; problems in getting people to do things properly (due to lack of knowledge, education, engagement or time); lack of ownership; lack of leadership; difficulties communicating within multidisciplinary teams; and people not believing the evidence.

*you do have that element where you know individuals would think that they work safely. So you know kind of like, "Well I always make sure my patient is safe".* (Senior Sister)

Accounts stressed the need for the changes in practices and attitudes encouraged by the **SPI** to become "taken for granted", and reproduced routinely in organisational settings without contestation or resistance, but many strategic stakeholders feared that the **SPI** would not enjoy the legitimacy it needed among front-line practitioners. The most common view (raised by 52 hospital stakeholders) was that "people" issues were hugely important, and that although some people in the organisation were enthusiastic, other sharp-end and often powerful stakeholders (especially doctors) were "cynical", "wary", "too busy" or "resistant", although only seven participants felt that people in their organisation primarily felt negatively or sceptical towards the initiative.

*I think the reaction has been mixed. I think there's been a lot of people who have felt it's nothing to do with them and therefore haven't shown any interest. I think there's been a lot of people who certainly at the outset felt it was just yet another thing that the management had signed us up for. I think there were some people who thought, "Yeah actually this has got some useful things in it but some of it's not for us", and I think there were one or two zealots who thought it was the greatest thing that they'd ever heard.* (Consultant)

*Really just the staff are upset about having more work to do. I would say it would be the general feeling on the on the ground.* (Senior Nurse)

Specific barriers to implementation identified in strategic stakeholders' accounts included increased work and bureaucracy, lack of communication about the initiative, trying to take on too many changes at once, lack of preparation, unrealistic time scales and concerns about maintaining momentum over longer periods of time. The "complainability" of these issues was seen as being a threat to legitimacy, in part because of the costs they imposed.

*When you start to implement change and bring maybe an additional form that they have to fill in, they automatically, you get negative feedback … "We're filling in enough forms and we've enough to do as it is. We haven't time for this"* (Pharmacist)

These doubts about the organisational feasibility of implementing the initiative were joined by doubts about the scientific legitimacy of some aspects of the programme in many (n=32) hospital stakeholder accounts. Scientific doubts arose in relation to difficulties measuring change and reaching targets in that relate to changes in behaviour, lack of evidence for some of the interventions, changes that did not fit easily with current practice, problems with definitions leading to difficulty in measuring outcomes (particularly in relation to infections), the abstract nature of the initiative, and problems with translating and applying an programme developed in the USA to the UK context. These accounts often appeared to point to the existence of professional subcultures that did not share the same view of specific issues as the **SPI**.

*[There has been] a lack of knowledge or a lack of belief if you like in the evidence, and there have been a couple of points in the surgical site infection bundle where people have questioned the evidence that has been given forward for doing something, and in fact in one of those instances they, with regards to beta blockers there is such widespread disbelief in the evidence that the Safer Patient Initiative has watered that down now.* (Medical Director)

*…we're not sure what an infection is you know and I feel strongly … maybe it's from a scientific background as opposed to a management background but if we're counting apples and oranges we need to know what an apple is. And not just a round fruit.* (Clinical Director)

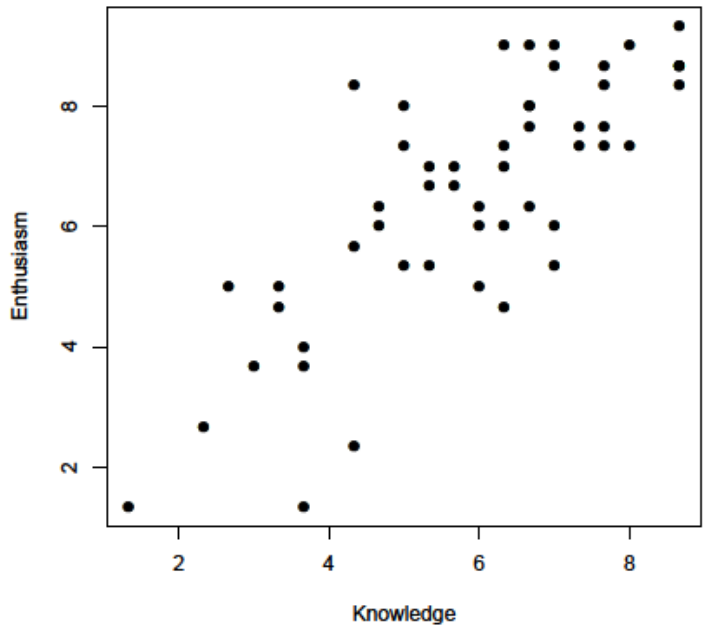
Structural barriers related to the organisation/trust were raised by 50 participants and included the large size of the organisation, lack of systems already in place, lack of resources (money or staff) and the mobility of medical staff.

*I think you know staffing issues is a big thing, you plan to staff a ward, you get last minute sickness, you get bank and agency nurses, and that I think is the difference really if you're getting agency nurses you haven't always got that ownership … You know it's*

### B.3.4 Quantitative analysis of interviews

Seventy-three percent (44/60) of participants scored above 5 on the knowledge scale and 83% (50/60) scored above 5 on the enthusiasm scale. The correlation between knowledge and enthusiasm varied depending on the rater (the ICCs 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 for enthusiasm, 0.63; and 0.58, 0.70 and 0.71). The overall correlation (Figure B.2) between knowledge and enthusiasm was 0.79 (p>0.0001).

**Figure B.2: Correlation between knowledge and enthusiasm for SPI among strategic hospital stakeholders (some dots represent more than one interviewee n=60)**



## B.4 Discussion

Strategic stakeholders generally saw the aims of the **SPI** as legitimate and sound. They accepted that there was a need to control risk (the likelihood of harm occurring) and that patient safety was an important priority for hospitals. Only seven of the 60 (12%) 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 initiative. Most, for example, gave a reasonable account of the PDSA cycle (see 3.6.3.1). There was considerable enthusiasm for the initiative, but 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. The quantitative analysis corroborates the qualitative analysis such that those participants that exhibited greater knowledge where more enthusiastic about the **SPI.**

A limitation of this sub-study was that the number of strategic stakeholders at each site was relatively small (on average 15 per hospital). Also the assessment of the agreement between the accounts of external stakeholders and hospital stakeholders was limited by the small number of external stakeholders interviewed. It is also possible that those who agreed to be interviewed were to some extent self-selected enthusiasts.

Nonetheless, this study does offer insights: First, it suggests that it is possible to get strategic-level individuals across geographically spread and organisationally diverse settings, and interviewed over a period of time, to understand and agree upon a shared model that

can be used in their organisations. Second, it suggests that people at a strategic level are able to recognise the competing interests within their organisations that forms the context of negotiation in which the programme of change is enacted. In particular, they are able to acknowledge the complexities of how to mobilise people, technologies, organisational structures, resources and "culture" around a new effort.

In Suchman's terms (1995), the initiative enjoyed both "pragmatic legitimacy", where it was seen to be of benefit to participants' organisations, and "moral legitimacy", where it was the right thing to do. However, participants also had specific doubts, particularly relating to feasibility of implementation and scientific legitimacy of some elements of the initiative.

# Appendix C: Ethnographic study

## C.1 Introduction

Participants at the "sharp end" of healthcare may not always share the same understanding of the intervention as those at the "blunt end". This sub-study explores the impact of the intervention on these participants and seeks to explain how they identified and classified patient safety risks through observations on medical wards.

## C.2 Methods

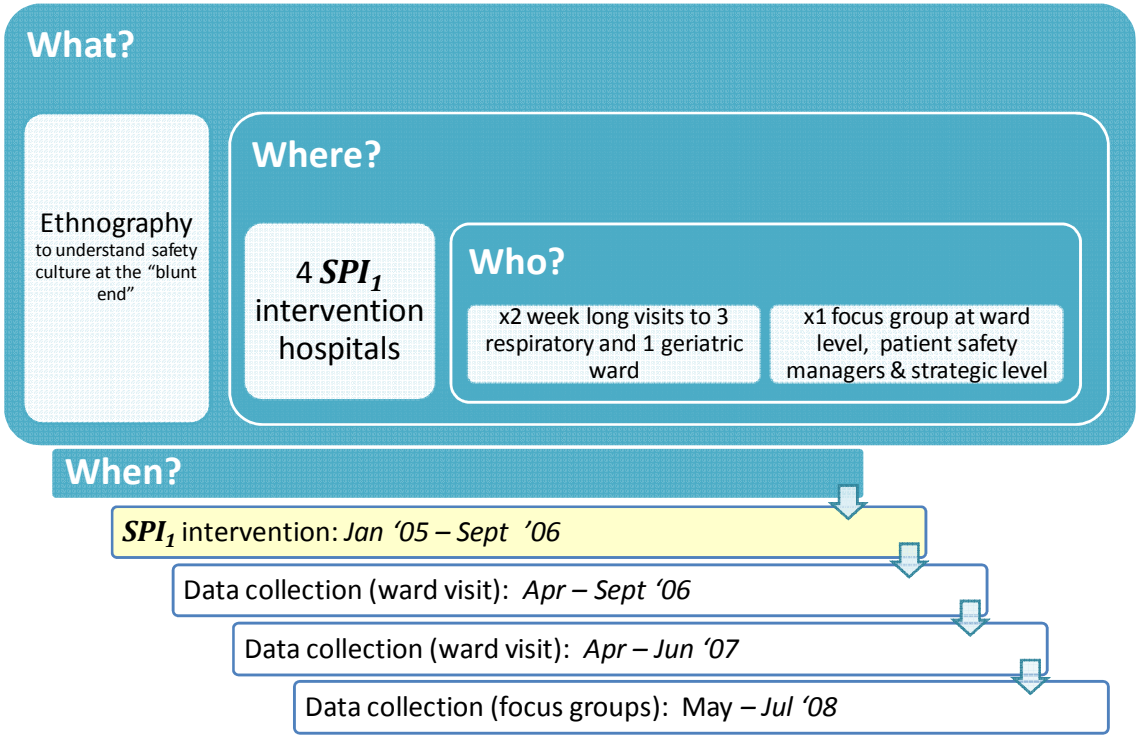
An ethnographic study was conducted at ward level in each of the four ***SPI1*** hospitals. Two respiratory wards, 1 general medical ward, and 1 ward for the care of the elderly were studied. These wards were selected so that they mirrored the data collection in the case-note review of acute respiratory patients (Chapter 6) and so that triangulation between the sub-studies could be made. Three rounds of data collection were undertaken (Figure C.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 ***SPI1*** was being rolled out;
- a week-long second visit to each ward, involving a further approximately 150 hours of observations and 41 interviews, this time with a particular (although not

exclusive) focus on making observations of 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 April-June 2007, during the "embedding and spread" phase of ***SPI$_1$***;

- a third visit involving three focus groups at each hospital (one at study ward level, one involving people with patient safety/**SPI** responsibilities and one at strategic (hospital-wide) 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 between May and July 2008, towards the end of the formal completion of the **SPI** programme in the ***SPI$_1$*** hospitals (September 2008).

**Figure C.1: A summary of the ethnographic study**



Data analysis was based on the constant comparative method (Glaser and Thomson, 1967). 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 fieldnotes, 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 hospital.

### C.2.1 Ethical considerations

Please see B.2.1.

## C.3 Results

The first two visits provided insights into the sharp end of practice on the wards, while the focus groups undertaken on the third visit also provided insights into the other layers of management and strategy in the hospitals. The ethnography was 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 strategic stakeholder interviews (see Chapter 4), the focus groups across the four hospitals agreed that the senior people in the hospitals were committed and enthusiastic about the **SPI**, made a significant strategic contribution, gave weight to the initiative and generally set a good example for staff.

> *if these guys aren't behind it very quickly your clinical directors and you know other directors, you know (other) senior people start to fall by the wayside and 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 in the **SPI** was seen as crucial in lending credibility and support to the implementation of the **SPI**, and was much valued as a source of expert knowledge and expertise.

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

Despite the enthusiasm and support at a strategic level, the management layer (often ward sister or consultant level) between the blunt end and the sharp end appeared less engaged in the **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)

The challenges of engaging middle management and frontline staff were not because they were not interested in or concerned about patient safety. In interviews, these staff gave detailed accounts of the kinds 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'd be in anywhere where you get a big establishment with a lot of people - you've got the doctors, the phyios, the OTs - it's communication. Like the doctors will quite often come and do their ward round and they'll go around and they'll say, "Oh Mr so and so you can go home today", but they won't tell [..] anyone else.*(Interview with ward staff)

Some of problems of engaging middle-management staff and ward are likely to be explained by the day-to-day nature of their work, which was often focused on managing highly complex clinical and organisational demands. Observations suggested that the wards were often very busy and stressed places in which to work, and staff interviews pointed to problems of managing with limited resources, including and especially inadequate staffing or problems of skill mix and shortages of equipment (Dixon-Woods et al., 2009). This meant that staff perceived that they were often too stretched to give priority to things other than the tasks that required to be done immediately. There were suggestions that the effort to improve patient safety needed to be focused on improving aspects of structure rather than on processes:

> *I was telling somebody that if I had one more doctor and a couple of more trained nurses. What I want is trained nurses on the ward so I can manage the ward better and we can discharge many patients sooner and we can reduce the number of complaint letters and of things going off*. (Interview with ward staff)

It was clear also that there were problems of "initiative fatigue" and, for middle managers in particular, challenges in constantly balancing a raft of competing priorities.

*there's so many things now that the nurses in the wards have to do and every time you bring something in, I mean 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, you know, our contingency arrangements, emergency evacuation…[..]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 and that we're just … 'cos 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 has improved the way we work from a point of view that we have you know a better system of monitoring our equipment, where is our equipment going, 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 from that point of view.* (Interview with ward staff)

However, the ethnographic work suggested that the impact of the **SPI** at the level of medical wards were mostly difficult to detect. Apart from improved monitoring of patients using the 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 (top-down) rather that something that they had been involved in developing (bottom-up). 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 left feeling excluded.

> *SPI was a select group of twenty people. I think we could only bring down twenty 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, there is 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 WalkRounds 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 problematic. They appeared useful and welcome in raising senior managers' understanding and awareness of life at the sharp end.

> *so you know if I wanted to find out what was going 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 WalkRounds 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.

*HAVE YOU HAD ANY LEADERSHIP WALKROUNDS ON YOUR WARD?*

> *A: Yeah we had erm … what's his name? One of the guys came down with [name] not that long ago, about a month ago or something it was, he came down for a walkaround.*

*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 on-going study*
>
> *B: Yeah 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 'cos they went along the whole floor on different days and visited all the wards along the medical floor and I think it was quite common ground that 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 stakeholder interview and focus groups 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 know, you haven't sort of broken any rules because in hospitals we are very much bound by, is this the accepted way, is this the allowed thing 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 rather lost.

It appeared that 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 on the surface very simple 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 therefore as not providing enough of a prompt for change – though it may also be the case that claims of problematic evidence were being used simply as a means of resisting change and reinforcing inertia. Legitimacy problems also arose when staff did not recognise that the problem being tackled was a 'real' one requiring

a particular response, or they did not feel that the resources that would be required to implement the intervention was a legitimate use of staff time in light of other priorities. There were also suggestions that that legitimacy varied among different staff groups; junior doctors were often seen (by nurses in particular) as especially difficult to engage.

> *I think it's very variable, I think there's some senior clinicians that are very supportive and I think on the whole at that level they are supportive and they seem to be supportive and they show that support but we still don't seem to be solving the problems with the junior doctors.'* (Focus Group)

The second related to how "trackable" any improvements were. One of the reasons why the EWSS seemed to penetrate practice was that it 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 SBAR communication technique seemed to have relatively low uptake. In the focus groups, the apparent "failure" of the safety briefings and SBAR 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*** *it wasn't a measurement, 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 that:

> *A: nursing-wise it's quite consistent but in medical staff, the doctors …*
>
> *B: … we're having still poor response, yeah, poor compliance.*
>
> *A: I think plus we've had a lot of bank and agency working haven't they and they sometimes they don't seem to be as hot on it do they as kind of the regular staff because we're aware of it.* (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 ever 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 big 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 organizational 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. The Safer Patients Initiative had, I think, something like 29 simultaneously (…) I know the American experience was they could choose only six and now with the latest experience they can choose from twelve so we were trying something very new and I just wonder if we tried too much and weren't able to devote enough focus on every element of it.*

> *I kind of 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 you know. I think we're at an advantage nearly now because we have learnt a lot from what we have done over the past three years that 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 meetings, the monthly meetings between medical director and the clinical director and also the clinical managers on a monthly basis 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 'cos 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 and people are saying well "I really have to do that, so I think this, I think we're getting close to that".* (Focus Group)

Avoiding "paperwork" associated with patient safety work 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 well if you asked a member of staff later on that day "what was 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 because 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 learnt from SPI but stop calling things SPI committees and things… … 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 gonna 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 gonna 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 'cos we're introducing training and a few other things as well as… … and they'll be improvement team support for the directorates and the Safety Leads in the directorates, so it's gonna be a different way of looking at it in … with the aim of trying to embed it more.* (Focus Group)

## C.4 Discussion

The analysis shows that there is a disconnect between the strategic stakeholders and front-line staff perception of the **SPI**. Strategic stakeholder focus groups reinforced the findings of the interviews in Chapter 4 and showed enthusiasm and dedication to the **SPI**, whilst, staff at the "sharp-end" were unable to describe the **SPI** in any great detail. A possible reason for this was the lack of engagement of middle management, who described a concern for patient safety but were balancing these concerns with other competing priorities.

This gap was also noticeable in the description of the intervention – strategic stakeholders were able to articulate the **SPI** in detail and describe components such a WalkRounds and PDSA cycles. However, the EWSS was the only aspect that was readily observed on the wards, perhaps as EWSS created a physical record, whilst, there is no such artefact for other components such as safety briefings and SBAR and thus no imperative to engage in these types of activity.

Front-line staff also had little knowledge of the other parts of the intervention and perceived the **SPI** as being handed down to them rather than the participatory model espoused by the IHI. In fact there was a perception of elitism and better career prospect for those that were actively involved in the **SPI**.

Another barrier to compliance was a lack of perceived scientific legitimacy of the methods used in the **SPI** and the inability to sustain momentum because of continuous staff changes.

Nevertheless, there was a positive impact with the **SPI** sites, primarily a managerial recognition and focus on patient safety.

Overall the ethnography within the evaluation suggests that staff at the sharp end 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.

However, a limitation of this study is that it focuses on medical wards and it is unknown if this experience of the **SPI** is shared by frontline staff working within surgical theatres and the ICU – areas which were subject to more targeted components of the intervention, i.e. the peri-operative care bundle and the critical care bundle respectively.

## Appendix D: Staff survey questions

The following 13 questions were identified as being relevant to the **SPI** evaluation *(*these scores were not included in the* ***SPI$_1$*** *evaluation)*. Six of these thirteen scores are straightforward percentages:

1. "Percentage 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 (Guzzo et al., 1985; West et al., 2006)

2. "Percentage 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 (Borril et al., 2000)

3. "Percentage 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 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 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; and

6. "Percentage 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.

Six of the other seven scores were calculated as the mean of a number of separate questionnaire items, each scored from 1 to 5 representing answers form “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 hospital 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 health care organisations (Michie and West, 2004)

12. “Fairness and Effectiveness of Incident Reporting Procedures” is a measure of the extent to which employees their hospital’s procedures for reporting and dealing with errors, near misses and incidents are effective and fair. This is based on seven items.*

One other variable was also measured on a similar scale, but with some slight differences:

13. “Availability of Hand Washing Materials” is a measure of the extent to which hand washing materials (hot water, soap and paper towels, or AHR) are available when needed by different groups. This was originally measured on a scale from 1 to 4 representing answers from “never” through to “always”, and than adjusted to fit a 1 to 5 scale for consistency with the other scale scores. *

## Appendix E: Prescribing errors - stage of drug process and their definition

- **Need for a drug:** includes the following:

1) 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

2) Drug no longer needed
Continuation of a prescribed drug for a longer duration than necessary.

3) No indication for drug prescribed
Prescription of a drug without a corresponding indication.

4) 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:

1) 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.

2) 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.

3) 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.

4) 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:

1) Failure to specify maximum dose
Failure to specify the maximum dose for a drug prescribed to be given as required.

2) 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.

3) Dose/rate mismatch
Prescription of a drug to be infused on a mg/kg/hr basis, where the ml/hr rate calculated does not correspond to the dose required.

4) 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.

5) 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.

6) 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.

7) 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:

1) 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.

2) 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.

3) Route not specified

Failure to state the route of administration for a drug that can be given by more than one route.

4) Prescription not signed
An inpatient or discharge prescription that has not been signed by the prescriber.

5) Controlled drugs prescription requirements
Failure to write a discharge prescription according to the controlled drugs requirements.

## Appendix F: Difference at baseline for quality ratings, errors and adverse events

**Table F.1**: Ratings and rates of adverse effects and errors: differences between ***SPI2*** hospitals and control hospitals at baseline; and changes between Epoch 3 and baseline in the control hospitals (99% CIs are in parenthesis)

| | **Comparisons at baseline** * (1) | **Changes in Controls** * (2) |
|---|---|---|
| | **Intervention – Control** | **Epoch 3 – Baseline** |
| | | |
| **Quality Ratings** | | |
| Admission Rating † | 0.12 (-0.27, 0.50) | 0.11 (-0.32,0.26) |
| Management Rating † | 0.14 (-0.33, 0.61) | 0.28 (-0.29, 0.84) |
| Pre-discharge Rating † | 0.00 (-0.54,0.54) | 0.11 (-0.38,0.60) |
| Overall Care Rating ‡ | 0.10 (-0.30, 0.48) | 0.29 (-0.12, 0.69) |
| **Errors / Adverse Events** | | |
| Number Errors Φ | -5.78 (-23.84, 12.28) | -14.35 (-32.42, 3.71) |
| Number Adverse Events Φ | -1.42 (-5.81, 2.97) | -1.70 (-7.37, 3.96) |

Errors can be of multiple categories

* Effects are estimated from a mixed effects model (see methods for details) and represent differences at baseline (1) and the effect of time (2). Baseline refers to the average scores over Epoch 1 and Epoch 2

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

‡ Score scale: 1 (unsatisfactory) to 10 (very best care)

Φ Number of errors and number of adverse events are per 100 patients (patients could experience more than one error and more than one adverse event)

**Table F.2:** Rates per 100 patients of errors identified by broad category of error: differences between ***SPI2*** hospitals and control hospitals at baseline; and changes between Epoch 3 and baseline in the control hospitals (99% CIs are in parenthesis)

| | **Comparisons at baseline** *(1) | **Changes in Controls** * (2) |
|---|---|---|
| | **Intervention – Control** | **Epoch 3 – Baseline** |
| | | |
| Diagnosis/ Assessment/ Admission Error | -3.28 (-27.15,20.60) | -13.08 (-36.31, 10.14) |
| Hospital acquired infection | -0.00 (-0.93,0.93) | 0.88 (-0.28,2.04) |
| Technical/ management | -3.58 (-10.50, 3.34) | -1.17 (-9.66,7.31) |
| Medication /Maintenance/ Follow-up | -1.08 (-11.24, 9.07) | -8.54 (-21.43, 4.35) |
| Clinical reasoning | -4.90 (-18.56, 8.76) | -10.93 (-24.84, 2.97) |
| Discharge information | 0.62 (-9.43, 10.67) | -5.63 (-16.14, 4.87) |

Errors can be of multiple categories

* Effects are estimated from a mixed effects model (see methods for details) and represent differences at baseline (1) and the effect of time (2). Baseline refers to the average scores over Epoch 1 and Epoch 2.

## Appendix G: Details of fitted models

**Table G.1**: Fitted models for rate of soap and AHR (litres) consumption per 1,000 bed days

| | **Soap** | | **AHR** | |
|---|---|---|---|---|
| | **Coeff (se)** | **p-value** | **Coeff (se)** | **p-value** |
| Constant | 41.76(13.3) | 0.000 | 3.80 (10.5) | 0.708 |
| Intervention | 0.73 (13.9) | 0.941 | 10.90 (12.2) | 0.371 |
| Time | 0.73 (1.82) | 0.623 | 3.91 (1.28) | 0.002 |
| Time^2 | -0.03 (0.08) | 0.657 | -0.12 (0.06) | 0.034 |
| Time^3 | 0.00 (0.00) | 0.501 | 0.00 (0.00) | 0.065 |
| Intervention*Time | 0.08 (0.44) | 0.760 | -0.05 (0.38) | 0.889 |

**Table G.2**: Fitted models for observed to expected mortality ratio (exponential scale) and mean length of stay for patients admitted to ICU unit

| | O/E mortality | | Mean LOS | |
|---|---|---|---|---|
| | Coeff (se) | p-value | Coeff (se) | p-value |
| Constant | 1.28 (0.12) | 0.000 | 180.4 (19.7) | 0.000 |
| Intervention | -0.14 (0.08) | 0.068 | -39.4 (17.2) | 0.022 |
| Before | -0.07 (0.06) | 0.258 | -12.9 (8.49) | 0.128 |
| Intervention*Before | 0.09 (0.08) | 0.250 | 5.9 (11.11) | 0.598 |
| APACHE II Score | 0.01 (0.01) | 0.138 | 0.34 (1.18) | 0.774 |
| Physiology Score | -0.01 (0.01) | 0.015 | -1.34 (0.87) | 0.123 |

**Table G.3**: Fitted models for APACHE II and ICNARC physiology scores for patients admitted to ICU unit from a ward within the hospital

| | **APACHE II score** | | **ICNARC score** | |
|---|---|---|---|---|
| | **Coeff (se)** | **p-value** | **Coeff (se)** | **p-value** |
| Constant | 18.47 (0.72) | 0.000 | 20.95 (1.00) | 0.000 |
| Intervention | 1.20 (0.98) | 0.225 | 2.32 (1.36) | 0.087 |
| Before | 1.85 (0.81) | 0.022 | 1.77 (1.19) | 0.136 |
| Intervention*Before | -0.83 (1.09) | 0.449 | -2.26 (1.60) | 0.158 |

**Table G.4**: Fitted models for rate of C. diff (per 1,000 bed days) and MRSA infections (per 100,000 bed days)

| | C. diff | | MRSA | |
|---|---|---|---|---|
| | **Coeff (se)** | **p-value** | **Coeff (se)** | **p-value** |
| Constant | 0.94 (0.22) | 0.000 | 15.36 (2.51) | 0.000 |
| Intervention | 0.05 (0.28) | 0.853 | 2.37 (0.14) | 0.420 |
| Time | -0.13 (0.07) | 0.051 | 0.26 (0.50) | 0.601 |
| Time^2 | -0.01 (0.01) | 0.264 | 0.01 (0.03) | 0.789 |
| Time^3 | 0.00 (0.00) | 0.784 | -0.00 (0.01) | 0.208 |
| Intervention*Time | -0.01 (0.02) | 0.652 | -0.05 (0.14) | 0.693 |

## Appendix H: Patient survey questions

Five scores were identified as being 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?" (same response options);
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).