The Scale and Scope of Preventable

Hospital Deaths

Thesis submitted to the

University of London for the

Doctor of Philosophy Degree

PhD candidate: Dr Helen Hogan

Supervisor: Professor Nick Black

Department of Health Services Research and Policy,

Faculty of Public Health and Policy

1st May 2014

Declaration

I, Helen Hogan, confirm that the work presented in this thesis is my own.

Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature:

Date: 1st May 2014

Abstract

In 2008, the lack of a robust estimate for the proportion of patients experiencing preventable deaths in English acute hospitals was fuelling debate and hindering progress in tackling the underlying problems associated with serious patient harm. In this thesis a narrative literature review and a study of harm measures in a single acute hospital are used to guide the choice of method for a study to determine the proportion of preventable hospital deaths.

A subsequent retrospective case record review (RCRR) of 1000 randomly sampled deaths from 10 English acute hospitals found the proportion of preventable deaths to be 5.2% (95% CI, 3.8% to 6.6%) which would equate to 11,859 (95% CI 8712 to 14 983) preventable deaths per year in NHS hospitals in England, 60% of whom had a life expectancy of less than 1 year. The proportion was lower than previous estimates based on US RCRR studies but consistent with a recent Dutch study which reviewed 3,983 hospital deaths.

The majority of underlying problems in care were related to clinical monitoring, diagnostic error and drug and fluid problems, and 44% occurred during ward care. Problems were more likely to occur in surgical than medical patients (23.6% vs 12.7%). Three-quarters were omissions, rather than commissions, in care and accumulated throughout the hospital episode. While there was a strong positive correlation between proportions of preventable deaths in hospitals and MRSA bacteraemia rates (r=0.73; p<0.02) there were no other significant associations with common measures of safety, including HSMR.

Improvements are needed to reduce human error and to provide better quality of care for acutely ill older people to reduce serious harm in acute hospitals. A national mortality review process, based on this study, is to be rolled out across the NHS and will provide one mechanism for monitoring progress.

Acknowledgments

Table of Contents

List of Tables

List of Figures

List of Boxes

Abbreviations

AE   Adverse Event

AHRQ   Agency for Healthcare Research and Quality

CMO   Chief Medical Officer

DH   Department of Health

GTT   Global Trigger Tool

HMPS   Harvard Medical Practice Study

HSMR   Hospital Standardised Mortality Ratio

ICD   International Classification of Diseases

IHI   Institute for Healthcare Improvement

NCEPOD   National Confidential Enquiry into Patient Outcome and Death (formally

                National Confidential Enquiry into Perioperative Deaths)

NHS   National Health Service

NPSA   National Patient Safety Agency

NRLS   National Reporting and Learning System

PSI   Patient Safety Indicators

QAHCS   Quality in Australian Healthcare Study

RCRR   Retrospective Case Record Review

SHMI   Summary Hospital-level Mortality Indicator

Glossary

Acts of Commission: affirmative actions related to the active delivery of care, such as incorrect treatment or management

Active Failures: errors committed at the interface between the health professional and the patient

Acute Hospital: a hospital providing care for both elective and emergency patients across a broad range of medical and surgical specialties

Adverse Event: an unintended injury or complication resulting in prolonged hospital stay, disability at the time of discharge, or death caused by healthcare management rather than by the patient's underlying disease process

Care Delivery Problem: a problem that arises in the process of care due to actions or omissions by staff

Clinical Technical Processes: the processes related to diagnosis and management that are targeted at the patients' presenting problems

Complications: unexpected harm where care had been delivered to an acceptable standard and was error free, such as a drug reaction on the first dose of a new medication

Contributory Factors: underlying or intervening variables at individual, team or organisation level which lie behind the failure of processes of care

Explicit Review: case record review that uses predetermined criteria to assess processes of care

Harvard Medical Practice Study: the largest retrospective case record review study which examined 30,121 randomly selected case records from 51 New York acute hospitals in the 1980s

Hospital Standardised Mortality Ratio: a measure calculated from the ratio of a hospital's observed death rate to an expected death rate derived from a regional or national average. Scores above the average are said to identify organisations with excess deaths

Implicit Review: reviews of the case records conducted without any pre-set criteria, and which use clinician judgements based on knowledge and experience to assess whether processes of care were of an acceptable standard

Latent Failure: error- provoking conditions within the organisational environment that can increase the likelihood of patient harm

NHS Outcomes Framework: a collection of national indicators used by the English Department of Health to hold the NHS to account

National Patient Safety Agency: a national agency established in the early 2000s to lead and contribute to improvements in patient safety across the NHS

National Reporting and Learning System: a national patient safety incident reporting system housed by the National Patient Safety Agency

Non-Technical Processes: processes related to the wider aspects of healthcare delivery beyond the clinical doctor- patient encounter and determined by organisational factors such as leadership, modes of communication or teamwork

Omissions: inactions such as failure to diagnose or treat

Patient Safety Incident: an unintended or unexpected incident that could have harmed, or led to harm of NHS patients

Problem in Care: patient harm resulting from:
a) Acts of omission (inactions), such as failure to diagnose or treat
b) Acts of commission (affirmative actions), such as incorrect treatment or management
c) Unintended complications of healthcare.

Retrospective Case Record Review: traditionally a two-stage process consisting of an initial nurse-led screening stage followed by an in-depth review of screen positive records by a senior doctor to identify whether patients were harmed by healthcare

Service Delivery Problem: a problem associated with decisions, procedures and systems at organisational level

Summary Hospital-level Mortality Indicator: a measure derived from routine hospital data calculated from the ratio of a hospital's observed death rate to an expected death rate derived from a regional or national average. It differs from HSMR by including all in-patient deaths and deaths within 30 days of discharge in the analysis

System-Related Harm: harm that arises from factors related to the structure or organisational aspects of care delivery

Chapter 1 Introduction

1.1 Context

Around 15 million people are admitted to hospital each year in England and Wales, and the majority are treated safely and discharged, satisfied with the outcome.[1-2] Unfortunately, for some patients something will go wrong, resulting in harm and sometimes death. Florence Nightingale was one of the earliest figures to attempt to assess the degree of harm caused by healthcare through closely observing the outcomes of her patients. Yet it was not until after the Second World War, when new and more complex therapies were rapidly being introduced, that consciousness began to spread amongst health professionals of the risks attached to such therapies. In 1964, Schimmel described as 'noxious episodes' all untoward events, complications, and mishaps resulting from diagnostic or therapeutic procedures instituted in the hospital.[3] His clinical staff reported that 20% of patients receiving hospital care experienced a noxious event. Despite the gradually emerging evidence, a strong belief in technological innovation as a force for good led to a prevalent attitude amongst healthcare professionals that healthcare benefits outweighed the risks of harm.[4]

Rising litigation costs in the US from the 1950s onwards saw a renewed focus on the scale and scope of healthcare related harm. In California, as part of an investigation into the feasibility of a state-wide, 'no fault' insurance scheme, the first large-scale investigation into the proportion of such adverse events was undertaken using retrospective case record reviews (RCRR).[5] The seminal Harvard Medical Practice Study (HMPS), conducted in the late 1980s, examined 30,121 randomly selected records from 51 acute hospitals in New York and for the first time established a baseline for such events.[6] Over the next two decades, replica healthcare harm studies were undertaken across the developed world and, more recently, such studies have begun to emerge from the developing world.[7]

Although awareness of healthcare related harm grew steadily amongst healthcare professionals during the half century after the second World War, it was not until the publication in the US of the Institute of Medicine report, *To Err is Human: Building a safer healthcare system* in 1999,[8] that acknowledgement of its potential scale spread to a wider community of politicians, policy makers, patients and the general public. The report estimated that between 44,000 and 98,000 Americans died each year as a result of the healthcare they received, and concluded that this harm represented the eighth leading cause of death in US hospitals. In 2000, the UK Department of Health's (DH) Chief Medical Officer (CMO) published a review of patient safety in the National Health Service (NHS), entitled *An organisation with a memory*,[9] in which he extrapolated figures from the US studies to estimate that between 60,000 and 250,000 patients might be suffering severe injury or death as a result of NHS care. The report indicated that settlement of the resulting clinical negligence claims cost the NHS around £400 million per year, and that the additional hospital bed days cost as much as £2 billion annually. Furthermore, highly publicised failures such as at Bristol Royal Infirmary (high death rates following paediatric surgery)[10] and Stoke Mandeville Hospital (deaths following an outbreak of Clostridium difficile),[11] were leading to increasing concern that the true burden of healthcare related harm in the NHS had not been uncovered.

During the early 2000s, debate increased over the scale and scope of severe harm, and its ultimate outcome - preventable death in acute hospitals. The estimates in the CMO report had been drawn from the findings of two studies conducted in the 1980s and 1990s: the HMPS study and a subsequent study of 14,000 patient records in Utah and Colorado.[6 12] A range of alternative estimates for the proportion of deaths in the NHS associated with healthcare harm were also in circulation. In 2001, the Bristol Inquiry report quoted 25,000 deaths annually.[10] This figure was based on the US estimate of 98,000 deaths per year approximately adjusted for the size of the UK population. In 2004, Aylin *et al* cited the figure of 40,000 deaths,[13] which was followed in 2005 by citation of the figure of 34,000

deaths in the National Audit Office report, *A Safer Place for Patients: Learning to improve patient safety*.[14] Both of these estimates were based on a pilot RCRR conducted by Vincent *et al* in 1999, which reviewed the records of 1014 patients from two London hospitals.[15] The accuracy of these estimates is questionable, reliant as they are on relatively small numbers of deaths in the study sample. In contrast, a Dutch RCRR undertaken in 2005 of 8,415 patient records, including nearly 4,000 deaths, found preventable harm associated with death in 4.1% of patients.[16] This would equate to a figure for the NHS of 11,250 preventable deaths per year.

Although traditional RCRR studies are designed to measure the proportion of preventable harm in patients who die, they are not designed to assess the causal association between the preventable harm and the subsequent death, i.e. whether the harm caused the death. This would require the reviewer to also take into consideration factors such as the acuity of presentation, co-morbidities or typical prognosis that also have a bearing on the risk of death. A single study from the US, which examined 116 deaths across Veterans Administration System hospitals, did consider these factors and found the proportion of preventable deaths to be 6%.[17] This study also showed that the majority of patients experiencing preventable deaths had very limited life expectancies. This was in contrast to the Institute of Medicine report which, by suggesting that US deaths caused by healthcare harm are equivalent in number to two jumbo jets crashing every day, created the impression that the problem was as likely to impact on the young as the old.[8]

Hospital safety incident reporting systems can offer another approach to understanding the nature of serious healthcare-related harm. The National Patient Safety Agency's National Reporting and Learning System (NRLS) collected 6,688 reports of incidents associated with serious harm or death from acute hospitals in 2005-06, accounting for 1.3% of all reported incidents from a total of 526,599.[18] A special analysis of the serious incidents associated with death identified 425 deaths that were potentially avoidable.[19] Unfortunately, serious

under reporting occurs in such systems and this is particularly the case for harm at the severe end of the spectrum if staff fear they will be blamed.[20]

Following the Bristol Inquiry, the DH had put much faith in the power of publicly available comparative data on hospital mortality to identify outliers for quality and patient safety. The Hospital Standardised Mortality Ratio (HSMR) and more recently the Summary Hospital-level Mortality Indicator (SHMI), case-mix adjusted ratios of observed to expected hospital deaths, easily calculated using routinely collected hospital administrative data, were developed for this purpose. These statistics have been used to infer that hospitals towards the higher end of the ratio distribution have higher levels of avoidable deaths, an assumption that has been questioned primarily because of the lack of rigorous evidence to back it up. In fact, there have been only four studies, all from North America and published between 1987 and 2008, that have looked at the association between HSMR and the proportion of preventable deaths detected by case record review and none for SHMI. Across the studies, the preventable deaths were from selected specialties,[21-22] diseases,[23] or interventions,[24] and therefore had a limited capacity to predict relationships between HSMR and preventable deaths in broader groups of patients. The studies were also limited by sample size (ranging from182-347 patients, except for one with 739 patients[22]) and did not analyse the relationship at the level of individual hospitals, but as aggregated data from groups of high and low HSMR hospitals. Three of the studies either found no correlation,[21-22] or a non-significant negative correlation.[24] Only one study, the smallest, found a significant positive association in hospitals with a high HSMR and preventable deaths, and that was confined to patients in a single disease group (pneumonia).[23] Conducting similar correlation studies in England has been limited by the lack of a reliable estimate for preventable deaths in hospitals.

In 2006, the House of Commons Public Accounts Committee commented that the '…*lack of accurate information on serious incidents and deaths makes it difficult for the NHS to*

evaluate risk or get a grip on reducing high-risk incidents...',[25] a perspective reiterated in 2009 by a House of Commons Health Committee which also looked at patient safety.[26] This uncertainty applied not only to the numbers of preventable deaths, but also to the problems in healthcare that led to these deaths and to the subpopulations most affected. At the time there was no information available as to whether preventable deaths were occurring predominantly in those with an already limited life expectancy, or were foreshortening lives by a substantial number of years. Since then, the publication of failings at the Mid Staffordshire Hospitals Trust,[27] has ensured that the debate around the scale of serious harm including the number of preventable deaths has remained active amongst the public, politicians and policy makers. Clearly it is incumbent upon health services to minimise the risk to patients, and seek to implement good systems that prevent unnecessary harm, including death. Continuous debate over the numbers of preventable deaths in acute hospitals in England, and whether current measures such as HSMR and SHMI, correlate with these deaths, is at best a distraction, and at worst leads to inappropriate decisions on priority setting for improving safety. Providing a clear picture of the size and nature of safety related deaths in hospitals requires a robust multicentre study of deaths in order to develop a good understanding not only of the size and impact of the problem, but also the main causes and underlying causal factors.

1.2 Aims and Objectives

The overall aim of this thesis is to identify the most appropriate method for measuring severe harm, principally in regard to preventable death in acute hospitals, and to determine the proportion of preventable deaths, the nature of those deaths and how this proportion correlates with other patient safety indicators.

The specific objectives are:

- To describe the strengths and weaknesses of current measures of patient safety for identifying harm in hospitals
- To compare the scale and scope of hospital harm identified by different measures of patient safety
- To determine the proportion of preventable deaths, causes, contributory factors, subpopulations affected and years of life lost in acute hospitals in England
- To determine whether the proportion of preventable deaths across hospitals correlates with other patient safety indicators

1.3 Conceptual Framework

In his 2008 report *High Quality Care for All*, Lord Ara Darzi identified patient experience, effectiveness of interventions and patient safety as key domains of the quality of healthcare provided in the NHS.[28] This thesis is grounded in the third domain of patient safety, and focuses on the area of harm measurement.

1.3.1 Defining Harm

There are many different ways that healthcare can harm patients, and any overarching definition has to have a broad scope if it is to be applied to a general population of adult inpatients. The Oxford English Dictionary defines harm as a *'physical injury, especially that which is deliberately inflicted'*.[29] Harm in the context of healthcare has been defined by the World Health Organisation/ World Alliance for Safer Healthcare as *'harm arising from or associated with plans or actions taken during the provision of healthcare, rather than an underlying disease or injury'*.[30] This definition clearly links such harm to the provision of

healthcare, but fails to incorporate harm that arises from failures to undertake plans or actions. For the purpose of this thesis, I use a broad definition that includes any harm arising from the provision of healthcare. I further classify healthcare harm as harm due to:

- Acts of omission or inactions such as failure to diagnose or treat
- Acts of commission or affirmative actions related to the active delivery of care such as incorrect treatment or management
- Unintended or unexpected complications of healthcare: the occurrence of harm despite care that was delivered to an acceptable standard and was error free, such as an adverse drug reaction following the first dose of a new medication.

Since the HMPS, the occurrence of healthcare harm has been customarily labelled as an 'adverse event' (see Box 1 for definition).[6] More recently, the NPSA customised this term for use in the NHS, naming harm events as 'patient safety incidents'.[31] My research uses a novel term, 'problem in care', to describe healthcare related harm. Its definition is designed to extend the focus on harm beyond discrete incidents, thus ensuring that it includes injury resulting from multiple omissions in care, particularly if these occur over days or weeks.

Box 1.1: Definitions of patient harm

- **Adverse Event**: An unintended injury caused by healthcare management rather than the patient's disease that resulted in temporary or permanent disability, death or prolongation of hospital stay (Harvard Medical Practice Study)
- **Patient Safety Incident**: Any unintended or unexpected incident that could have or did lead to harm for one or more persons receiving NHS healthcare (National Patient Safety Agency)
- **Problem in Care**: Patient harm resulting from acts of omission (inactions), such as failure to diagnose or treat, or from acts of commission (affirmative actions), such as incorrect treatment or management, or harm as a result of unintended complications of healthcare.

1.3.2 Design of a Conceptual Model for Safety and Harm Measurement

Understanding the theories underlying the generation of harm in healthcare is essential when considering the development of research into patient safety measurement. As part of developing the Donabedian Model, a framework for measuring quality in healthcare, the physician Avedis Donabedian described healthcare organisations as having structures, processes, and outcomes.[32] An organisation's structure is the context in which care is delivered, and reflects both physical (e.g. facilities) and organisational (e.g. proportion of trained staff) characteristics. Processes are activities related to the provision of healthcare including the actions of doctors, nurses and patients, and spanning from prevention to cure. These processes are sometimes broken down into technical processes that describe how care related to diagnosis and treatment is delivered to the patient, or non-technical processes reflecting wider aspects of healthcare delivery, encompassing the interactions between people and knowledge sharing determined by organisational factors such as modes of communication, teamwork and leadership.[33] Structure and processes combine to culminate in health outcomes. Harm can be considered an adverse outcome of structural and process factors within healthcare organisations.

In a paper exploring the conceptualisation of patient safety, Brown *et al* adapted Donebedian's model to show how breakdowns at each point in the framework can become part of the causal chain in harm generation (Figure 1).[34] The causal chain model explains how structural factors influence clinical technical processes and thence harm mediated by intervening variables (also known as contributory factors) such as shift work, team structures or modes of communication (non-technical processes).[35] This model also draws on the work of James Reason, a psychologist, who has been an influential thinker in the field of patient safety. He conceptualises a healthcare organisation as a complex system made up of the activities taking place between clinicians and patients, the organisation's design and procedures and the influence of external factors. He describes errors that occur at the

interface between the clinician and the patient as 'active failures' and those which are related to how the organisation is run as 'latent failures'. These 'latent failures' can be seen as creating the climate in which 'active failures' are more likely to occur.[36]

Figure 1.1 Conceptual model outlining points where harm can occur and preventative interventions focused
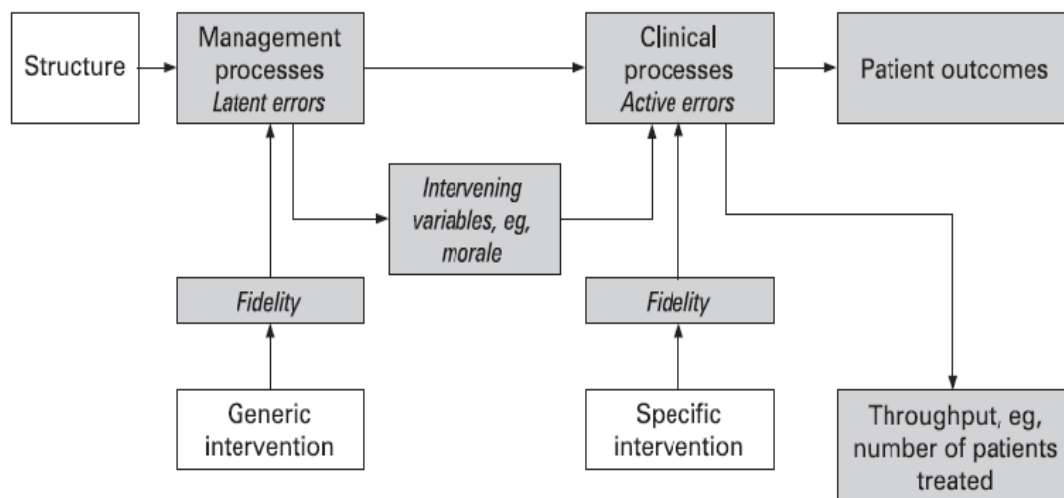
Thomas and Petersen proposed a framework for harm measures which also builds on the work of Reason (Figure 2).[37] The framework places different measures along a continuum. At one end are measures that provide information on the context in which the harm occurred, therefore shedding light on error provoking environments or latent/ system failures. These include malpractice claims files, incident reports, and morbidity and mortality meetings. Although such sources can provide valuable information on system-level issues, they cannot be used to determine incidence because of reporting and selection biases. At the other end of the spectrum are methods that collect information on harm prospectively, including direct observation and prospective clinical surveillance, which will be more likely

to identify active failures at the level of patient and practitioner. These approaches can be used to measure incidence, and are therefore better placed to allow the measurement of the impact of interventions to improve safety.

Figure 1.2 Thomas and Petersen's framework for harm measures

Latent errors   Active errors   Adverse events

•Incident reporting   • Chart review   • Direct   • Clinical
                                                      observation   surveillance
•Autopsies and   • Administrative
morbidity and   data analysis
mortality conferences
                             •Information
•Malpractice claims   technology
files analysis
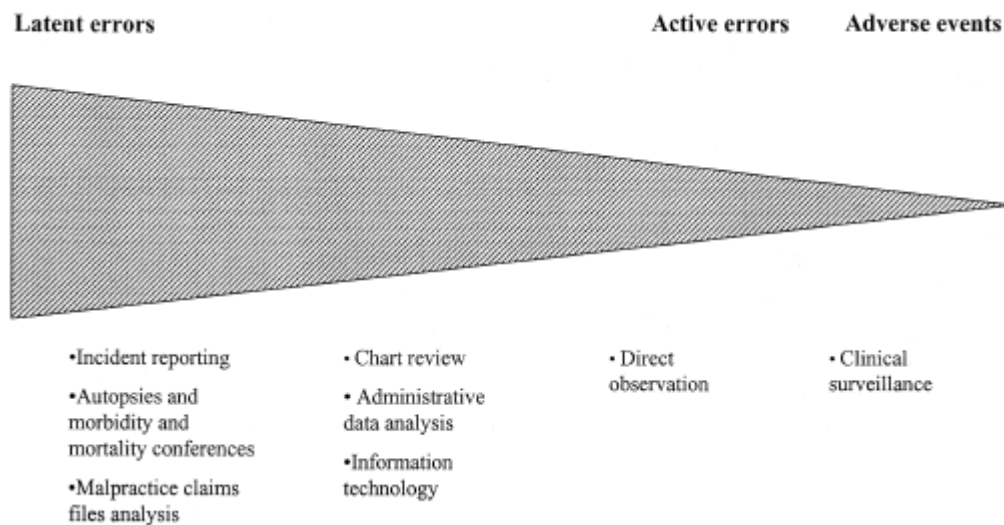
Reproduced from Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. *J Gen Intern Med* 2003;18(1):p.64.

I have combined these two models to create an overarching conceptual model for healthcare-related harm measurement, in order to inform my research (Figure 3). The proposed conceptual framework models the causal chain of harm generation after Brown *et al*,[34] and acknowledges that different harm measures are likely to identify different types of problems in care (system or clinical) as outlined by Thomas and Petersen.[37] It extends the models proposed by these two groups by adding other safety related measures to the components of the causal chain, and categorising harm measures by health services or research orientation. Interpretation of this conceptual framework leads to the prediction that there should be a correlation between measures related to components of the causal chain and harm.

Figure 1.3: Conceptual model outlining points where factors in the generation of healthcare-related harm and harm itself can be measured

Structural measures:
e.g. staffing levels,
existence of Safety Policies

Management process
measures:
e.g. safety culture

Clinical process measures:
e.g. staff hand-washing

Health Service orientated
harm measures

Routine Data
Incident Reporting
Global Trigger Tool
Morbidity and Mortality Meetings
Confidential Enquiries

Claims Files
Retrospective Case
Record Review
Prospective Surveillance

Structure → Management Processes → Clinical Processes → Patient Outcome: Harm Preventable & Non-Preventable

Intervening Variables

Latent Failure / System
Problem: problems related to
the organisation of healthcare
delivery that create
vulnerability to unsafe care

Direct Observation
Case Control
Patient Reported Harm

Active Failure / Clinical
Problem: error of
omission or commission
affecting processes related
to diagnosis and
management that are
targeted at the patient's
presenting problems

Research orientated
harm measures

Measures for intervening variables
e.g. team communication measures
using direct observation

Measures to identify latent failures:
Incident Reporting
Malpractice Claims File Analysis
Mortality and Morbidity Meetings
Prospective Surveillance with Staff
Interviews

Measures to identify active failures:
Retrospective Care Record Review
Routine Data
Electronic Mining of Case Records
Global Trigger Tool
Incident Reporting
Morbidity and Mortality Meetings
Malpractice Claims File Analysis
Direct Observation
Patient Reported Harm

Based on 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(3):158-62)

The evidence linking structural factors to safety, though still limited, has been accumulating since the 1970s. It suggests that factors such as numbers and qualification levels of nursing staff,[38] work scheduling for junior doctors,[39] presence of hospital quality improvement systems,[40] and hospital design features[41] all have an impact on levels of patient harm. Amongst organisational measures, the most developed are those for evaluating safety culture. Safety culture can be conceptualised as the values, attitudes and behaviours that influence an organisation's commitment to patient safety improvement.[42] Organisation scores using safety culture measurement tools have been shown to be associated with frequency of hospital acquired infection, pressure ulcers and drug errors.[43-45] Direct observation has been used to measure important intervening variables, such as teamwork, and has identified the importance of multidisciplinary team composition and knowledge sharing in harm prevention in intensive care units and surgical operating theatres.[46-47] The links between clinical processes and patient harm are well established through numerous RCRR studies of harm.[48]

The conceptual framework can also help highlight one of the key debates in patient safety measurement; whether it is better to measure patient harm, or the underlying errors (both active and latent) that lead to that harm. Harm measurement might seem the obvious approach, responding as it does to the fundamental principle of care provision, i.e. that it causes no unnecessary injury. Its measurement can prove an effective way of gaining the attention and involvement of healthcare professionals in quality and safety improvement. This is particularly the case in systems such as the NHS, where errors can be more easily regarded as trivial, or an inevitable part of its unreliable functioning.[49-50] However measurement of harm is complex, particularly at the severe end (including death), as such events are relatively rare phenomena. Moreover, different approaches to defining and detecting harm lead to different findings. Even with the most clear cut events such as death, the debate over the use of HSMR highlights the challenges of case mix adjustment when such measures are used to compare organisations.[51-53]

In contrast, others argue that measurement of error is superior because errors are more common than harm, therefore offering greater precision in measurement.[54] Furthermore, the study of errors in the processes of care, can more easily identify exactly where such care needs improving in relation to both human knowledge and skills at the clinical interface, and for system level issues.[55-56] Outcome measures such as harm, being dependent on multiple variables, can be poor indicators of where to target improvement.[57] However, there are a number of disadvantages of error measurement that mean it is unlikely to ever replace harm measurement as a reliable metric. Firstly, like harm, definitions of error are subject to debate. Over time there have been a plethora of definitions, from error as an underlying causal factor, to error as an event (process definition), as well as error as an outcome.[58] Only in the last decade, influenced by the work of James Reason, has the process definition of error become the most commonly adopted view.[59] Confusion between the different notions of error can lead to a loss of clarity over what should be measured, and against which standards (if these standards exist at all). Secondly, a focus on error rather than harm also has the potential to stigmatise staff and reinforce a culture of blame.[60] Despite the evidence that system or 'latent' factors often underlie an individual's error (active error), the individual's error is often more visible and focussed on, especially in organisations with weak safety cultures.[61] Because of its multifactorial origins, attention on harm moves the safety improvement focus away from individual error towards the less stigmatising identification of system flaws.[62] It also acknowledges that not all harm is caused by underlying errors and provides an opportunity to work towards increasing the safety of care through risk reduction.[63]

My conceptual model is crucial in fulfilling the aims of this thesis, and will be used to direct an exploration of different measures of patient harm. In turn this will enable the development of an understanding of which harm measure might best fulfil the requirements of a study to establish a baseline proportion of preventable death, and identify underlying

problems in care contributing to such deaths. It will also help identify which patient safety measures one might expect to correlate with preventable hospital death.

In this thesis the focus will be on deaths that occur during a hospital admission. Examination of hospital deaths would seem a logical approach to measuring the quality and safety of hospital care, being an easily defined outcome of such care and one held to be important by the public, politicians and clinicians alike. However, deaths occur in less than 5% of hospital admissions,[64] and many of these deaths are expected, as up to 50% of the UK population will come to hospital to die.[65] Outcome measures can be poor at indicating where interventions for improvement should be focused and many other factors apart from quality and safety can influence these measures, for instance differing lengths of stay and availability of alternative provision for end of life care will influence the proportions of deaths that occur in hospital.[53] [66]

1.4 Structure of the thesis

Chapter 2 presents a narrative literature review which examines the background to harm measurement, describes the most common measures, outlines the epidemiology of harm derived from these measures, looks at issues to be considered when measuring harm, and compares the performance of the different harm measures. Chapter 3 (Research paper 1) describes an exploratory study conducted in one hospital, which examined the utility of a range of information sources to provide information on patient harm. This work, in association with the literature review, contributed to the development of the methodology for measurement of preventable deaths. Chapter 4 draws on the findings from Chapters 2 and 3 and describes the development of the methodology used in the study to determine the proportion of preventable hospital deaths and their nature. Chapters 5, 6 and 7 comprise three research papers. The first covers headline findings from my RCRR study of 1000

deaths across ten English acute hospitals, related to the proportion of preventable hospital deaths and their causes, the second describes in more detail the causes of preventable deaths, and the third explores correlations between hospital preventable death proportions and other measures of patient safety. The final Chapter is an overview of the main findings and discusses the limitations of the thesis, as well as opportunities for future research, along with policy and practice implications.

1.5 Contribution to the thesis

I undertook the background literature review. I took the lead in the design of all studies which make up this thesis and was supported in this by the research study co-authors. I collected all data for Paper 1(Chapter 3). Dr Sisse Olsen and Dr Graham Neale acted as second reviewers for the case record reviews. Dr Frances Healey was second reviewer for the case narratives explored in Paper 3 (Chapter 6). I undertook all data analysis and was provided with statistical support by Dr Jenny Neuburger and Mr Andrew Hutchings. All co-authors of the research papers contributed to data interpretation. Professor Charles Vincent and Dr Frances Healey provided guidance on presentation of the findings in Research Paper 3 and Professor Nick Black for Research Paper 4 (Chapter 7). I produced the first draft of each research paper and made changes in response to co-authors' and peer reviewer feedback.

The main study was funded by National Institute of Health Research (NIHR), under the Research for Patient Benefit Programme. The candidate was the Chief Investigator, and Professors Black, Vincent, Thomson, and Drs Neale and Healey were co-investigators.

1.6 Overall contribution of the thesis to the field of study

My thesis examines the field of harm measurement to identify which approach is best in determining the proportion of preventable deaths in acute hospitals. Chapter 3 (Research paper 1) describes the findings when using a range of approaches to measuring harm in a single acute hospital Trust. The paper concludes that different information sources identify different patient harms, but that these are not harnessed in tandem to allow the development of a better understanding of key risk areas.

In Chapter 4, my thesis outlines the methodology behind the development and implementation of the largest study of preventable death ever undertaken in the UK. The results of this study, including a robust estimate of the proportion for preventable death in English acute hospitals, are presented in Chapter 5 (Research paper 2). A proportion of 5.2% was found, which was lower than previous estimates based on extrapolations from US studies, but consistent with findings from a more recent Dutch RCRR study. The majority of problems in care that contributed to preventable death were related to clinical monitoring, diagnostic error, and drug and fluid problems. The study methodology has been actively drawn upon to guide the development of an approach that can be used to measure such deaths at a national level. This measure will become a new NHS Outcome Framework indicator in 2014, entitled 'hospital deaths due to problems in care'.[67]

Chapter 6 (Research paper 3) presents the findings of a novel content analysis of the case narratives for each preventable death collected during the case record reviews, and reveals more detail of the nature of underlying problems in care linked to such deaths. Problems that have been previously identified, such as failure to monitor anticoagulant medication, poor management of fluid balance, and failure to adequately assess patients and to give indicated drugs, appear still to be common within the NHS, despite a number of national policies and campaigns directed at such issues over the last decade.[31 68-70] The fact that around 70% of

such problems were related to omissions of care indicates that there are persistent failures to tackle reliability within the health service.

Chapter 7 (Research paper 4) outlines findings from an examination of the association between the proportion of preventable deaths found in acute hospitals and other safety measures including the HSMR. No significant correlations were found, with the exception of MRSA bacteraemia rates. This finding casts doubt on previous assumptions that HSMR/ SHMI measure preventable deaths. One recommendation of the Keogh Review[71] of 14 NHS acute hospitals carried out in 2013 in response to quality and patient concerns, was that my colleagues and I should extend our RCRR of hospital deaths to a further 24 hospitals. Combined with the findings from the first ten hospitals, there will then be adequate statistical power to determine if a clinically and statistically significant association exists between preventable deaths identified by case record review and HSMR/ SHMI.

Chapter 2 Measuring Harm in Healthcare: Background

2.1 Introduction

This chapter is framed partly by my conceptual framework, and partly by drawing on previous work on specifications for quality measures.[72-73] It draws on an extensive body of literature drawn from a wide ranging search strategy (see Appendix 1). The chapter begins with a brief description of the history of harm measurement. Then, guided by my conceptual framework, I identify the main measures of patient harm, reviewing current use and scope for measuring different types of harm. These measures are grouped by measurements more commonly used in hospital practice, and those with more of a research focus. There follows an overview of the epidemiology of patient harm, with a focus on findings from RCRR studies.

As harm can be regarded as an outcome indicator for poor quality care,[34] the next section explores harm measurement against criteria that have been developed to assess the technical attributes of healthcare quality indicators. These criteria draw on those developed by the Institute of Medicine in the US as part of its Medicare Quality Assurance Programme, and the World Health Organisation Performance Assessment Tool for Quality Improvement in Hospitals (PATH).[72-73] The harm measures are then directly compared against these criteria and each other.

The chapter concludes with consideration of which approach may be best for measurement of the scale and scope of preventable deaths in hospitals in England. The work presented in this chapter and the following chapter guided me in the choice of harm measure for a study to ascertain a national estimate of the proportion of preventable hospital deaths in England and their underlying causes.

2.2 The Development of Harm Measurement

Receiving healthcare can be a hazardous business for a patient. The driving force of all patient safety initiatives is to prevent patient harm as a consequence of healthcare. More specifically, patient safety can be defined as '...*the avoidance, prevention and amelioration of adverse outcomes or injuries stemming from the processes of healthcare.*'[74] Attainment of such goals is dependent on identifying problems that occur in healthcare, the harm that results and their frequency. Florence Nightingale and later, Ernest Codman, a 19th Century US surgeon, can be seen as the founders of the modern safety measurement movement. Both took an interest in the outcomes of healthcare interventions, particularly in the numbers of patients who died following such interventions, and the causes of those deaths. However, it was the emergence of peer review organisations in the US from the 1970s, combined with a growing interest in establishing the contribution of hospital related harm, that really stimulated the development of harm measurement.[5 75] Retrospective methods for analysing the contents of case records emerged and became formalised as the RCRR approach, traditionally a two stage process consisting of an initial nurse-led screening stage followed by an in-depth review by a senior doctor of screen positive records. Researchers from Boston, when designing the first large scale epidemiological study based on RCRR, the HMPS, built on these foundations; they improved reliability by introducing a structured review form, systematic training of reviewers, and a proportion of double reviews at the screening and full record review stages. Over 50 years later, the RCRR remains the internationally recognised method for the measurement of patient harm in hospital settings.

In the UK, the traditional forums for examining patient harm were Mortality and Morbidity meetings, principally run amongst surgical specialties. In the 1930s, the first national Confidential Enquiry was established, with the aim of identifying problems in care associated with maternal deaths in a more systematic fashion.[76] Many more national

confidential enquiries were introduced, including the National Confidential Enquiry into Perioperative Deaths (NCEPOD) in 1986.[77]

Over this period, growing concern aroused by increasing litigation in the NHS led to the development of hospital risk management programmes, which had already successfully reduced the number of health-related legal claims in the US.[78] Such programmes specified the need for an incident reporting system similar to those used to improve safety in the high risk industries of aviation, nuclear power, and oil.[79] In *An organisation with a memory*, published in 2000, the CMO for England exposed the scale of harm in the NHS and put forward a national programme for its amelioration.[9] One specific action was to set up a national incident reporting system, and the NRLS was established in 2004. It was to be run by a new agency, the NPSA. The hope was that the NRLS would initiate a step change in learning from harm. Much effort was put into promoting openness about error and patient harm in NHS organisations which, in turn, it was hoped would lead to good reporting rates.[31] New mechanisms were devised to disseminate learning and ensure that recommendations were acted upon. The NRLS drew on information fed in from local reporting systems, as well as independent NHS staff reports, via a web-based portal. Although incident reporting has gradually increased over the last decade, reaching a total of just over a million reports made by 2012/13, analysis of reporting patterns indicates persistent under-reporting. Combined with the lack of denominators, the result is that the utility of this source in providing an accurate picture of harm in the NHS is limited.[80-81]

The latter half of the 20th Century proved a fertile period for innovation in safety measurement. James Reason's highly influential work exploring the nature of error and harm from a psychological perspective was instrumental in influencing how harm is conceptualised and investigated.[59] Interest in organisational culture and its influence on the proportion of errors and harm has led to a burgeoning of approaches to evaluating safety culture by surveys or direct observation within hospitals.[44][82] The addition of questions on

the witnessing of harm and incident reporting to the annual NHS staff survey acknowledged the patient safety intelligence held by NHS frontline staff. The emergence of new dangers, such as healthcare acquired infections (e.g. MRSA or Clostridium difficile) have led to the establishment of new monitoring systems to track incidence and promote a sense of urgency in addressing these serious problems.[83]

In the last decade, there has been a movement towards developing a more systematic understanding of patterns of mortality as part of a suite of approaches that can be used to identify preventable harm and other quality failings across hospitals and, in doing so, focus improvement efforts.[84-85] Death statistics were first published over one hundred years ago, and have appeared intermittently in the public domain since then.[76] Following the Government's increased interest in the potential of these statistics to benchmark hospitals and provide an early warning system for poor care, The Dr Foster organisation began to publish comparative HSMRs for all acute hospitals in England from 2002.[86] Debate has continued since the introduction of this measure as to the validity of these estimates, and the nature of the 'excess deaths' identified, in terms of just how many of these are actually preventable. With the emergence of reports of hospitals gaming the system, and research indicating that many other factors apart from the quality and safety of patient care have an impact on the value of these statistics, calls have been made to abandon the use of HSMR.[51] [87]

Drawing on the work of the US Institute for Healthcare Improvement[88] and the UK's Modernisation Agency,[89] the NHS Institute for Innovation and Improvement has subsequently advocated the use of case record based mortality reviews for identifying patient harm and focussing safety efforts.[90] This approach was also recommended by NHS national safety campaigns in England and Wales.[69 85] Furthermore, the need to better understand their own fluctuations in HSMRs has generated bottom-up momentum within hospitals to expand the review of deaths beyond Mortality and Morbidity meetings. The

differential impact of these forces has resulted in the emergence of a variety of approaches to the mortality review process. Morbidity and Mortality meetings, RCRR, incident reporting and HSMR have emerged as the main approaches to measuring serious healthcare-related harm in the UK. The next section describes these measures and others in common use in more detail, including an exploration of their scope.

2.3 An Exploration of Harm Measures

Harm measures can be divided into those that have been developed for use in a health service setting, and those that are currently more frequently used in patient safety research. Some measures, such as RCRR, analyses of claims and inquest records, and prospective surveillance span both spheres and others are likely to move from research into the health services over time. Different measures are likely to identify different types of error or harm.

2.3.1 Health Service Orientated Measures

2.3.1.1 Routine Data

Routine data must be collected by hospitals as part of corporate resource management, and therefore provide a cheap and easily accessible source of information on hospital activity. The utility of the data can be enhanced through links with other data sources such as Office of National Statistics (ONS) mortality data. There are three main harm measures derived from routine data: the 41 ICD-10 diagnosis codes for adverse events and misadventures, standardised mortality ratios and Patient Safety Indicators (PSI). Such measures can provide information on harm at individual consultant, department or hospital levels, as well as for different patient subgroups.[91] However, only standardised mortality ratios are used as an indicator of safety in today's NHS.

The 41 adverse event and misadventure codes are primarily limited to surgical and obstetric harm, with the codes identifying problems during and after procedures or complications related to devices, grafts or foreign bodies. None of the codes relate specifically to death. One English study of hospital episode statistics from between 1999-2003 and covering over 50 million episodes of care found at least one of these codes in 2.2% of all admissions.[13] A similar study from Australia found the codes in 4.75% of admissions.[92]

In the US there has been much more interest in developing indicators of patient safety from routine data than in the UK, which probably reflects the better quality and depth of coded activity information available. Early on, attention focused on deriving case mix adjusted measures of hospital mortality using sophisticated algorithms that standardised for age, deprivation, gender, urgency of admission, co-morbidities and diagnosis. Professor Brian Jarman and his team at Imperial College developed the first such measure for use in the UK, the HSMR, in 1999. [93] The measure was calculated from the ratio of a hospital's observed death rate to an expected death rate derived from the national average. Values above 100 are interpreted as 'excess deaths' with the assumption that at least some of these deaths are preventable. Linkage between hospital administrative data and the Office of National Statistics data also allows inclusion of deaths within 30 days of discharge.

In the 1980s, cardiothoracic surgeons in New York were the first medical specialists to make their death rates, at individual surgeon level, available to the public.[94] Since then there has been increasing international political interest in the use of such data for hospital performance management and benchmarking.[95] The UK Government believed publication of league tables for hospital mortality rates would lead to the earlier recognition of problems with hospital safety, and avert future scandals similar to the one uncovered at Bristol Royal Infirmary.[76] Marshall *et al,* in their review of the impact of publicly released mortality data in the US found that publication of such information did lead to changes, both in the processes and outcomes of care; responses being driven by factors such as sensitivity to

public image and the legal risk posed by underperforming doctors.[94] A subsequent review of outcomes following feedback of mortality data to cardiothoracic surgeons in England, also found a reduction in mortality from cardiac surgery.[96] However, other studies have found little or no impact of such information.[97] Smith highlighted that there were also a range of unintended consequences as a result of publishing such data, including measure fixation, tunnel vision, misinterpretation and gaming.[98] Like other summary outcome measures, HSMR is limited in its ability to indicate where resources for improvement should be focused. To date, HSMRs continue to be used for benchmarking hospitals in England alongside the SHMI, a measure similar to HSMR whose calculation is based on a broader range of in-hospital deaths and also takes into account deaths within 30 days of discharge. More recently, the Care Quality Commission, the national organisation charged with ensuring quality and safety standards are maintained in NHS organisations, has commenced monitoring of disease-specific mortality ratios. The measures are being used as triggers for further investigations of hospital practice.

An alternative approach to harm measurement, again largely developed in the US, has been to use signal or indicator codes known to be linked to hospital harm.[99] The algorithms for PSI are created by combining primary and secondary diagnoses with procedure codes. Sophisticated systems can identify those diagnoses present only after admission.[100] The first indicators focused on specific causes of harm, such as hospital acquired infection and drug errors.[101-102] In the 1990s, Iezzoni *et al* in the US broadened the scope to 27 indicators including post-operative haemorrhage, post-operative pneumonia, sepsis and wound infection as part of the Complications Screening Programme Study.[103] Building on this work, the US Agency for Healthcare Research and Quality (AHRQ) created a suite of PSIs for use in inter-hospital comparisons. Again, the majority of PSIs are designed to identify complications following surgical procedures and obstetric trauma rather than medical harm, as more of the codes used in these specialties are clearly linked to harm.[104] Two PSI codes look for potentially avoidable deaths: 'failure to rescue' includes deaths per 1,000 patients

with specified treatable complications of care (e.g. pneumonia, sepsis, gastrointestinal bleeding) develop during hospitalisation, and 'deaths in low mortality diagnosis related groups' includes in-hospital deaths per 1000, in patients with an admission diagnosis that has a less than 0.5% expected mortality. Interest in using PSI is developing in the UK, and some initial work has been done to adapt the algorithms to make this possible.[105]

Harm identified by adverse and misadventure codes or patient safety algorithms is more likely to be related to acts of commission rather than omissions, and clinical rather than system-based. The introduction of the electronic patient record holds promise for more sophisticated data linkage, and measurement of a wider range of harms using routine data.[106]

2.3.1.2 Incident Reporting

Incident reporting systems encompassing the reporting, collating and learning from safety incidents were initially designed to identify specific, usually rare, events like blood transfusion reactions,[107] or problems occurring in the high risk settings of anaesthetic rooms or intensive care units.[108] The UK was the first country in the world to develop a national, voluntary, confidential incident reporting system in 2004.[9] Other countries now have similar systems, but none on the scale of the NRLS. Reports to the NRLS provide descriptive details of the incident, contributory factors and mitigating actions along with an assessment of the degree of harm. Such systems are able to shed light on harm as a consequence of system factors such as low staffing levels, as well as those with a clinical origin, but do tend to pick up more harm linked to acts of commission. A national system is particularly useful for identifying rare harms and high risk areas, and for tracking responses to interventions that address these problems over time.[109]

Although the number of reports has increased steadily since its inception, patterns of reporting have remained similar with two-thirds of reports being no-harm incidents. Falls remain the most common type of incident reported, at approximately 30%.[18] Less than 0.1%

of reports relate to a death.[110] Analysis of trends in patient safety incident reporting to the 46NRLS, shows that hospitals with the highest reporting rates overall (in the top 25%) report fewer incidents linked to no-harm and falls as other types of incident reports take their place. Higher reporting rates are considered to be a feature of a positive safety culture.[111] However, across all organisations, doctors report fewer incidents than nurses do, leading to an under representation of incidents linked to clinical diagnosis, assessment and management.[112-113]

2.3.1.3 Morbidity and Mortality Meetings

In the early 20th century, Ernest Codman, an American surgeon, began to keep records of outcomes following surgery, documenting errors and subsequent harm. This approach developed into the modern day Mortality and Morbidity (M&M) meetings. M&M meetings are the traditional forums for discussing and learning from unexpected deaths or serious complications that occur in surgical and anaesthetic specialties, and have been a prerequisite in NHS hospitals hosting surgical training programmes since the 1960s.[114] Increasingly, meetings to review deaths have been adopted by other specialties, attracting participation from the wider multidisciplinary team and, more recently, reframed as an approach to pinpoint patient safety risks in healthcare provision.[115]

The meetings are a potentially rich source of information on serious harm, especially when post mortem findings are also available. A systematic review has shown that up to 25% of post mortems reveal an unsuspected principal diagnosis, or primary cause of death, providing evidence of missed diagnoses.[116] With the decline in frequency of post mortems, this valuable source of learning is being lost.[117] Although there is potential to collect information on system-related harm as well as clinical harm, to date there has been little systematic collection and analysis of this information. Recent initiatives have looked to standardise the process of case selection, analysis and feedback, in an effort to improve institutional learning alongside educational and peer review elements.[118]

2.3.2 Health Service and Research Orientated Measures

2.3.2.1 Retrospective Case Record Review

The HMPS was the first rigorous application of the RCRR method. The study reviewed 30,121 randomly selected records from hospitals across New York State; it was designed to investigate the epidemiology of healthcare related harm and to build on the findings from earlier smaller studies that had used non-random samples. Traditionally the method consists of a nurse-led initial screening process followed by a detailed clinical review by one or more senior physicians. Doctors are asked to make judgements as to whether harm occurred as a result of healthcare rather than a patient's own illness, the degree of harm, and its preventability. The rich material found in the record can provide the reviewer with a picture of care from admission through to discharge, in addition to information on the context in which care was delivered, and other contributory factors. The technique uses implicit review, whereby reviews of the case record are conducted without any pre-set criteria, and use clinician judgements, based on knowledge and experience, to assess whether processes of care were of an acceptable standard.

Seen as an approach that can shed light on a broad array of harms, especially those at the more severe end of the spectrum, those generated by the actions of doctors and those caused by omissions in care which are difficult to identify using other measures, it has often been used as the 'gold standard' against which other measures are compared. However, case record content is more likely to contain information on technical aspects of care, encompassing processes related to diagnosis and management that are targeted at the patient's presenting problems. Limited information on non-technical aspects of care which relate to the way care is delivered at the clinician-patient interface and on the organisational context in which it is delivered make it more likely that reviews will identify problems related to an individual's actions as opposed to those due to underlying system failure.[119-120]

Confidential Enquiries can be seen as a form of implicit review designed to determine if adverse outcomes, particularly serious harm and death, were associated with the processes of care delivery for particular specialties or procedures. Some enquiries look at all deaths within a specialty, such as obstetrics, whilst others will undertake themed reviews. In recent years the NCEPOD has undertaken investigations into deaths following in-hospital cardiac arrest, deaths in older patients, and deaths as a result of acute kidney disease.[121-123] The aim of these investigations is to make recommendations that will address identified problems, and in so doing will improve safety. These studies can be useful for generating hypotheses around the nature of preventable mortality and key contributory factors that can be subsequently tested. Wider generalisations can be limited by a lack of denominator data and controls.[77]

2.3.2.2 Global Trigger Tool

Trigger Tools are a form of explicit review in which sentinel events or 'triggers' linked to harm, are pre-specified in a list which is then used to screen a case record. The approach originated to address concerns that traditional RCRR was too resource intensive, both in terms of review time and the requirement for senior doctors to undertake assessments. In the 1970s, Jick *et al* first developed a pre-determined list of sentinel words or conditions associated with medication harm that could be used to find high risk records for further more extensive review.[124] Subsequently, Classen *et al* used this trigger list to search electronic patient records. [125] Under the auspices of the US Institute of Healthcare Innovation (IHI), Rozich went on to develop a tool which could be applied to a wider range of harms, and named it the Global Trigger Tool (GTT). [126-127]

Since its introduction, use of the GTT has spread to a number of developed countries including the UK, and new trigger lists have been developed for use in subsets of patients including those in intensive care or children, along with flexibility for customisation to suit

local needs.[128] The choice of triggers in the GTT determines the range of specific harms (e.g. cardiac arrest, surgical site infection or medication-related harm), it can be used to measure.[129-130] It is designed to find harm related to acts of commission, and generally those that are clinically focused. As yet it has not been adapted for use in patients that die during admission.

The GTT approach involves a 20 minute screen of the case records, searching for up to 32 triggers including cardiac arrest, prescription of naloxone, or deep vein thrombosis, followed by a limited review of the case record by a doctor if any triggers are found to allow confirmation of harm. Nurses usually undertake the initial screening review, but other staff, including clinical audit personnel, can be trained to do so. When compared to patient safety indicators and incident reporting systems, GTT appears to be more efficient in finding harm events confirmed by case record review, finding up to 10 times more such events.[131-132] Care has to be taken that hospitals do not confuse triggers with actual harm when tracking trends. IHI introduced the tool for internal quality assessments only and not for comparisons of harm between hospitals, due to limits in the specificity and sensitivity of some of the triggers.[34 133]

2.3.2.3 Claims Files

Closed claims files contain a range of information including case record extracts, patient, relative, clinician and lawyer statements, and descriptions of final judgements. Interest in the potential of information collected as part of medical negligence claims to identify patient harm has been apparent since the 1980s, when the American Society of Anaesthesiologists set up the first database of information from all closed claims related to the specialty.[134-135] Using data on the nature of problems in care, subsequent harm and contributory factors, problems causing serious harm related to intubation, equipment misuse and nerve injury were brought to the attention of the profession.[135] As well as a source of information about

serious harm including death, closed claims files pick up system factors underlying harm. In the case of the American Society of Anaesthesiologists, this enabled the development of new safety standards.[136] They can be most useful in identifying rare causes of harm such as retained medical instruments or wrong site surgery.[137-138] Analyses of claims records have also highlighted the role of missed or wrong diagnoses in generating serious harm, and the contribution to this harm of unsupervised patient assessments by junior doctors.[139] One group studying drug related adverse events leading to claims found that 73% of the events identified were preventable.[140] Measuring harm using claims files can be hindered by their relative inaccessibility. In addition, more claims are related to surgery, orthopaedics, obstetrics and accident and emergency than other specialties, and there is a bias towards the more serious forms of harm.[141] This may be related to the fact that harm in these circumstances is more visible to both the patient, relatives and clinical teams, thus prompting more litigation, against a backdrop of the relatively small proportion of patients harmed by healthcare who take legal action.[142]

2.3.2.4 Prospective Surveillance

Based on approaches that have been used to track hospital acquired infections and surgical complications, prospective methods for identifying other errors and harm have been developed.[101 143] These methods, which often triangulate direct observation, interaction with staff, and review of records, have the advantage over retrospective methods of being able to estimate incidence. Generally confined to research because of resource intensiveness, a recent study from Canada has explored the possibility of using this approach for the routine measurement of harm on hospital wards. [144] Nurse researchers gathered reports of harm from staff, case records or during periods of direct observation on four different specialty wards (general internal medicine, obstetrics, intensive care and cardiac surgery intensive care) across a single hospital. The approach not only identified more diagnostic and therapeutic type harm events than traditional RCRR, but was also able to elucidate different types of

errors underlying harm in each ward. This finding opened up the potential to better tailor interventions to improve patient safety. Two previous studies comparing prospective methods to RCRR found that the methods identified both similar numbers and severity of harm events.[145-146] Prospective approaches have the advantage of being able to collect more information on immediate and wider contextual system factors through interactions with staff, which is helpful in supporting decisions around preventability. However, investigator training is critical to the success of this method and even then, rare or more slowly emerging harms may be missed.[144] [147]

2.3.3 Research Orientation

2.3.3.1 Direct Observation

Qualitative ethnographic methods based on direct observation have been used to measure error and harm associated with particular tasks such as drug administration,[148] or in particular settings such as intensive care,[149] or accident and emergency.[150] The approach can be useful in providing a picture of communication patterns or teamwork issues and other underlying contributory factors.[151] Direct observation works best for detecting harm when care tasks are predictable, or where staff have defined roles. Structured data collection tools improve reliability, especially if more than one observer is involved in data collection, and training is also vital in maximising the identification of errors and harm.[152]

2.3.3.2 Patient Reported Harm

Interest in patient reporting of errors and healthcare related harm initially developed in the US, when consumer questionnaires completed after hospitalisation revealed just how commonly patients identified such events.[153] Weingert *et al* interviewed 228 adult patients during admission to a general medicine ward in a US teaching hospital, and by telephone ten days after discharge. Patients were asked about problems, mistakes or injuries they had experienced. Eight per cent of patients reported harm and, in a quarter of cases, this harm

had serious consequences. Only half of the 20 harm events identified were documented in the case record, and none were documented in the hospital incident reporting system.[154] In a similar study in England, 80 inpatients reported an average of three patient safety incidents each, encompassing both clinical problems and those associated with the organisation of care such as communication. Overall 83% of these adverse events were found in their records.[155] Forster *et al* interviewed 400 patients after discharge, and combined their narratives with discharge summaries and laboratory reports to form case histories. These histories were reviewed for adverse events by physicians, and 19% of patients were found to have experienced such an event.[156] Patients report similar rates of hospital acquired infection, pressure ulcers and drug errors to those found by RCRR studies.[157] Methods used to gather information from patients include surveys (phone or web), face to face interviews, and patients' own written reports. Face to face interviews using open ended questions and conducted during or soon after discharge have been found to elicit more reports of adverse events than those conducted by telephone, or several months after the admission.[158] However, health professionals' attitudes and the interview setting have a strong influence on how patients feel about reporting harm.[159-160]

2.3.3.3 Case Control Studies

Case control studies compare cases with an outcome of interest, such as healthcare related harm or death, with those that do not and, in doing so, determine if there are differences in particular factors between cases and controls. These studies have been traditionally used to explore the relationships between harm associated with drug events or hospital acquired infection, and lengths of stay or costs.[161-162] A small number of studies have looked at the relationship between death and specific harms, such as hospital acquired infection or a wider range of adverse events, in order to identify the relative risk of adverse events in hospital patients who die compared to those discharged alive.[163-164]

Garcia-Martin and colleagues examined the case records of 529 adult patients that had died during hospital admission, and a similar number of controls who were discharged alive. Cases and controls were matched on date of admission and primary diagnosis. The reviews identified adverse events that occurred during hospitalisation, which were defined as problems in care potentially related to clinical or organisational aspects of management rather than underlying disease. Case and controls were similar in terms of patient sex, length of stay and quality of medical records, but cases were older and admitted with more severe illness. The study found a significantly increased number of adverse events amongst cases (57.1%) compared to controls (42.9%). The presence of at least one adverse event was significantly associated with a 60% increased risk of death. This relationship was stronger in patients with lengths of stays of more than 48 hours, than for those with shorter stays. As part of the analysis undertaken of data collected from a Dutch RCRR study Zegers *et al* compared the proportions and types of adverse events found in patients who died and those discharged alive. They found twice as many adverse events in those that died (10.7% (95% CI 9.8-11.7) vs. 5.4% (95% CI 4.7-6.1)), with the proportion of harm events linked to diagnostic processes, medication and clinical management also higher, and those linked to surgical or other technical procedures lower in the deceased.[165]

This design is particularly useful for identifying the risk factors associated with rare outcomes, such as retained instruments. Using closed claims files, Gawande *et al* identified cases where a retained instrument or swab had been found after surgery and matched these cases with controls.[166] They were able to identify a number of risk factors for retained instruments and swabs including emergency surgery, a change of plan during surgery, and high body mass index. Issues of selection bias and confounding such as quality of care threaten the validity of the case control method in patient safety research.

2.4 The Epidemiology of Harm

RCRR has been the main method used for epidemiological investigations of healthcare-related patient harm over the past 50 years. Table 1 shows the findings from the largest of these studies. Across patients, the majority of whom were discharged alive, the proportion of harm events has varied widely from 3.2% (Utah and Colorado) to 16% (Quality in Australian Healthcare Study (QAHCS)). A systematic review of the first eight major RCRRs, covering 74,485 randomly selected patients, found a median proportion of 9.2% patients experiencing adverse events, with 43.5% of these events classified as preventable.[48] These statistics are similar to the figures found by the two largest English RCRRs to date. The variation in proportion of harm found between studies is likely to be related to differences in how the RCRR methodology is applied. A comparison of the Utah and Colorado and QAHCS studies showed differences in sample selection (the inclusion or exclusion of pre-admission harm), and in the cut off levels used by reviewers to determine levels of severity and preventability.[167] When records from the QAHCS study were reanalysed taking the Utah and Colorado approach, the proportion of adverse events in that study fell from 16.6% to 10.6%.[167]

The systematic review found that whilst more than half of patients experience minimal harm (56.3%) as a result of an adverse event, 7% sustained permanent disability, and in a further 7% the adverse event was associated with death.[48] Across studies, harm associated with surgical operations was the largest category (39.6%) followed by drug-related events (15.1%). Drug, diagnostic and management related harm tend to have higher preventability. Older people appear more likely to suffer harm during hospital admission, and this harm is more likely to be severe or result in death.[168] This phenomenon seems to be linked to both complexity of condition and longer lengths of stay. Exploring underlying causes of harm, both the HMPS and the QAHCS study found omissions in care, were more common than errors of commission,.[169-170] Omissions were also found to be more often preventable.

Hutchinson *et al*, in a UK based RCRR, identified the greater impact on harm generation of multiple small omissions spanning the care pathway, rather than via single acts of commission.[171]

It is important to acknowledge that in these studies, relatively few records reviewed were from patients who died during their admission, as death is an infrequent outcome (around 2% of all admissions[172]). This reduces the reliability of any extrapolations made from this subsample of cases. Furthermore, traditional RCRR studies were not designed to make a specific determination of whether any deaths were preventable. Nonetheless, sub analysis of the causes of harm associated with severe outcomes, including death, from these studies can be helpful in giving an impression of the likely causes of preventable death. This analysis reveals that diagnostic problems are the most common cause of harm in patients who die during admission, in contrast to harm resulting from technical problems related to surgery, the most common underlying cause in patients who leave hospital alive.[170 173] More recently, a Dutch RCRR study of 8,415 cases records which purposively oversampled patients who had died (3983 patients), confirmed this finding. The researchers found that the proportion of patients who died having experienced harm due to problems with the diagnostic process was twice the proportion compared to those discharged alive (14.8% vs. 6.3%).[174]

Table 2.1 Summary of findings on harm of the largest international RCRR studies

Name	Harvard Medical Practice Study[6]	Quality in Australian Healthcare Study[169]	Utah and Colorado Study[12]	Vincent *et al* Study (London)[15]	Adverse Events in New Zealand Public Hospitals[175]	Canadian Adverse Events Study[176]	Danish Adverse Events Study[177]	Sari *et al* (York)[178]	Dutch Adverse Event Study[174]	Dutch Adverse Event Study (follow up)[179]
Country	US	Australia	US	England	N. Zealand	Canada	Denmark	England	Netherlands	Netherlands
Year	1984	1992	1992	1998	1998	2000	2000	2005	2005	2008
Sampling	Representative	Representative	Representative	4 specialties	Representative	Representative	Representative	Random	Random	Random
No. of cases reviewed	30,121	14,179	14,700	1,014	6579	3745	1097	1006	8400	4023
% patients with adverse event	3.8	16.6	3.9	10.8	11.2	6.8	9.0	8.7	4.1	6.2
% patients with adverse events that were preventable	1.0	8.5	0.9	5.2	4.8	2.8	3.6	2.7	1.8	1.6
% of all adverse events that were minimal harm	56.8	45.6	53.3	66.4	7.9	61.6	73.8	56	56.8	86.6
% of all adverse events that were moderate harm	16.5	29.8	31.6	19.1	3.5	19.0		21.0	30.4	
% of all adverse events that were severe harm	6.5	13.4	8.4	6.4	0.3	10.2	26.2	11	5.0	4.8
% of all adverse events that were associated with death	13.6	4.8	6.6	8.2	0.6	4.5		10	7.8	8.6
% patients admitted experiencing preventable death (where calculable)	0.3	0.55	0.12		0.36	0.13			0.7 overall 4.1% (for deaths)	

2.5 Key Issues in Harm Measurement

Two key requirements for any measure employed to estimate preventable hospital deaths are firstly that its scope is broad enough to detect problems in care contributing to such deaths across the full range of adult acute inpatients, and secondly that it collects enough contextual information to allow the assessment of preventability. Judgement of harm measures against these two criteria are discussed in the next section, along with judgement against additional criteria that have been derived from work completed by the US Institute of Medicine and WHO, on technical specifications for quality of care indicators.[72-73] Findings are summarised in Table 2.

2.5.1 Scope of Harm Measured

Different harm measures identify different harms.[37] This may be as a result of design. For instance approaches such the GTT, where measurement is based on the identification of specified events or triggers, will not be useful in identifying omissions in care.[127] Even though prospective surveillance and RCRR find the broadest range of harms, harms are more likely to be clinical and related to individual error, than systems-based.[119-120] This is also a feature of Mortality and Morbidity meetings.[115] Prospective surveillance will tend to miss harms that are slowly emerging, and if based on direct observation, are necessarily narrow in focus, while incident reporting systems tend to under-report severe harm.[110 148 180-181] Given the variety of healthcare harm it is unlikely that the full spectrum can be captured by a single measure. Different methods identify different numbers and types of harm, with surprisingly limited overlap between sources.[145-146 178 181] From a single organisation's point of view, triangulating information from a variety of sources is likely to provide the best insights into harm.

The scope of harm measured will also be determined by the definitions used. The majority of studies have defined patient harm as an adverse event encompassing 'unintended injuries

or complications which lead to longer lengths of stay, disability at the time of discharge or death and are caused by healthcare management rather than by the patient's underlying disease'.[6 12 169] Use of this definition allowed traditional RCRR studies to take a broad view of harm, inclusive of the range one finds in hospitals. Although this definition has the advantage of being well-known and in widespread use, which can be helpful when making comparisons, it tends to draw reviewers' attention towards 'discrete' harm events.[171] This is a limitation in light of current theories of harm generation that recognise the key role of multiple small omissions in harm generation.[171] Poor clinical decisions, delays in key investigations, or inadequate monitoring can combine across an admission to cause harm, especially in the frail and sick whose defences against such small insults are not as robust as those of younger, fitter patients.[182] The term 'cascade iatrogenesis' has been coined in the US to describe this process.[183]

In some early RCRR studies, harm had to be accompanied by evidence of healthcare falling below an acceptable standard in order to be counted (similar to definitions used in the medicolegal world as part of the determination of negligence) .[6] In others, the interpretation of harm was much broader, encompassing the consequences of known complications of appropriately delivered care such as unpredictable adverse drug reactions or surgical complications, as well as minor and major errors.[169] The consequence of using narrow definitions is highlighted by a study in a Massachusetts hospital. When record reviewers were oriented towards identifying cases of where processes fell below accepted standards, they found a rate of harm of 2.7%. However when the same records were examined using a different approach that recorded any harm from healthcare, the rate rose to 11%.[167 184]

2.5.2 Judgment of Preventability

It has been argued that if the main purpose of harm measurement is to improve safety, then the focus of such measurement should be on the preventable element of harm.[173]

Preventability in this context would be the degree to which harm could be avoided if different healthcare processes were employed. However, such a focus on preventability could lead to reduced efforts to decrease the risks to patients posed by appropriately delivered care such as complications of certain surgical procedures and thus stall the development of interventions that might tackle such high risk areas.

Preventability is not absolute, and is usually measured with an interval scale to reflect its probabilistic nature. A six point Likert scale is the norm, where '1' represents 'virtually no evidence of preventability' and '6' represents 'virtually certain evidence of preventability'. The cut-off point for preventability of harm on the Likert scale can be a reflection of decisions by researchers to take either a broader or a narrower view of harm measurement. Most RCRRs, emulating the HMPS, have used the cut-off point of '4' which is defined as 'preventability more likely than not, more than 50:50'. On the other hand, studies including QAHCS and the Adverse Events in New Zealand Public Hospitals study, elected to use the lower cut-off point of '2' ('slight evidence for preventability'). As a consequence, the QAHCS found approximately half of all harm events detected were deemed at least partly preventable, compared to one third in the HMPS.[169] The explicit aim of both the QAHCS and the New Zealand studies was to initiate quality improvement, and therefore these studies aimed to capture the wider harm burden, even those harms with a limited chance of preventability in current times.

The estimation of preventability depends on the availability of adequate information about the context in which harm occurs. The richest source of contextual detail comes from the direct observation of healthcare processes. However, direct observation is resource intensive, which can lead to a narrowing of its focus. Written accounts from case records can provide a reasonable substitute but may not contain the level of detail required. For some measures, such as those derived from routine administrative data or incident reports, the lack of contextual information hinders a direct assessment of preventability and for others, such as

GTT, collecting this information is not a requirement of the method.[127] Assessment of preventability of hospital deaths can be particularly difficult, as many patients who die in hospital are elderly and frail with complex conditions and multiple co-morbidities.[17] The contribution to a death by a problem in care can be challenging to separate from the underlying illness, with a large number of such patients ultimately dying whether or not a problem in care had occurred. Measures that include the largest amount of contextual information, such as prospective surveillance or RCRR, are likely to be the best methods for untangling the complex interplay of factors contributing to patient death and allowing such judgements to be made.[17 37 185]

2.5.3 Validity

The validity of a measure is the extent to which it measures what it is supposed to measure.[186] Three key categories of validity important in harm measurement are: face validity, construct validity, and criterion validity:[187] (see Box 2)

Box 2.1: Definitions of validity categories applied to harm

Face Validity: Is there a consensus amongst users and experts that the measure covers an important dimension of patient safety?

Construct Validity: Is there evidence of expected relationships between the measure and other relevant or linked factors?

Criterion Validity: Is there evidence that the measure is associated with other measures of patient safety?

Adapted from the WHO Performance Assessment Tool for Quality Improvement in Hospitals (PATH) Implementation Manual, p10: WHO Europe, 2006 http://www.pathqualityproject.eu/upLoad/file/path_implementation_manual_2006.pdf Accessed 4th January 2013

With regard to face validity, Walshe surveyed 150 doctors working in clinical medicine and public health in the UK, and found broad support for the use of adverse event measurement to monitor safety.[188] Preventable deaths, being at the most serious end of healthcare-related harm, are seen as undoubtedly worthwhile to measure by patients, the public and politicians. Acute hospital Trusts have been extending the use of mortality reviews beyond traditional Mortality and Morbidity meetings in recent years as an approach to identifying high risk areas,[189] and more recently the Department of Health has been exploring the potential of creating a national safety indicator of avoidable deaths based on these reviews.[67]

To test construct validity, one would expect to find that patient harm is more likely to occur in patients admitted as emergencies, those undergoing complex procedures, and those with complicated underlying conditions, and as a consequence, such patients would be subject to longer lengths of stays and higher treatment costs. A range of studies of different harm measures has shown that these relationships hold true. The HMPS study found that harm was more common in the elderly, and a further study indicated that older patients were more likely to experience permanent disability or death from harm than younger patients.[168][170] Moreover harm does occur more commonly in patients admitted as emergencies.[190] International RCRR studies have provided estimations of the impact of harm on length of stay and cost of care. The QAHCS estimated that patients exposed to harm stayed an average of 7.1 extra days in hospital as a result.[169] Following a study of harm across 1014 London inpatients in two acute hospitals, Vincent *et al* estimated that the NHS was spending up to £1 billion per year on three million additional bed days, as a consequence of harm.[15] Zhan and Miller also showed an association between the AHRQ PSIs and increased lengths of stay and extra costs.[191] Postoperative sepsis alone accounted for an average of 10 extra bed days and $57,727 in extra charges per patient. Raleigh *et al* applied the PSI to English NHS inpatient data, and confirmed that patients found to have these indicators had longer lengths of stay (the stays ranged from 0.2 days for obstetric trauma to 17.1 days for

postoperative hip fracture, p<0.001) and increased mortality (the rates ranged from 5.7% for infections due to medical care to 27.1% for postoperative sepsis, p<0.001).[105]

Turning to criterion validity, there have been relatively few correlation studies looking at the relationships between harm measures, and those that do exist are limited in scope. These studies have provided evidence for both positive, negative and no correlation. The largest collection of studies focus on measures derived from routine hospital administration data, such as HSMRs and PSI. Surprisingly few correlation studies exist which compare harm identified by RCRR with other harm measures, despite RCRR often being regarded as the gold standard measure for harm. Four North American studies compared HSMR with preventable deaths identified through RCRR. Three studies either found no correlation,[21-22] or a non-significant negative correlation.[24] One study (the smallest) found a statistically significant association between hospitals with a high HSMR and preventable deaths for one out of three medical conditions (pneumonia but not stroke or acute myocardial infarction) (5.70% vs. 3.25%, p<0.05).[23]

For PSIs, most studies have been conducted as part of indicator validation. In the Complications Screening Programme, a US study, cases flagged as positive for one or more indicators underwent record review, in order to confirm the presence of harm. Indicators were found to be better at identifying surgical rather than medical complications with relatively high specificities but low sensitivities. Only 27% of the medical complications identified were found to be harm events, and just 16% were due to problems with the care provided.[192] Two PSIs are focused specifically on death. 'Failure to rescue' and 'deaths in low mortality diagnosis related groups (DRGs)'. 'Failure to rescue' has been shown to be associated with HSMRs and with other structural measures, such as the bed to nurse ratio, trained nurse mix, hospital size and status, and quality of care measures amongst Medicare patients.[193-194] Evidence remains patchy as to the validity of 'deaths in low mortality DRGs' as an indicator of safety problems due to low sensitivity.[195] As a confirmation of the

potential of 'failure to rescue' for use as an indicator of harm, Isaac *et al* found it to be the only indicator out of four (including deaths in low mortality diagnosis related groups, decubitus ulcer, and infections related to intravenous lines and catheters) to be positively correlated with quality of care variables and disease specific death rates.[194]

The ability of any measurement tool to measure what it is supposed to measure is limited by other systematic factors that cause deviation from true measurement. The validity of the findings from the Mortality and Morbidity review process are compromised by the small numbers of cases that are examined, case selection bias, and lack of involvement of all members of the multidisciplinary team. [115 196-197] Incident reporting is well known to offer a poor reflection of overall harm and to miss serious harm due to under-reporting, particularly by doctors.[178 20] Time pressures, perceptions that notification will not lead to any improvement and fear of blame have been found to play a role in this under-reporting.[198-199] Information in claims files can lack objectivity, and is found to often reflect past care practices due to the length of time it takes for claims to be resolved.[139] Direct observation will miss harm that occurs when the observer is not present and also those harms that take a longer time to evolve.[152] All measures dependent on interrogation of the written record will be vulnerable to information bias if the record is not complete. Finally, any retrospective method will be prone to hindsight bias, whereby a reviewer's judgment is altered by prior knowledge of the outcome. These biases are discussed in more detail in the following sections.

2.5.4 Reliability

Reliability is dependent on the explicitness of measure specification, uniformity of data collection, and the consistency of judgements.[187] The explicitness of trigger definition is particularly important in determining the sensitivity and specificity of the GTT. Being dependent on explicit screening criteria allows a uniformity of data collection when these

tools are used by different professional groups.[200] Structured data collection forms have also been found to improve the reliability of harm measurement in RCRR and direct observation.[152 201]

The uniformity of data collection within incident reporting systems hampers wider use of this dataset for tracking harm. Studies show variable levels of reporting to such systems.[202] Under reporting is found at both ends of the harm spectrum, and amongst doctors to a greater degree than nurses.[79] Prompted incident reporting, either by direct verbal encouragement or via electronic prompts, has led to an increase in adverse event reporting by junior doctors.[145 203] An additional issue is the level of inaccuracy in classification and coding of self-reported submissions, particularly when attempting to code levels of harm.[81]

The consistency of judgements around the origins of patient harm and its preventability are captured in the measurement of inter-rater reliability, or the degree to which measures taken are reproducible between two different reviewers.[187] The Cohen Kappa statistic is normally used to assess inter-rater reliability in these studies, and is categorised into poor (K<0.2), fair (K>0.21 to K<0.4), moderate (K.0.41 to K<0.6), substantial (K>0.61 to K<0.8), and good (K>0.8).[204] Poor inter-rater reliability can result in over- or underestimation of harm. Measurement methods such as RCRR that rely on subjective judgement of complex healthcare scenarios where not all relevant information may be accessible, are prone to poor inter-rater reliability.[187] Concerns about the reliability of reviewer judgements of harm and preventability go back as far as the HMPS, which achieved moderate agreement between reviewers for the presence of an adverse event (kappa=0.57), but only weak agreement on preventability (kappa=0.24).[184] The QAHCS produced similar results with a kappa of 0.55 for adverse events identification, and 0.33 for preventability.[169] An RCRR study by Hofer of 111 hospital deaths, which focused on the identification of preventable deaths, found an inter-rater reliability for judgments of preventability of K= 0.34.[17] It is important to put these findings into context; although reliability of RCRR falls below an ideal level, it is similar to

findings seen when two reviewers are asked to interpret common diagnostic tests such as mammography.[205]

Analysis of the performance of reviewers in the HMPS indicated that agreement was higher between more experienced reviewers, and for some types of harm events (e.g. wound infections and drug side effects) than others (e.g. omitted therapy or failed diagnosis).[184] Reviewers were also found to be consistent outliers for high or low numbers of events detected. Approaches to decision making also make a difference. Consensus and majority decisions, as opposed to unanimous decisions, will identify more adverse events.[206] However, Hofer and Hayward showed that whilst discussion between a pair of reviewers increases agreement between them, when the same records are reviewed by a new pair of reviewers, there is no improvement in agreement between the two sets of reviewers.[207] Rubin found that only when five or more reviewers were used did the classification of reliability reach levels of 90% accuracy.[208] In a subsequent study, Hayward *et al* trained 12 internists to undertake an RCRR of 675 patient admissions with 20% duplicate reviews.[209] They concluded that unless resources permit the use of many reviewers per case record, the reliability of single or double review (whilst inadequate for making judgments in a single case) is adequate for comparing mean performance across wards or hospitals. The findings from a recent large Dutch RCRR study confirmed no advantage for reliability in having paired reviewers as opposed to a single reviewer.[210]

Lilford *et al* conducted a systematic review of inter-rater reliability of case record review.[211] The evidence suggested that the use of explicit rather than implicit criteria improved reliability (mean Kappa change 0.39 to 0.62). Separate studies have found that the GTT does have good inter-rater reliability for both identification of triggers and subsequent harm. However, the reliability of some individual triggers is poor, particularly those that are less well-defined or recorded in the case record.[212] In a recent report on the use of hospital records to assess quality of care, Hutchinson *et al* found inter-rater reliability to be moderate

to good for explicit criterion based measurement and good for implicit scores, as long as reviewing pairs for the implicit reviews were from the same professional group and training was provided.

Senior doctor time is an expensive resource, and RCRR can be a time consuming process. As a consequence, interest has developed in whether other healthcare personnel can be used without jeopardising reliability. Hutchinson found only weak to moderate agreement on quality of care, whether assessed by implicit or explicit review of case records with mixed reviewing teams.[200] In line with previous research, he found that during the reviews, nurses and doctors tend to identify problems related to different aspects of care.[200 213] Nurses are more likely to focus on issues such as whether routine clinical observations were completed, or the timing of medication, whereas doctors focus on the wider clinical aspects of care such as whether the correct diagnosis was reached, or whether appropriate treatment was implemented, and were more likely to give global ratings of overall quality. Hutchinson goes on to advocate a nurse-doctor pairing, if resources permit, to maximise opportunity for identifying patient harm.[200]

Ashton identified a number of ways to improve the reliability of judgements in RCRR, including the provision of guidance on the types of information a reviewer should look for, and on assessment criteria, use of a structured data collection form or algorithms developed to identify specific events.[201] Training, particularly focused on reviewing and discussing cases, has also been shown to improve reliability.[169 214] Rubenstein stressed that this should be to ensure that reviewers fully understand the process overall, definitions and rating scales rather than to change their judgements.[215] Investigator training is also critical to the success of prospective approaches, including direct observation, in identifying errors and harm.[144 147]

2.5.5 Retrospective versus Prospective Approaches

Hindsight bias is an inherent weakness of harm measures that have been designed for use in investigations after the problems of care have taken place (retrospective). Retrospective methods include RCRR, Mortality and Morbidity meetings, case control studies and reviews of claims files. Reviewers have been shown to detect more harm when the final outcome was poor, as in the case of a patient death.[17 216] Researchers have attempted to avoid this bias by masking observers to patient outcome through practices such as the re-dictation of case notes. However, this risks the omission of important information and is expensive to undertake.[54] Alternatively, training can be used to reduce hindsight bias by orientating reviewers to examine processes of care before ascertainment of harm, and by getting them to imagine 'walking in the shoes' of the patient's original clinical team as the story unfolds, rather than jumping to premature conclusions.

Retrospective measures can be conceptualised as reactive and generally identifying issues from the past that may no longer apply to current performance.[217] Prospective approaches can avoid these pitfalls. By combining activities such as daily interrogation of case records or administrative databases, interviews with staff or patients, or direct observation, these methods allow harm events to be both measured and investigated as they occur. The fact that staff are more likely to remember key details surrounding the origins of the harm if interviewed around the time of occurrence is useful for quickly grasping an understanding of underlying contributory factors, and for allowing the speedy resolution of investigations. Comparative studies have found that prospective methods are able to identify similar numbers of harm events as RCRR, in addition to finding other events and contributory factors that are not captured in the written record.[146-147] However, as well as potential to miss those harms that take longer to evolve, prospective approaches are more disruptive to ongoing ward care, can change staff behaviour, and are open to observer bias as the investigator builds relationships with staff.[147] Furthermore, for those studies with more

reliance on observation, episodes of harm will be missed when observers are absent. The high costs involved when using prospective approaches has limited their current use to date.

2.5.6 Content and Context

Missing content, whether due to failure to record details, poor legibility, sparse entries or inconsistent filing, presents the possibility of information bias. For RCRR, an incomplete record may lead to more harm missed compared to a complete record, which has the potential to result in organisations with better quality records having an apparently higher proportion of harm, particularly for severe harm at the clinical interface between doctor and patient.[120 218] Missing content is a particular issue when using claims data to investigate patient harm. Within the file, descriptions of problems in care or harm may be from a limited number of subjective perspectives. Furthermore, the long period for claim resolution provides an opportunity for key material to go astray.[141] Incident reports can also be difficult to interpret because of a lack of contextual information or details of contributory factors supplied on the reporting form.[80 219] Routine data sources and GTT collect little or no information on the context of harm, which limits their use in determining preventability.

The validity of routine hospital administrative data is challenged by the completeness and accuracy of coding within the system.[220-221] Studies have found up to 40% of diagnostic codes used in US hospital discharge data are wrong when compared to the recorded diagnoses in the case records.[222] In addition, a lack of flags to indicate if particular diagnoses were present on admission makes identifying the temporal sequence of events difficult.[223] In the late 1990s, McKee *et al* pointed out that the use of a 'failure to rescue' indicator was not feasible in the NHS due to the poor standard of secondary diagnosis coding.[224] However, a recent analysis would indicate that improvements in coding over the last decade have changed this situation, especially for surgical patients.[225]

2.5.7 Denominators

The clinical researcher Peter Pronovost, reflecting on the challenges of measuring safety, identified the absence of information on the population at risk (the denominator) as a key issue when calculating rates of harm from a number of measures including incident reports, the GTT, and claims files.[226] Although these systems can be useful for identifying rare harms when large numbers of reports are pooled, they cannot be used to make comparisons between organisations unless a denominator is created. The difficulty then arises as to which denominator is most appropriate, especially when different subpopulations of patients may have different risks of exposures to particular harms. The choice of denominator used in the calculation of rates will have a substantial impact on the outcome of such analyses be it admissions, bed days or particular healthcare processes. As rates play a key role in comparisons, it is important that the impact of using different denominators is appreciated. Focusing on the measurement of specific harms using agreed definitions and denominators that represent the patients at risk of such harms is probably the best option. This approach has been shown to improve the specificity of PSI.[91]

2.5.8 Sampling

Some harm measures are applied to whole patient populations but, depending on the resource intensiveness of the measure, this may not always be possible. Random sampling will allow generalisation of findings to other settings; the use of stratification ensures that relevant subpopulations are represented, and that populations sampled reflect the range of patients receiving treatment. It should be acknowledged that random samples may miss rare harm events. When the aim of the investigation is local quality improvement, rather than external comparisons, purposive sampling such as examining the records of consecutive patients admitted to a certain area will be adequate.

Sampling can be adjusted to improve the yield of harm events. Findings from previous RCRRs have indicated that higher levels of problems in care resulting in harm are found in patients who die.[174] In Wales, where hospital mortality review is well established, a range of techniques have been used to target the identification of deaths with higher risks of harm, namely targeting mortality review at deaths in low risk diagnostic categories, deaths identified using the Institute for Healthcare Innovations Categorisation Matrix (which categorises deaths according to whether they are a palliative care death, and an expected or unexpected death) and deaths in specialties with high standardised mortality ratios.[227] However, although this approach might be useful where resources are stretched, it is likely to uncover a relatively small fraction of preventable deaths, as the majority of these deaths are still to be found in low risk groups.

Clarity of inclusion and exclusion criteria for a sample is an important consideration if there is an intention of making comparisons with the findings from previous studies. A study that investigated the differences in proportion of harm in the Utah and Colorado RCRR[12] (3.2%) and QAHCS[169] (16.6%) RCRR studies found that the US study counted only those adverse events that led to the index admission or those which occurred and were discovered during that admission. On the other hand, QAHCS counted events whether they occurred before, during or were discovered after the index admissions.[167] Alternatively, a large Dutch RCRR included adverse events with their origins before admission only if they were associated with a previous admission to hospital.[174]

Sampling also needs to consider the impact of inclusion of certain population subgroups. All previous RCRR studies have excluded psychiatric patients, and a number have also excluded obstetric patients and children under 18 years.[174 176] Justification for exclusion of the latter groups was based on the fact that these specialties are responsible for very few hospital deaths and including them requires the recruitment of specialist rather than generalist reviewers. When drawing samples for cases and controls in case control studies,

consideration of patient characteristics is vital. Failure to adequately match on characteristics associated with harm, such as severity of illness, will lead to biased findings.

2.5.9 Case Mix Adjustment

As patient characteristics such as age, co-morbidities or disease severity are independently associated with harm outcomes such as death, case mix adjustment then becomes important for comparing such outcomes across organisations. Indicators of harm derived from routine administrative data such as HSMRs or PSIs are particularly vulnerable to criticisms regarding the adequacy of case mix adjustment, especially as essential information for stratifying risk, such as severity of primary and secondary diagnoses, is not adequately captured in algorithms used to derive the measures.[228] Complexities involved when comparing organisations' harm rates have led to suggestions that the focus should instead be on comparison of processes of care.[54] However, scientific evidence linking processes, such as skin marking prior to operation, with outcomes, such as wrong site surgery, often does not exist.[226]

For RCRR, the appropriateness of the healthcare the patient receives is assessed before judgements of harm occurrence are made. As such decisions on standards of healthcare processes are independent of patient age and co-morbidity, the need for case mix adjustment is avoided. However, Richard Lilford and colleagues have recently challenged the assumption that error and harm rates can be measured without consideration of the opportunity for error or harm (e.g. opportunity for harm would be more for longer lengths of stay which are themselves case mix dependent).[54] These researchers have proposed that opportunity for error or harm would be a more appropriate denominator. This proposal remains theoretical, as other academics have challenged the feasibility of discerning a set of opportunities for error or harm in patients with complex conditions given the normal level of detail found in patient records.[200]

Table 2.2 Comparison of harm measures

	Adverse Event Misadventure Codes	Patient Safety Indicators	Hospital Standardised Mortality Ratio	Incident Reporting	Surveillance / Direct Observation	Patient Reported Harm	Retrospective Case Record Review	Case Control	Global Trigger Tool	Mortality & Morbidity Meetings	Claims Files
Scope	Mainly related to surgery and acts of commission. No indication of degree of harm. Good for rare harms, which can be tracked over time.	Mainly related to surgery and acts of commission. Difficult to determine degree of harm.	Values above 100 interpreted as the proportion of 'excess deaths' compared to national average.	Majority of incidents are no or low harm. Patient falls form biggest category overall, followed by drug errors. Mixture of omissions and commissions, but bias towards acts of commission. Good for rare harms, which can be tracked over time.	Narrow focus on particular area or process (direct observation) or broader care processes (surveillance). Direct observation and staff interaction can capture context and contributory factors, particularly system factors such as communication or teamwork. Identifies both omissions and acts of commission and allows assessment of degree of harm.	Identifies wide spectrum of harm particularly that related to system problems such as communication or coordination. Less good at identifying diagnostic error.	Wide spectrum of harm identified, especially clinical problems. Some information on contributory factors. Finds a mixture of omissions and commissions. More omissions than most sources. Good at identifying serious harm not reported elsewhere.	Establishes risk factors associated with harm (very useful for rare harms). Determines relative risk poor outcomes associated with problems in care.	Pre-specified list of triggers may be broad or narrow. Only detects errors of commission. Identifies harm across the spectrum.	Mainly focused on diagnostic errors and technical clinical problems. Contributory factors may not be discussed. Main focus is on severe harm and preventable death	Majority of harm is related to diagnosis. Can identify some contributory factors. Claims most often result from surgery, orthopaedics, obstetrics and A&E. Mixture of omissions and commission. Good for rare harms
Judgement of Preventability		Not possible		Not reported directly. Limited inference can be drawn from mitigating actions and contributory factors	Can be determined	Patients may not be able to determine in some cases without expert input	Can be determined if adequate contextual information available	Usually does not assess	Does not assess	Can be determined	Can be determined if adequate contextual information available

68

	Adverse Event Misadventure Codes	Patient Safety Indicators	Hospital Standardised Mortality Ratio	Incident Reporting	Surveillance	Direct Observation	Patient Reported Harm	Retrospective Case Record Review	Case Control	Global Trigger Tool	Mortality & Morbidity Meetings	Claims Files
Content and Context	Lack of completeness and accuracy of coding limits adequacy of case mix adjustment. No information on context.			Reports can be difficult to interpret due to staff failure to include key contextual and contributory factors.	Triangulation across sources minimises risk of missing information. Best source for details of context as staff can be consulted.	Some events missed, as observers cannot be in place all the time. Good level of detail on context.	Information may be forgotten if there is a delay in seeking patient views. Some information on context.	Information may be missing due to failure to record, misfiling, or lost sections. Records may be illegible. Some information on context.	Information may be inadequate for case/ control matching and identification of adverse events.	Information may be missing from case records.	Missing information from case record may hamper review. Good level of detail on context.	Multiple reports from different individuals' perspective. May be missing vital information. Length of time to close a case may lead to missing information.
Validity	Harm codes are narrow in scope. Some important harms are inadequately coded, e.g. hospital acquired infection. Original data source not collected for this purpose.	Associated with longer stays and increased mortality. Poor coding limits sensitivity and specificity. Inability to identify if diagnosis present on admission decreases specificity. Source data not collected for this purpose.	Correlations with other measures of quality and safety show a mixed picture - with positive, negative and no correlations. Many factors other than quality and safety impact HSMR, including depth and style of coding, referral and discharge practices etc. Use of different	Incidents more likely to be reported if not attributable to direct staff action (such as falls). Doctors report fewer incidents than nurses, limiting capture of the full range of harm related to clinical problems.	Triangulation of three approaches maximises identification of harm. Identification of harm in real time and in its own context possible. Slowly evolving and rare harms such as preventable deaths may be missed.	May miss slowly evolving and rare harms or those that happen when the observer is absent. Difficulty detecting fleeting or simultaneous events. Resource intensiveness leads to restricted focus.	Proportion of reported harm due to pressure ulcers, drug error and hospital acquired infection harm found to be similar to RCRR. Diagnosis and monitoring harm is less visible to patients. Severity of illness, co-morbidities, socioeconomic status, health professionals'	Wide range of harms can be identified, but bias towards identification of clinical/ technological harms at doctor-patient interface. Less likely to find harm related to systems issues, as well as minor harm from drug errors or falls less likely to be recorded. Prone to hindsight bias.	Wide range of harms can be identified, but bias towards identification of clinical/ technological harms at doctor-patient interface. Less likely to find harm related to systems issues as well as minor harm from drug errors or falls as less likely to be recorded. . Prone to hindsight and selection bias. Impact of confounding factors such as	Not designed to find harm related to single or multiple omissions. Sensitivity and specificity of individual triggers varies widely, and lower for less specific triggers. Time limit on GTT process may lead to missed harm.	Lack of involvement of multi-professional team limits exploration of harm and contributory factors. Fear of blame and stigma may lead to failure to explore the full details of a case. Bias towards harm caused by individuals such as clinical error. Selection bias.	Submission of a claim by a patient does not necessarily indicate that harm has occurred. Hindsight bias.

69

	Adverse Event Misadventure Codes	Patient Safety Indicators	Hospital Standardised Mortality Ratio	Incident Reporting	Surveillance	Direct Observation	Patient Reported Harm	Retrospective Case Record Review	Case Control	Global Trigger Tool	Mortality & Morbidity Meetings	Claims Files
			HSMR algorithms leads to different hospital rankings. Incomplete case mix adjustment. Source data not collected for this purpose.				attitudes, and interview setting influence patient reporting. Selection bias in respondents possible.			length of stay.		
Reliability	Large and stable database. Poor coding.	Standardised algorithms applied to a stable database. Poor coding.		System relies on self-reported submissions, leading to inaccuracies in classification and coding especially for harm levels. Under-reporting at both ends of the harm spectrum.	Complexity of clinical scenarios may lead to poor inter-rater reliability around judgements of harm and preventability. Observers may have inherently different thresholds for judging whether harm has occurred. Observer and staff behaviour may change as the investigator builds relationships with staff.		Face to face interviews, using open-ended questions, elicit more reports. Best conducted as soon as possible after discharge to aid recall.	Complexity of clinical judgements leads to generally poor inter-rater reliability around judgements of harm and preventability. Harm and preventability definitions and thresholds can change the proportion found.	Complexity of clinical judgements leads to generally poor inter-rater reliability around judgements of harm.	Using predetermined criteria to assess processes of care increases the reliability of judgments around quality of care, errors and harm.	Small number of cases examined Lack of standard approach to information gathering	Harm identified by physician judgement. Complexity of judgements leads to poor inter-rater reliability around judgements of harm and preventability. Observers may have inherently different thresholds for judging whether harm has occurred
Timing		Retrospective		Retrospective	Prospective	Prospective	Retrospective/ Prospective	Retrospective	Retrospective	Retrospective	Retrospective	Retrospective
Denominator	Usually per admission or	Usually those	Ratio of observed to	No, but sometimes a	Usually per admission or hospital bed day.		Usually per patient. May	Usually per admission or	Depends on design.	No, but sometimes a	No denominator	No denominator

70

	Adverse Event Misadventure Codes	Patient Safety Indicators	Hospital Standardised Mortality Ratio	Incident Reporting	Surveillance	Direct Observation	Patient Reported Harm	Retrospective Case Record Review	Case Control	Global Trigger Tool	Mortality & Morbidity Meetings	Claims Files
	hospital bed day.	patients at risk of the particular PSI.	expected deaths.	denominator is created.			be restricted to population at risk.	hospital bed day.		denominator is created.		
Sampling	Can be purposive or random. Higher specificity if sample from a population at highest risk.			Purposive	Usually purposive but can be random.		Can be purposive or random.	Can be purposive or random.	Matching of cases and controls.	Can be purposive or random. Opportunity for harm differs between patients. It has been advocated by some that the denominator should be opportunity for harm.	Purposive	Can be purposive or random
Case Mix Adjustment	Incomplete case mix adjustment.			N/A	May be focused on particular specialty or sub population of patients.		May be focused on particular specialty or sub population of patients.	May be focused on particular specialty or sub population of patients.	May be focused on particular specialty or sub population of patients.	May be focused on particular specialty or sub population of patients.	May be focused on particular specialty or sub population of patients.	May be focused on particular specialty or sub population of patients.
Comparison with RCRR	No studies found	Two validation studies looked at the compared 21 AHRQ PSIs to RCRR. First study found PSIs correctly	Four US studies look at association between HSMR and preventable deaths determined by RCRR. The studies focused on selected	Findings from RCRR of 1006 records from a single English hospital compared with incident reports. RCRR found harm in 10.9% of	A US study in one teaching hospital over five months asked house staff to submit reports of harm, as soon as possible after the event.	Direct observation studies tend to focus on specific harms. In a study of 36 US hospitals and nursing homes an observer accompanie	Weingert *et al* interviewed 228 US adult patients during admission and by telephone 10 days after discharge.[154] 8% reported harm and a		Extension of RCRR method	GTT is a tool applied to case records.	M & M meetings are partly based on material in case records.	A comparison of the number of harm events found in hospital quality assurance, risk management or claims records compared to RCRR was undertaken in a

Adverse Event Misadventure Codes	Patient Safety Indicators	Hospital Standardised Mortality Ratio	Incident Reporting	Surveillance	Direct Observation	Patient Reported Harm	Retrospective Case Record Review	Case Control	Global Trigger Tool	Mortality & Morbidity Meetings	Claims Files
	identified harm in 68.4% of surgical patients and 27.2% of medical patients.[103] Second study found that of all patients screening positive for at least one indicator only 11% were confirmed as experiencing harm on RCRR.[229] Sensitivity of individual indicators ranged from 4% to 68%.	specialties (general medicine, surgery),[21-22] diseases (pneumonia, CVA, MI) [23] or interventions (CABG).[24] Three of the studies either found no correlation or a non-significant negative correlation. Only one study found a significant positive association- the proportion of preventable deaths for pneumonia was significantly higher in high HSMR hospitals. .	admissions, the incident reporting system picked up only 10% of these.[178] A study comparing findings from RCRR of 228 case records discovered that 26 harm events were found by the incident reporting system.[230]	Reports were verified by senior resident physicians via RCRR. The prospective method identified similar rates of harm to RCRR (2.8% vs. 2.7%).[145] A French study of 778 patients across seven hospitals used nurse investigators to check notes and ask staff about harm. Similar rates of harm to RCRR were found (15.4% versus14.5%).[146] A third study in one hospital across four specialist areas also included	d nurses on drug rounds and found 71% of the 35 errors thought to be clinically significant. Only 9% of these were also found by chart review. [148]	quarter of these were reports of serious harm. Only half of the 20 harm events were found on RCRR and none were found in the incident reporting system. In a similar study in UK, 80 inpatients reported an average of three patient safety incidents each including clinical and non-clinical aspects of care Patients were able to identify 83% adverse events found in their records.[231]					large teaching hospital.[232] RCRR identified 80% of the events found across all three other sources. Harm not found on RCRR tended to be associated with more minor and less costly malpractice claims. 21% of harm found on RCRR had never given rise to litigation or risk management reports, despite being judged as due to negligent care.

	Adverse Event Misadventure Codes	Patient Safety Indicators	Hospital Standardised Mortality Ratio	Incident Reporting	Surveillance	Direct Observation	Patient Reported Harm	Retrospective Case Record Review	Case Control	Global Trigger Tool	Mortality & Morbidity Meetings	Claims Files
					direct observation and found more harm related to diagnosis and treatment problems than RCRR.[144]							
Issues to consider when measuring preventable death	**Advantages** Accessible and cheap. Could be used to track trends over time. **Disadvantages** Limited number of indicators that focus on death. Coding completeness, accuracy and depth. Lack of information on temporal sequence of diagnoses. Incomplete case mix adjustment. No information on preventability. Need for follow up with RCRR to confirm preventability.			**Advantages** Possible to identify rare causes of potentially preventable deaths. Possibility of tracking trends over time. Prompted reporting might increase levels of serious harm detected. **Disadvantages** Under reporting of serious incidents and preventable deaths. Lack of	**Advantages** Surveillance identifies more diagnostic and therapeutic harm events than RCRR, which are more common in preventable deaths. More information on context and contributory factors than RCRR can improve judgements around preventability **Disadvantages** As preventable death is rare the number of patients observed or period of observation may need to be substantial. Resource required in terms of training and manpower. Direct observation has a limited focus and surveillance may not be possible on more than a few wards at a time. Easier to identify common rather than rarer harm such as preventable death. Lag time between error and	Not applicable	**Advantages** Identifies as much harm as prospective approaches. Easily assessable source of information. Generally enough information on context to make judgements on preventability. **Disadvantages** Missing information. Time consuming. Resources needed for training and manpower. Hindsight bias. Moderate	**Advantages** Allows determination of risk factors associated with harm **Disadvantages** Missing information Time consuming Resources needed for training and manpower Hindsight bias Moderate Inter-rater reliability Bias towards clinical technical harm, rather than harm due to systems problems. Need to review	**Advantages** Less resource intense than RCRR. Useful for screening the case notes and identifying high risk populations. **Disadvantages** Poor recognition of harm due to omissions, or complex interacting problems which are common in preventable deaths. Will miss harm if occurred out with the 20	**Advantages** Natural focus on severe harm and death. Able to get input on context and contributory factors from multidisciplinary team **Disadvantages** Small number of cases discussed. Selection bias. Failure of healthcare professionals to disclose full details of case.	**Advantages** Files contain information on severe preventable harm, particularly around diagnosis. **Disadvantages** Long delays in closure of claims limit access. Often have missing information which makes judgements around preventability limited. Lack of verification of subjective reports. Hindsight bias Selection bias

Adverse Event Misadventure Codes	Patient Safety Indicators	Hospital Standardised Mortality Ratio	Incident Reporting	Surveillance	Direct Observation	Patient Reported Harm	Retrospective Case Record Review	Case Control	Global Trigger Tool	Mortality & Morbidity Meetings	Claims Files
			denominator data prevents establishment of rates and comparisons. Limited involvement of medical staff in reporting. Relatively fewer reports of technical clinical problems.	harm. Risks of observer bias.			Inter-rater reliability. Bias towards clinical technical harm, rather than harm due to systems problems.	cases and controls, so is more expensive	minutes allowed for case review. Does not collect information on preventability of harm.		as claims are only made by a small proportion of all patients harmed by healthcare.

2. 6 Comparison of Harm Measures for the Measurement of Preventable Deaths

Though there are a number of measures that could be used to assess preventable deaths, for the majority there are significant limitations to consider. Mortality and Morbidity meetings and claims files will only capture small numbers of cases, and are subject to selection bias. Explicit reviews such as GTT are limited to detecting those harms that are pre-specified, rely on case record review to confirm preventability and have not been designed for use in patients that die. Use of incident reporting systems is restricted by under reporting of severe harm.

Poor coding and difficulties establishing the temporal sequence of events complicate the gauging of preventable death from routine data. HSMR was heralded as an easily accessible measure that could provide an indication of the number of 'excess deaths' occurring in a hospital, and be used for hospital comparisons. However, a recent study has shown that four of the commonly used algorithms for calculating HSMR provide different rankings when applied to a single population of acute hospital patients in Massachusetts.[233] Lilford *et al* have used modelling to estimate that preventable mortality would have to be 15% or more, to enable HSMRs to have a predictive value of even 30%.[234] Moreover it has been observed that, amongst the few publications that show a decrease in standardised mortality ratios linked to interventions targeted at preventable deaths, these changes were in fact more likely to be the result of regression to the mean and concurrent changes in coding, than the interventions themselves.[52] Some measures hold more promise as potential screening mechanisms for preventable deaths in the future. With improvements in methods for data capture, coding and the sophistication of the algorithms, routine data and GTT might offer an efficient approach to screening large numbers of records and identifying high-risk cases for review, but this has yet to be demonstrated.

Comparative studies have shown that prospective surveillance and RCRR find more severe harm than other methods, and are the approaches most suited to capturing preventable deaths.[145-146] Three studies specifically compared RCRR and prospective methods for measuring harm across a broad spectrum of patients. [144-146] The studies were conducted within acute hospitals and used similar definitions of harm, and a similar two stage RCRR process, to the original HMPS.[4] In the first study from the US, 3,141 medical admissions to a major teaching hospital were examined over a five month period. For the prospective component, house staff were asked to gather information on patient harm soon as possible after an event.[145] Team leaders reminded junior staff on a daily basis. A senior resident physician verified if harm had occurred, through interaction with junior staff and by reading their reports. The prospective method identified adverse events in 2.8% of patients, which was similar to the percentage identified by RCRR (2.7%). Both prospective recording of harm and RCRR methods identified similar types of events with similar distributions across levels of severity. In a later French study covering 778 patients (278 medical, 263 surgical and 237 obstetric) across seven hospitals over a two month period, nurse investigators visited wards each week, and identified adverse events through interaction with ward staff and monitoring of case notes.[146] Any harm detected was verified by a physician, who also visited the wards on a regular basis. Again, levels of detection were similar (15.4% for prospective surveillance versus 14.5% for RCRR). The prospective method was more effective at identifying events in medical cases than surgical cases. A third study across four specialist areas in one hospital, which used direct observation alongside the questioning of staff and record review for the prospective approach, found more harm related to diagnosis and treatment problems by these methods compared to RCRR.[144]

Prospective surveillance and RCRR identify similar numbers of harm events, both omissions and commissions. Each method allows collection of contextual information to enable judgements around the preventability of deaths to be made. However, by triangulating up to three different sources of data (direct observation, interaction with staff, and daily case

record review) prospective measures can enable a clearer understanding of this context and the contributory factors / intervening variables integral to the origins of harm, making judgements of preventability easier than for RCRR where such information may be missing. Prospective surveillance also has the advantage of allowing estimation of incidence and is timelier than retrospective methods.

Although the two methods compare favourably on the scope and scale of harm identified, prospective surveillance is at a disadvantage when it comes to measuring preventable deaths. Being a rare event, either a large number of patients have to be monitored prospectively, or the period of surveillance needs to be extensive to ensure adequate numbers of cases are captured. In addition, it is more labour intensive than RCRR, particularly if direct observation is undertaken. Given these constraints and their impact on cost, RCRR would seem the more pragmatic choice allowing a large number of deaths to be examined in a relatively short time frame with little disruption to services. A case control design could be seen as an enhancement of the RCRR method. It would allow not only the determination of the proportion of preventable deaths due to problems in care, but also the determination of the risk of problems in care amongst those that died compared to those that were discharged alive. However, both the technical challenges of selection bias and dealing with confounding, and the added cost of doubling the number of reviews that would need to be undertaken, are key disadvantages to adopting this approach.

2.7 Conclusion

An understanding of how harm is generated has led to the development of a range of approaches to its measurement. RCRR was the method developed for the first large scale epidemiological studies of healthcare-related harm. These studies have shed light on the burden of severe harm and preventable death resulting from healthcare errors. The most

commonly used harm measures have been described and compared against a range of criteria to determine their relative strengths and weaknesses. RCRR compares favourably with other approaches, particularly when used to measure serious harm and preventable death, and has the advantage of allowing comparisons with previous studies. The method has three main weaknesses; missing information, poor to moderate inter-rater reliability,[211] and hindsight bias.[216] Actions to mitigate the latter two, such as the use of structured data collection tools and reviewer training programmes, have been developed in previous studies and can be used to improve the reliability and validity of measurements.

Chapter 3 Research Paper 1

3.1 Introduction to Research Paper 1

This first research paper was published in the *Quality and Safety in Healthcare* journal, which later became *BMJ Quality and Safety*. It presents a study to identify the range of measures available to assess harm in a typical acute NHS hospital in England. The method includes semi-structured interviews with staff, direct examination of data sources and policy documents, and attendance at clinical governance meetings at one acute hospital.

The study found that there was a wide range of information on patient harm available, but poor quality of coding, delays in collection, narrow scope, time limited data collection, and lack of central collation were barriers to using many of these sources. The paper looked in particular at seven sources of data that had the potential to provide hospital-wide information on harm, namely the five databases held in the Trust's local risk management system (clinical incidents, health and safety incidents, complaints, claims, and inquests), the patient administration system (PAS), and case records. These sources were found both to identify different kinds of harm, and harm of differing levels of severity with little overlap between sources. The study highlighted some of the issues to be considered when using different types of harm measurement, such as a lack of temporal sequencing available in PAS data, and the subjective nature of information on harm held in claims files. RCRR was found to identify the largest number of harm events with the broadest scope.

This study was important in developing an understanding of harm measurement and, in particular, some of the limitations associated with different measures. It was influential in guiding the choice of methodology for the main study.

COVER SHEET FOR EACH 'RESEARCH PAPER' INCLUDED IN A RESEARCH THESIS

Please be aware that one cover sheet must be completed <u>for each</u> 'Research Paper' included in a thesis.

1. **For a 'research paper' already published**
   Title: What can we learn about patient safety from information sources within an acute hospital?

   H. Hogan, S. Olsen, S. Scobie, E. Chapman, R. Sachs, M. McKee, C. Vincent, R. Thomson

   1.1. Where was the work published? ..Quality and Safety in Healthcare

   1.2. When was the work published? ....2008.

      1.2.1. If the work was published prior to registration for your research degree, give a brief rationale for its inclusion
      The work described in this paper made an important contribution to my understanding of the scope of healthcare-related harm measures in acute hospitals and their advantages and disadvantages. During the study, I undertook a retrospective case record review and gained appreciation of how the method could be applied in future studies.

   1.3. Was the work subject to academic peer review? ....Yes

   1.4. Have you retained the copyright for the work? No (evidence of permission from copyright holder included)

      If yes, please attach evidence of retention.
      If no, or if the work is being included in its published format, please attach evidence of permission from copyright holder (publisher or other author) to include work

2. **For a 'research paper' prepared for publication but not yet published**

   2.1. Where is the work intended to be published? ....

   2.2. Please list the paper's authors in the intended authorship order

   2.3. Stage of publication – Not yet submitted / Submitted / Undergoing revision from peer reviewers' comments / In press

3. **For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)** *See attached*

   I led the design of the study in collaboration with all the co-authors and, in particular, Professor Richard Thomson, who was the Head of Research and Epidemiology at the NPSA at the time. I conducted the research independently under the supervision of Professor Thomson, apart from the RCRR where I worked with an expert clinical reviewer, Dr Sisse Olsen. In addition, Jane Chapman and Richard Sachs helped to recruit staff for interview and identify key hospital documents for me to review. I was responsible for the

analysis of data, production of the first draft of the paper. All co-authors made comments on successive drafts and approved the final version before journal submission. I acted as a guarantor of the final published version.

NAME IN FULL (Block Capitals) ......Dr Helen Hogan

STUDENT ID NO: 10586

CANDIDATE'S SIGNATURE .... Date ...1st May 2014

SUPERVISOR/SENIOR AUTHOR'S SIGNATURE (3 above) ....

What can we learn about patient safety from information sources within an acute hospital: a step on the ladder of integrated risk management?

H Hogan,[1] S Olsen,[2] S Scobie,[3] E Chapman,[4] R Sachs,[5] M McKee,[6] C Vincent,[2] R Thomson[3]

[1] London School of Hygiene and Tropical Medicine, London, UK; [2] Imperial College, London, UK; [3] National Patient Safety Agency, London, UK; [4] Whitecot Consultancy Ltd, London, UK; [5] North West London Hospitals Trust, London, UK; [6] European Observatory on Health Systems and Policies, London School of Hygiene and Tropical Medicine, London, UK

Correspondence to:
Dr H Hogan, Department of Public Health and Policy, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK; helen.hogan@lshtm.ac.uk

Accepted 19 August 2007

ABSTRACT

Objective: To assess the utility of data already existing within hospitals for monitoring patient safety.

Setting: An acute hospital in southern England.

Design: Mapping of data sources proposed by staff as potentially able to identify patient safety issues followed by an in-depth analysis of the content of seven key sources.

Data source analysis: For each data source: scope and depth of content in relation to patient safety, number and type of patient safety incidents identified, degree of overlap with incidents identified by different sources, levels of patient harm associated with incidents.

Results: A wide range of data sources existing within the hospital setting have the potential to provide information about patient safety incidents. Poor quality of coding, delays in reports reaching databases, the narrow focus of some data sources, limited data-collection periods and lack of central collation of findings were some of the barriers to making the best use of routine data sources for monitoring patient safety. An in-depth analysis of seven key data sources (Clinical Incident database, Health and Safety Incident database, Complaints database, Claims database and Inquest database, the Patient Administration System and case notes) indicated that case notes have the potential to identify the largest number of incidents and provide the richest source of information on such incidents. The seven data sources identified different types of incidents with differing levels of patient harm. There was little overlap between the incidents identified by different sources.

Conclusion: Despite issues related to the quality of coding, depth of information available and accessibility, triangulating information from more than one source can identify a broader range of incidents and provide additional information related to professional groups involved, types of patients affected and important contributory factors. Such an approach can provide a focus for further work and ultimately contributes to the identification of appropriate interventions that improve patient safety.

An integrated approach to risk management requires healthcare organisations to gather information on risk and safety from a range of information sources so that the scale and nature of key risk areas can be assessed. At a national level, the National Patient Safety Agency (NPSA) has established a Patient Safety Observatory to quantify, characterise and prioritise patient safety issues by bringing together information held by different organisations.[1] At a local level, despite the fact that hospitals in the UK collect a wealth of data on many aspects of patient care, these data have been seen as an underutilised source of information on patient safety.[2] The majority of hospitals in England and Wales rely on voluntary reports of patient safety incidents (defined by the NPSA as unintended or unexpected incidents that could have led or did lead to harm for one or more patients[3]) to Local Risk Management Systems (LRMS) to identify trends and areas for further investigation. The NPSA's national database (the National Reporting and Learning System) consists almost entirely of data derived from this system. However, this approach has been criticised as potentially misleading. A few studies have compared the number and types of incidents identified by LRMS with those identified by other sources including case notes,[4] internal departmental incident reporting systems[5] and computerised hospital administrative records.[6 7] These have shown that LRMS can fail to pick up serious incidents and are more likely to identify incidents not attributable to direct staff action such as falls than those related to clinical care. Evidence also suggests that nurses report more incidents than other staff groups, with significant under reporting by doctors.[8]

A variety of methods have been used to identify adverse events affecting hospitalised patients including retrospective case note review,[9–11] in-person collection of information from staff and case records on the wards,[12] direct observation,[13] screening of administrative data[14] and staff and patient surveys.[15] Comparisons of incidents detected by different methods have shown relatively little overlap between sources.[4 7 16] These findings suggest that there may be a value in bringing together information on patient safety from a wider range of sources. While most hospitals do not have the resources to institute some of the methodologies used in these studies, it is plausible that they are able to make better use of the data sources that they currently have. This study investigated the range of sources information relevant to patient safety found within a single acute hospital in England, the scope of information held by these sources and how it might be used to examine key areas of patient safety.

METHODS

Design

A mapping exercise, including semistructured interviews with 33 clinical and non-clinical staff, direct examination of data sources and attendance

Error management

at clinical governance meetings, was used to identify potentially useful hospital data sources in a large district general hospital in southern England with 850 beds and approximately 40 000 admissions per year.

By judging each source against the criteria of number and types of incidents that could be identified, mode of data collection, accessibility and content, seven sources were selected for more detailed analysis. Retrospective data collection was undertaken in respect of adult medical and surgical inpatients admitted between 1 April 2004 and 31 March 2005. Data sources were assessed to identify the completeness of information found in each source, the number and types of incident detected, patient harm resulting from each incident and the degree of overlap between incidents identified by different sources. An incident was registered if the coded event suggested the potential to cause patient harm, even if harm was not explicitly recorded, as information on harm was not always available from some sources. All incidents detected were coded by category and by level of harm using the standard coding system employed by the National Reporting and Learning System (NRLS) (table 1).[17] Completeness of information was assessed by identifying how many of the data items required by the NRLS incident report form were present.[18] Using patient full name as an identifier, the degree of overlap between records held on each database was examined.

Data sources

The sources interrogated included:

Box 1 Examples of different categories of incident detected by different data sources

Clinical incidents: medication error
Patient given five doses of Co-dydramol in 1 day, although prescribed four times daily—no patient complaints. Drug chart needed rewriting; previous entries illegible.

Complaints: consent, communication and confidentiality
Patient's daughter raises concerns regarding doctor's attitude. When she spoke to doctor, she was told mother was doing very well, when in fact the doctor was describing a different patient.

Health and Safety incidents: patient accident
Patient being weighed on sitting scales, mobility poor and brakes loose, scales moved and patient fell on her side to the floor. Slight graze to existing haematoma on right elbow and possible injury to right hip.

Claims: treatment and procedure
Patient had right ankle injected instead of left.

Inquest: clinical assessment
Patient attended A&E three times and sent home with diagnosis of tonsillitis. Prescribed antibiotics. Finally presented with severe shock requiring resuscitation and surgery. Delay in diagnosing haemorrhage due to ruptured spleen. Patient died.

Adverse event from case note review: treatment and procedure
Cystic artery inadvertently cut during laparotomy cholecystectomy. Operation converted to a laparoscopy in order to control bleeding. Blood loss estimated as 2 litres. Postoperative blood transfusion given.

Five databases which made up the Trust's Local Risk Management System

The Clinical Incident database, Health and Safety Incident database, Complaints database, Claims database and Inquest database were searched for patient-identifiable entries linked to adult medical and surgical inpatients within the index year.

Patient Administration System

The Patient Administration System (PAS) was searched using the 41 three-digit ICD 10 diagnosis codes for complications and misadventures.[19] Within each record, complication or misadventure codes appearing in the first of the six diagnostic code boxes, which is normally used to designate the reason for admission, were considered to indicate a preadmission event and excluded.

Case notes

Two hundred and twenty randomly drawn records with an adult surgical or medical admission within the index year were reviewed. Reviews were undertaken using a method adapted from that described by Woloshynowych et al.[20] A 10% sample of case notes with no incident identified and a 25% sample of positive case notes were further reviewed by a second expert reviewer (SO). All positive cases were also discussed with two expert reviewers (SO and GN). The presence of a patient safety incident was ascribed only if all three reviewers were in agreement (75% of cases).

RESULTS

Data-quality issues

Table 2 describes the range of data sources that have the potential to provide information about patient safety incidents occurring in medical and surgical inpatients that exist within the hospital. These sources could be divided into four main types: Incident Reporting Systems, Surveillance Systems, Audits and others (including case notes, the Patient Administration System, minutes from Morbidity and Mortality meetings, written claims and inquest records). Poor quality of coding, delays in reports reaching databases, the narrow focus of some data sources, time-limited data-collection periods and lack of central collation of findings were some of the barriers that limited the scope for routine data sources to be used in monitoring patient safety.

Of the seven data sources selected for more detailed analysis, case notes contained the most detailed information on individual incidents including risk factors such as age, ethnicity or comorbidities, time, place and location, description of the incident, levels of harm and contributory factors. The Clinical Incident database also contained many of these data items, but information on risk factors and contributory factors was usually not entered. In addition, the category of healthcare professional who was involved in the incident or who made the report was often missing. It was difficult to make objective assessments of patient harm for incidents detected in the Complaints database or on the PAS system. For the former, information on harm was principally from the patient or carer's perspective only, and for the latter there was inadequate detail. Elucidating the temporal sequence of events was sometimes challenging with individual PAS records, occasionally leading to difficulty distinguishing preadmission diagnoses or comorbidities from in-hospital incidents.

Box 2 Examples of incidents graded for differing levels of patient harm

Clinical incident: no harm impact prevented
Patient's "to take home" medication mislabelled as 5 mg Amlodipine when contents in the bag were in fact 10 mg. Nurse informed pharmacy of error and returned the drugs.

Health and Safety: no harm impact not prevented
Patient found sitting on floor by staff. No obvious injury noted.

Complaints: low
Patient's son has written regarding her mother was who given a wrong wrist band in A&E on 2/4/04. He also has other concerns regarding a cannula which was left in his mother when discharged.

Claims: moderate
Patient is diabetic with circulatory problems. Allegation made that insufficient care was taken during his admission to prevent development of pressure sores, which became infected with MRSA.

Case note review, adverse event: severe
Patient who was post coronary artery bypass graft gradually deteriorated over 1 week with symptoms of shortness of breath and a discharging chest wound. Clinical team failed to investigate reason for deterioration. After 5 days became acutely unwell and found to have chest wound breakdown and a passageway between wound and chest cavity. Transferred to ITU and put on a ventilator. Recovery took several weeks.

Inquest: death
Patient died following a right hemicolectomy. Cause of death on post-mortem was (1) haemorrhage due to right hemicolectomy, (2) Crohn's disease.

Numbers of incidents and degree of overlap between sources

Table 3 shows the number patient safety incidents identified by each of the seven data sources in total and the number identified exclusively by each source. Case notes potentially identify the largest number incidents. Seventy-one patient safety incidents were found across 220 inpatient admissions (32.3%). Of these, 40 (18.1%) fulfilled the definition of an adverse event (an unintended injury or complication of care leading to prolonged admission, disability at discharge or death and caused by healthcare management rather than the disease process), the term most often used to describe incidents in previous studies. Sixty-five patient safety incidents were single events. Three patients experienced two incidents during the index admission. Based on these findings, it can be estimated that 8781(95% CI 6495 to 12 144) incidents could potentially be identified across the 27 270 adult medical and surgical admissions between 1.4.04 and 31.3.03. The second largest source of patient safety incidents was the Clinical Incident database with 484 incidents identified within the index year. This was followed by PAS (462

incidents), Complaints (221 incidents) and Health and Safety (176 incidents). Inquest and Claims records identified small numbers of events, 21 and 10, respectively.

Table 4 demonstrates that the degree of overlap between incidents picked up by different data sources was small.

Types of incidents identified by different data sources

Different data sources tended to identify different proportions of incidents in each category (table 5), and 37.5% of incidents identified via the Clinical Incident database were medication errors and equipment failures. Complaints provide an insight into incidents related to communication failures (22% of total). The PAS system, Inquests database and case notes identified many incidents linked to surgical interventions and to investigative procedures. The PAS system was also useful in identifying incidents related to infection control (see Box 1 and table 6, for examples of incidents detected by different data sources).

Levels of patient harm

Incidents were graded by severity; the proportion of incidents graded as causing death, severe, moderate, low or no harm varied among the different data sources (table 7). Incidents found in inquest and claims records tended to be associated with death or serious harm; case records identified incidents mainly causing moderate or low levels of harm and LRMS databases captured a higher proportion of incidents causing low levels of harm or no harm. Although there was not enough information available to code the majority of incidents detected by the PAS, 8.4% of patients with a coded complication or misadventure in their record died (see Box 2 for examples of incidents graded for differing levels of patient harm).

DISCUSSION

The aim of this study was to explore the feasibility of using a range of hospital data sources to identify patient safety incidents in order to provide a better picture of the scale and scope of incidents related to key safety issues in an English hospital. The study focuses on routine data sources that are available within the hospital setting.

A number of limitations should be considered. The study was carried out in a single acute hospital, and although data sources identified by the mapping exercise as containing potentially useful patient safety information are likely to be present in other sites, accessibility, quality of coding and completeness may vary from site to site. The study also focused on data sources linked to medical and surgical inpatients only. While these sources also provide information on incidents related to other specialties and outpatients, there will be additional data sources that can also be utilised for this purpose, and some tailoring in relation to the issue being explored is necessary.

Many routine hospital data sources collect data for purposes other than identifying patient safety incidents. These data are observational rather than experimental and are prone to biases

Table 1 National Reporting and Learning System's harm grading for patient safety incidents

No harm	Impact prevented: any patient safety incident that had the potential to cause harm but was prevented, resulting in no harm to people receiving NHS-funded care
	Impact not prevented: any patient safety incident that ran to completion but no harm occurred to people receiving NHS-funded care
Low	Any patient safety incident that required extra observation or minor treatment
Moderate	Any patient safety incident that resulted in a moderate increase in treatment
Severe	Any patient safety incident that appears to have resulted in permanent harm
Death	Any patient safety incident that directly resulted in death

Error management

Table 2 Description of data sources collecting information on patient safety in one hospital and issues related to using such sources to monitor patient safety

Data source	Incident-reporting systems	Surveillance systems	Audit data	Patient administration system	Case notes and other written records
Description	Includes the Local Risk Management System, containing coded information on clinical and Health and Safety incidents, complaints, claims and inquests. Pathology, diagnostic services (radiology, endoscopy, etc) and electrobiomedical departments had internal incident reporting systems	A number of mandatory and voluntary surveillance systems existed; these tended to be focused on key public health concerns such as infection control, for example MRSA bacteraemia surveillance, or specific high-risk areas, for example transfusion of blood products or the introduction of new drugs	A number of local hospital audits, for example Resuscitation Audit and national audits, for example National Confidential Enquiry into Perioperative Outcomes and Deaths collected information on outcomes of clinical care such as complications or errors in the process of care, residual disability or death	Clinical information was available as a by product of the hospital's administration and reimbursement system; 41 ICD 10 codes were being used to code for complications and misadventures	Included patient case notes, minutes of Morbidity and Mortality meetings, written records related to claims and inquests
Categories of incident commonly found in these sources	Patient accidents, medication errors, equipment failures, access to care, admissions and discharges, communication failures, complications or delays related to diagnostic testing	Infection control, blood transfusion and medication incident	Clinical assessment, implementation of care and monitoring, infection control, medication errors and treatments and procedures	Infection control, treatment and procedures	Access to care, admissions and discharges, clinical assessment, implementation of care and monitoring, infection control, communication, infrastructure, medication errors and treatments and procedures
Mode of data collection	Continuous	Majority continuous	Majority have limited period of data collection	Continuous	Continuous or recurring
Accessibility	Majority electronic	Majority electronic	Mixture of electronic and paper	Electronic	Paper
How complete are individual reports Graded 1–4*	Mainly 3	2–4	2–3	2	3–4
	Information about incidents in complaints and claims records is usually from the patient/carer's perspective only as is any reference to harm	Varies between sources from limited information required by MRSA bacteraemia surveillance to detailed reports required by Serious Hazards of Transfusion Reporting System	Sometimes clinical judgment is required to identify both incident and degree of harm which may not be referred to explicitly in the record; complication of care may be recorded but information on patient harm missing	Very little information on the incident; information related to patient harm limited to whether the patient was discharged alive or died	Case notes provide the most comprehensive information on incidents; inquest and Claims records can contain post-mortem reports and detailed hospital reports
Issues related to using data source	Under reporting especially by health professionals other than nurses	Narrow focus	Lack of central collation of findings	Misses complications of treatment not specifically coded using one of the 41 ICD 10 codes	Time taken to gather information from written records
	Poor quality of coding of incidents by type/subtype	Information often not collated and analysed locally	Poor quality of some local audits	Completeness and accuracy of use of these codes by medical coders	Missing records or parts of records
	Lack of risk factor information, for example age, gender, ethnicity		Short-term data collection and limited sample sizes	Judgment may be required to separate complications that were the main reason for an admission from those that occurred during an admission	Poor legibility
	Profession of reporter or person involved in the incident not always collected		National audits such as the confidential enquiries may only sample a small number of patients from each hospital	Bias introduced by the limited number of codes which will tend to identify more surgical than medical patients	Sensitivity of material and willingness of staff to share information with a wider audience
	Bias towards low harm no harm events				
	Delays in complaints and claims reaching hospital				
	Subjectivity of content of complaints and claims records				
	Information from departmental incident reporting systems not shared outside those departments				

*Completeness rating: 1 = no patient identifiable information; 2 = patient identifiable information, subjective/implicit information or scanty details on incident, link to harm not clear; 3 = patient identifiable information, more detailed information about incident (place, time, person), some links with harm, no prevention or further action; 4 = patient identifiable information, more detailed information about incident (place, time, person), harm described/graded for harm, action or prevention described.

Table 3 Number of incidents identified by each of the seven data sources for adult medical and surgical patient admissions between 1.4.04 and 31.3.05

Data source	Total no. of incidents identified	No. of incidents exclusively identified by source
Case notes*	8781	NA†
Clinical incidents	484	428
PAS	462	399
Health & Safety	221	197
Complaints	176	148
Inquest	21	10
Claims	10	7
Total	10 190	

*Estimate based on a sample of 220 inpatient admissions.
†Of the 71 patient safety incidents identified from case notes, 10 were also found by other sources.

introduced by differential reporting levels, the variable quality of coding and levels of completeness. The most useful supplementary information sources would be those that collect timely data on a continuous basis, which are accessible and have appropriate content. Timely and continuous data collection is important for analysis of trends over time. These conditions limit the use of many hospital sources where data collection is short-lived and non-recurring. Accessibility is also limited by the long delays in information reaching some sources, the lack of central collation, poorly coded electronic information and the amount of time required to review written records. Staff may also feel uncomfortable sharing some types of sensitive information, an attitude difficult to change if the hospital's culture is not perceived as open and fair. Important content, such as information related to risk factors including age, gender, ethnicity or profession of healthcare worker involved, is often not available. Furthermore, the patchy nature of data collection across any healthcare organisation and the narrow focus of

Table 4 Degree of overlap between incidents identified by each data source, April 2004 to March 2005

Clinical incidents (484)	Clinical incidents					
Claims (10)	3	Claims				
Inquests (21)	9	0	Inquests			
Complaints (176)	4	2	2	Complaints		
Health & Safety (221)	7	0	1	10	Heath & Safety	
PAS (462)	35	1	3	12	14	PAS
Case notes (71)	3	0	0	2	1	6

much of this data collection will inevitably mean that there will be gaps in information in certain areas.

Limited information makes it difficult to identify incidents. In the Complaints and Claims databases, the descriptions of events are mainly from the patient or carer's perspective with limited information from the hospital, diagnoses in PAS records do not indicate when they occurred, and information on patient harm, apart from death, is not available. To get the most out of the data available, a pragmatic approach was taken to identifying incidents, judging whether the event described had the potential to cause patient harm, even if that harm was not explicitly recorded. This may have led to some overestimation of the numbers of incidents. The relatively small sample of case notes reviewed, representing 0.8% of all adult medical and surgical admissions in the index year, also increases the degree of uncertainty around estimates of the total number of patient safety incidents that can be identified by this method. These problems of validity and reliability highlight why such sources should not be used for comparisons of facilities. However, they are less important when hospitals are using the data internally to pinpoint areas of concern as part of internal quality-improvement processes.

Table 5 Numbers and proportions of incidents in each category detected by different data sources, April 2004 to March 2005

NPSA incident categories	Clinical incidents	Complaints	Health & Safety	Claims	Inquests	PAS	Case notes
Access, admission, transfer	67 (13.8%)	15 (8.5%)					8 (11.3%)
Clinical assessment	39 (8.1%)	12 (6.8%)		2 (20%)	5 (23.8%)		13 (18.3%)
Consent, communication & confidentiality	33 (6.8%)	39 (22.2%)					
Disruptive, aggressive behaviour	2 (0.4%)	1 (0.6%)	2 (0.9%)				
Documentation	43 (8.9%)	3 (1.7%)					
Care & ongoing monitoring and review	31 (6.4%)	29 (16.5%)		1 (10%)	1 (4.8%)		7 (9.9%)
Infection control	2 (0.4%)	16 (9.1%)		2 (20%)	1 (4.8%)	181 (39.2%)	9 (12.7%)
Infrastructure	34 (7.0%)	29 (16.5%)			1 (4.8%)		2 (2.8%)
Medical equipment	73 (15.1%)		7 (3.2%)	1 (10%)		4 (0.9%)	1 (1.4%)
Medication error	100 (20.7%)	4 (2.3%)			1 (4.8%)	13 (2.8%)	15 (21.1%)
Patient abuse		3 (1.7%)					
Patient accident	15 (3.1%)	7 (4.0%)	212 (95.9%)		1 (4.8%)		
Self-harming behaviour	1 (0.2%)						
Treatment, procedure	44 (9.1%)	18 (10.2%)		4 (40%)	11 (52.4%)	264 (57.1%)	16 (22.5%)
Total	484	176	221	10	21	462	40

Table 6 Patient administration system: infection control incident

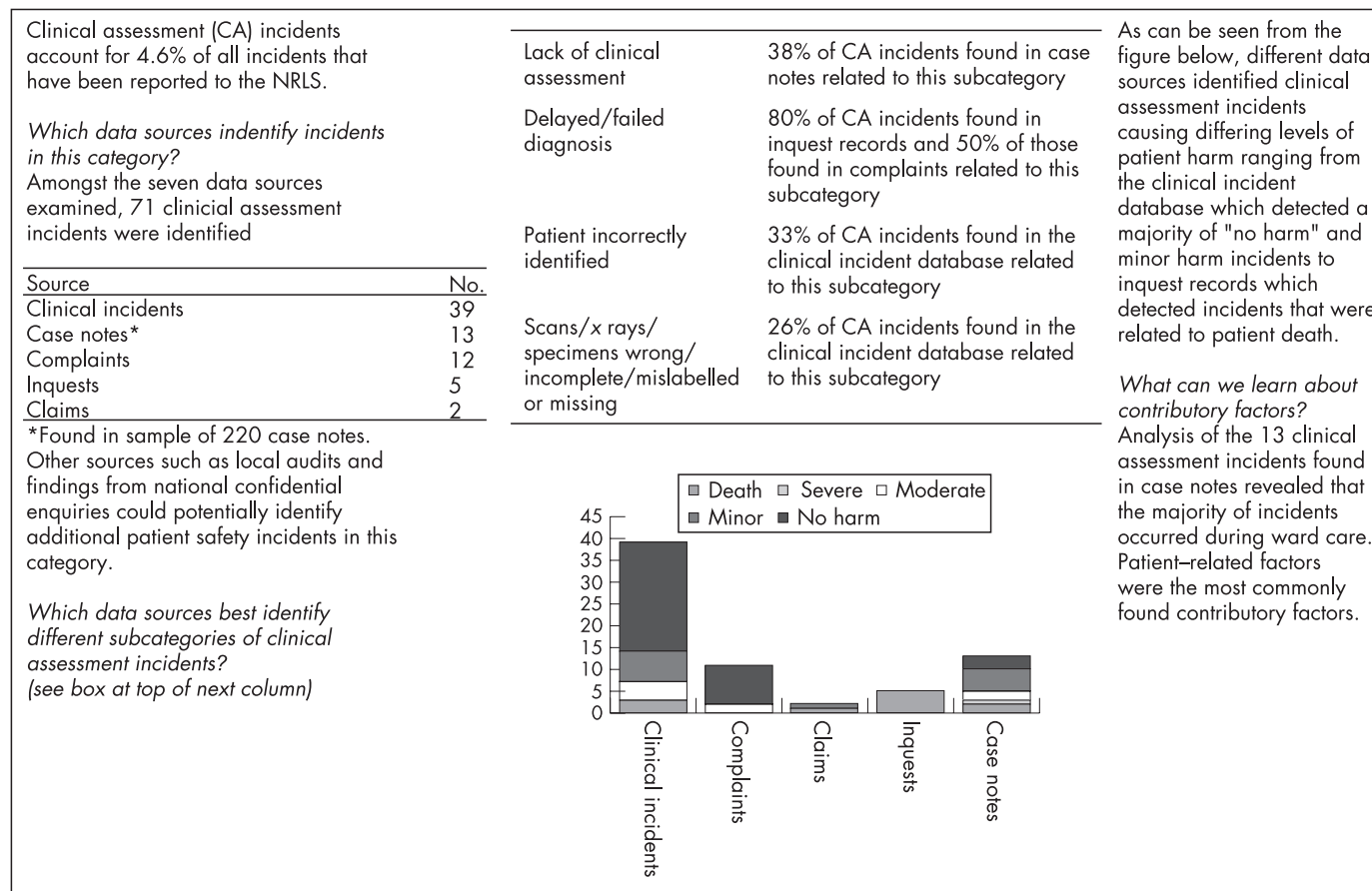| Diagnosis 1	Diagnosis 2	Diagnosis 3	Diagnosis 4	Diagnosis 5	Diagnosis 6	Procedure 1
I803 ~ Phlebitis and thrombophlebitis of lower extremities unspecified	I739 ~ Peripheral vascular disease unspecified	I10X ~ Essential (primary) hypertension	T814 ~ Infection following a procedure not classified elsewhere	B956 ~ Staph aureus as cause of dis classified to other chapters	N390 ~ Urinary-tract infection site not specified	L592 ~ Bypass of femoral artery by anastomosis of femoral artery to popliteal artery using prosthesis

Table 7   Proportion of incidents in different harm grades for each data source, April 2004 to March 2005

	Clinical incidents	Complaints	Health & Safety	Claims	Inquest	Case notes
Death	8 (1.7%)	0	0	2 (20%)	21 (100%)	2 (2.8)
Severe	9 (1.9%)	1 (0.6%)	2 (0.9%)	0	0	2 (2.8)
Moderate	28 (5.8%)	18 (10.2%)	3 (1.4%)	2 (20%)	0	18 (25.4%)
Low	114 (23.6%)	38 (21.6%)	86 (38.9%)	6 (60%)	0	35 (49.2%)
No harm impact not prevented	259 (53.5%)	107 (60.8)	127 (57.4%)	0	0	0
No harm impact prevented	66 (13.6%)	0	0	0	0	0
Not possible to code	0	12 (6.8%)	3 (1.4%)	0	0	0
Total	484	176	221	10	21	71

Our analyses, which focused on seven data sources, indicated that the Clinical Incident database, the main database used by many trusts to monitor patient safety incidents, identifies relatively few incidents overall. This finding is in line with previous studies.[4 7 21] In addition, each source picked up its own unique collection of incidents in terms of type and levels of harm, with minimal overlap between sources. From the mapping exercise, it is clear that there are many other data sources found within the hospital setting which have the potential to provide useful information on patient safety, particularly if use is tailored to the investigation of specific problems. Triangulating information from a wider range of data sources presents an opportunity to gain a greater understanding of key patient safety issues, including a better understanding of the common types of incidents, the healthcare professional groups and types of patients involved, and important contributory factors. It offers the opportunity to learn from events that cover the spectrum of patient harm. Using information from a range of sources can enhance investigations of key risk areas such as medication errors, diagnostic testing, infection control or treatments and procedures (see fig 1 for an example). It offers both a mechanism for ongoing monitoring and an opportunity to better focus clinical governance activities such as audit or targeted case note review. The future development of validated patient safety indicators, similar to those employed by the Agency for Healthcare Research and Quality (AHRQ) in the US, will increase the utility of information derived from administrative data such as PAS.[22] Collaborative working between clinical staff, clinical risk teams and information technology is essential to make the most of all the data sources available. Staff cooperation, in turn, depends on the presence of an open and fair culture, with an emphasis on learning from incidents rather than apportioning blame.

Clinical assessment (CA) incidents account for 4.6% of all incidents that have been reported to the NRLS.

Which data sources indentify incidents in this category?
Amongst the seven data sources examined, 71 clinical assessment incidents were identified

Source	No.
Clinical incidents	39
Case notes*	13
Complaints	12
Inquests	5
Claims	2

*Found in sample of 220 case notes. Other sources such as local audits and findings from national confidential enquiries could potentially identify additional patient safety incidents in this category.

Which data sources best identify different subcategories of clinical assessment incidents?
(see box at top of next column)

Lack of clinical assessment	38% of CA incidents found in case notes related to this subcategory
Delayed/failed diagnosis	80% of CA incidents found in inquest records and 50% of those found in complaints related to this subcategory
Patient incorrectly identified	33% of CA incidents found in the clinical incident database related to this subcategory
Scans/x rays/ specimens wrong/ incomplete/mislabelled or missing	26% of CA incidents found in the clinical incident database related to this subcategory

As can be seen from the figure below, different data sources identified clinical assessment incidents causing differing levels of patient harm ranging from the clinical incident database which detected a majority of "no harm" and minor harm incidents to inquest records which detected incidents that were related to patient death.

What can we learn about contributory factors?
Analysis of the 13 clinical assessment incidents found in case notes revealed that the majority of incidents occurred during ward care. Patient–related factors were the most commonly found contributory factors.

Figure 1   An approach to exploring incidents related to clinical assessment.

Information plays a vital role in identifying, monitoring and investigating levels of risk, promoting safer healthcare within organisations and enabling delivery of the continuous quality improvement that underpins the clinical governance agenda in the UK. Chief executives and directors of Trusts are now accountable for organisation-wide assessment of patient safety risks. This study highlights the advantages of triangulating information from a range of sources when making such assessments. Leadership from senior managers is vital to promote a culture that promotes the sharing of information derived from these sources among different departments. The directors are also in a position, based on the findings from local investigations or intelligence received from external sources, to identify key risk areas requiring further investigation anywhere in the hospital and to provide the resources needed to ensure that the methodology and findings from such investigations are disseminated and further in depth work such as audit is commissioned if necessary. Such a targeted approach to improving patient safety would allow for the most efficient use of scarce resources. Clinical governance teams could provide technical support to departmental staff, helping them to identify and exploit information sources that are relevant for a particular investigation, along with advice on search strategies, collation and analysis of these data. Work to overcome the limitations of some data sources, such as improvement of the quality and consistency of electronically coded information or institution of new data collection systems to address key knowledge gaps could be part of this support. Such an approach need not be restricted to acute hospitals although the challenges relating to sharing information across the primary-secondary care interface are likely to be more significant.

The NPSA's National Reporting and Learning System draws the majority of the incident reports it uses for monitoring national trends from LRMS. The agency acknowledges that incident reporting systems alone can never be relied upon to provide a comprehensive picture of patient safety.[1] At a national level, the NPSA draws more widely on intelligence from a range of sources both within and external to the NHS via its Patient Safety Observatory to better characterise patient safety issues. At the local level, data from a broader range of local sources would also seem to offer valuable supplementary information to the NPSA. Such findings could reach the agency via its network of Patient Safety Managers who currently work closely with individual Trusts.

In conclusion, gaining intelligence on patient safety incidents from a broader range of information sources has the potential to provide healthcare organisations with a better picture of key patient safety risks thus facilitating targeting of scare resources on appropriate interventions with the potential to improve patient safety.

REFERENCES

1. **National Patient Safety Agency**. *Building a memory: preventing harm, reducing risks and improving patient safety*. London: National Patient Safety Agency, 2005.
2. **National Audit Office**. *A safer place for patients: learning to improve patient safety*. London: National Audit Office, 2005.
3. **National Patient Safety Agency**. *Seven steps to patient to patient safety*. London: National Patient Safety Agency, 2004.
4. **Stanhope C,** Crowley-Murphey M, Vincent C, *et al*. An evaluation of adverse incident reporting. *J Eval Clin Pract* 1995;**5**:1–4.
5. **Ricci M,** Goldman A, de Leval M, *et al*. Pitfalls of adverse event reporting in paediatric intensive care. *Arch Dis Child* 2004;**89**:856–9.
6. **Classen D,** Pestotnik S, Evans R, Burke J. Computerised surveillance of adverse drug events in hospitalised patients. *Qual Saf Health Care* 2005;**14**:221–6.
7. **Jha A,** Kuperman G, Teich J, *et al*. Identifying adverse drug events: the development of computer-based monitor and comparison with chart review and simultaneous voluntary report. *J Am Med Inform Assoc* 1998;**5**:305–14.
8. **Lawton R,** Parker D. Barriers to incident reporting in a health care system. *Qual Saf Health Care* 2002;**11**:15–18.
9. **Brennan T,** Leape L, Laird N, *et al*. Incidence of adverse events and negligence in hospitalised patients. Results of the Harvard Medical Practice Study 1. *N Eng J Med* 1991;**324**:370–6.
10. **Wilson R,** Runciman W, Gibberd R, *et al*. Quality in Australian Health Care Study. *Med J Aust* 1995;**163**:472–5.
11. **Vincent C,** Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001;**322**:517–19.
12. **Bates D,** Cullen D, Laird N. Incidence of adverse drug events and potential adverse drug events. Implications for prevention. *JAMA* 1995;**274**:29–34.
13. **Taxis K,** Barber N. Ethnographic study of the incidence and severity of intravenous drug errors. *Qual Saf Health Care* 2003;**12**:342–7.
14. **Iezzoni L,** Daley J, Heeren T, *et al*. Identifying complications of care using administrative data. *Med Care* 1994;**32**:700–15.
15. **Ghandi T,** Weingert S, Borus J, *et al*. Adverse drug events in ambulatory care. *N Engl J Med* 2003;**348**:1556–64.
16. **Michel P,** Quenon J, de Saraqueta A, Scemama O. Comparison of three methods for estimating rates of preventable adverse events in acute care hospitals. *BMJ* 2004;**328**:199.
17. **National Patient Safety Agency**. Dataset download. http://www.npsa.nhs.uk/health/reporting/datasets (accessed 1 Jun 2005).
18. **National Patient Safety Agency**. Report an incident. http://www.npsa.nhs.uk/health/reporting/reportanincident (accessed 1 Jun 2005).
19. **Aylin P,** Tanna S, Bottle A, *et al*. How often are adverse events reported in English hospital statistics? *BMJ* 2004;**329**:369.
20. **Woloshynowych M,** Neale G, Vincent C. Case record review of adverse events: a new approach. *Qual Saf Health Care* 2003;**12**:411–15.
21. **Rozich J,** Haraden C, Resar R. Adverse drug event trigger tool: a practical methodology for measuring medication related harm. *Qual Saf Health Care* 2003;**12**:194–200.
22. **Zhan C,** Miller M. Administrative data based patient safety research: a clinical review. *Qual Saf Health Care* 2003;**12**(Suppl II):ii158–63S.

Chapter 4 Method

4.1 Introduction

This Chapter describes the development of a methodology for the determination of an estimate of the scale of preventable deaths in English acute hospitals, and exploration of the nature of that harm. Based on the findings from Chapter 2 and Chapter 3, RCRR was selected as the most appropriate method for the study. Given that there had been relatively few published RCRRs with a focus on deaths in the past,[17 174] I undertook a pilot study to test the feasibility of the approach and identify if there were adaptations needed. In this chapter I describe the findings of the pilot and its contribution to improving the reliability and validity of measurement, before describing the methods for the main study.

4.2 The Pilot Study

The aim of the pilot study was to establish what modifications might need to be made to the traditional RCRR method in order to ascertain a valid and reliable estimate of hospital preventable deaths. Forty-seven randomly selected deaths from two large London acute NHS hospital Trusts (20 from a teaching hospital and 27 from a large acute general hospital) were reviewed by myself (27 cases) and another expert reviewer (20 cases) in March and April 2007. A 10% sample of records from each reviewer was double reviewed by a third reviewer. Information was collected using a structured medical review form developed for the first English RCRR study that had been conducted in London in 1999.[235] Most of the patients in the sample were over 80 years old and had multiple co-morbidities. The study found 4 preventable deaths (8.5%), the majority of which occurred during general ward care

and were related to either a failure in clinical monitoring, hospital acquired infection, or a problem related to drugs or fluids.

The pilot study found that the case records were straightforward to access and were in hard copy (rather than microfiche), when deaths occurred within two years of the review. In addition, tracers in medical record databases enabled records missing from the Medical Records department to be located and retrieved for reviewing. However, specific instruction needed to be given to ensure that medical records staff made every attempt to retrieve these missing records. The average review time was one hour, with a review rate of six to seven case records per day. Most records (96%) were found to contain enough content to allow an assessment of the presence of harm and its preventability.

As a result of the pilot study, the following decisions were made about the design of the proposed main study:

1. New definition for harm events

    The pilot study confirmed that the term 'adverse event' tended to draw attention to acts of commission rather than omission, and to harms related to actions of individuals rather than those that were system-based. In consultation with Professor Charles Vincent and Dr Graham Neale, who had conducted the first English RCRR in London, and Professor Richard Thomson, the Head of Research for NPSA, a new definition of harm, termed 'problems in care', was developed and defined as:

    'Patient harm resulting from acts of omission (inactions), such as failure to diagnose and treat, or from acts of commission (affirmative actions), such as incorrect treatment or management, or harm as a result of unintended complications of healthcare'.

This definition was designed to be inclusive of the widest range of harms, including those that resulted from multiple, small sub-optimal failures in care, especially if these occurred over the period of the admission.

2. Questions to determine preventable death

In traditional RCRR studies, reviewers make a series of judgements related to whether the patient experiences harm, whether that harm was caused by healthcare and, if so, if it was preventable. At each point a decision is usually made using a six point Likert scale, reflecting the probabilistic nature of the decision. Again, based on the pilot study and in consultation with experts this series of questions was simplified. In the newly designed form the reviewer was asked to answer 'yes' or 'no' to a question enquiring whether problem(s) in care had contributed to the patient's death. An affirmative answer led to a further set of questions on the nature of the problem in care that culminated in a question on the preventability of the death, which reviewers were asked to rate on the six point Likert scale. The study design envisaged using a value of four or more on the Likert scale - deaths were judged preventable if reviewers felt that there was more than a 50% chance that the death was preventable.

3. Life expectancy

Descriptions of the burden of harm in healthcare have evoked images of a plane crash, or busloads of people dying every day, giving an impression that preventable deaths have as much impact on the young as on the old.[8] This assumption is unlikely, given that approximately 60% of deaths in England occur in hospital, and the majority of those are in elderly patients. To address this issue, an estimation of life expectancy was added to the review. A range of options for capturing this information was explored. The Charlson index is commonly used in elderly

populations, and captures life expectancy based on the number and severity of co-morbidities, but is not designed for use in gathering data on critically ill patients.[236] In contrast, the Acute Physiology and Chronic Health Evaluation (APACHE), and other associated tools, were designed to estimate life expectancy in critically ill patients, but completion requires access to a range of physiology results that might not be available in the patient record. Furthermore, both the Charlson Index and the APACHE score were created for case mix adjustment in population studies, and not for making estimates for individuals.[237] Disease specific tools also generally perform poorly if used for prognostication in elderly people with six months or less to live.[238] The same applies to performance status measures.[239]

Given the complexities of using a tool for life expectancy estimation in a cohort of elderly patients with multiple co-morbidities and towards the end of their lives, I adopted the approach taken by Hayward and Hofer in their small study of preventable deaths,[17] and the Dutch Adverse Event study,[16] both of which used physicians' judgements of life expectancy. Reviewers were asked to consider three factors when making this decision: the degree of urgency on admission (critical, urgent, semi-urgent, routine), the patient's functional state on admission using a seven point scale from 'normal' to 'requires special care' adapted from the Karnofsky Index,[240] and the patient's acute and chronic condition.

4. Measuring Quality of Care

To explore the relationship between overall quality of care, and problems in care and preventable death, a 'quality of care' rating scale designed by the RAND Corporation was included in the Review form.[215] The scale asked reviewers to rate the quality of each phase of care (from admission through to ward care), before giving an overall rating.

4.3 The Preventable Incidents Survival and Mortality Study: Main Methods

The final study design drew heavily on the approach used by Professor Charles Vincent and Dr Graham Neale for their RCRR of 1014 admissions, which took place in two London hospitals during the late 1990s.[15] Their design had, in turn, been based on the HMPS. Professor Vincent and Dr Neale were advisors to this study, and provided invaluable guidance.

4.3.1 Determining Sample Size

I based sample size estimation on the proportion of preventable deaths found in the US study of 111 deaths conducted by Hayward and Hofer (6%), as this was a more conservative estimate than that found in the pilot study.[17] For a simple random sample, a sample size of 347 deceased patients was required (alpha=0.05) to estimate a confidence interval of 2.5% to both sides. In an effort to make the study's results more generalisable, I decided that sub-samples of deceased patients should be selected from ten English acute hospital Trusts of contrasting size and location. The number of hospitals was a pragmatic choice based on study resources. With a sample size of ten hospitals, taking inter-hospital clustering into account, and using the intra-cluster correlation coefficient derived from the well powered Dutch Adverse Event Study[16] (0.037), the sample size needed to be increased at least three fold to 1000, in order to account for the two stage sampling strategy, and to maintain the same 95% confidence interval. The ten acute hospitals were randomly selected from a list of all English acute hospitals, which was then stratified by size, geographical location and HSMR quartile. An index year of 2009 was chosen for the sample, being the last complete calendar year before reviewing was scheduled to start, and recent enough to ensure that hardcopy records were likely to be available for review.

4.3.2 Sample Exclusions and Inclusions

The majority of previous RCRRs excluded psychiatric patients,[6 169] with some also excluding obstetric patients,[174] and others children under 18.[15] Given that obstetric, psychiatric and paediatric patients accounted for less than 5% of all hospital deaths in England and Wales,[64] and would require recruitment of specialist reviewers in those areas to undertake the case record review, I decided to exclude these patients from the sample. Additionally, patients that were admitted explicitly for palliative care were excluded. Deaths were included if they occurred within the first 24 hours after admission, as long as the notes indicated that the patient had been formally admitted to hospital and was not just an A&E attendance.

4.3.3 Hospital Recruitment

I sent a letter to each hospital Chief Executive outlining the aims of the study, the approach and how findings would be fed back. All hospitals were asked to nominate a lead (a senior member of the clinical governance team in most cases) who would be the main point of contact during the study. A protocol was developed covering patient sampling, and the location, tracing and retrieval of medical records at each site. Information Departments were asked to draw up a list of adult medical and surgical patients from 2009 that had died during their admission, using the Patient Administration System. The Microsoft Office Excel 2010 random numbers generator was used to select a random sample from the list. Oversampling was undertaken to ensure replacement records were available if those from the original list could not be traced, or had inadequate content. I gave instructions to medical records staff about the importance of tracing any missing records to ensure the sample was as complete as possible. It was stressed that this was particularly important if these records were located in the legal department. Each hospital maintained a list of all records that could not be traced, including details of age, sex, specialty, whether subject to a coroner's inquiry or claim, and last location, as well as the reason for their loss.

In previous RCRRs, reviews have usually been completed on the hospital site. This has the advantage of allowing access to the complete record, enabling the reviewer to track a problem backwards from the index admission if its origin falls outside this hospitalisation. In addition, missing elements such as drug charts or laboratory results can be searched for within the record itself, or accessed via the hospital's computer system. One RCRR study in Birmingham did obtain photocopies of the index admission, previously anonymised by hospital staff, and this allowed reviewing to take place remotely.[241] However, the resource implications of this approach were beyond the means of this study.

4.3.4 Reviewer Recruitment and Training

I considered whether each hospital's own doctors could undertake the reviews. Previous studies indicated that this was feasible and would have the advantage of leaving a legacy of trained staff.[200 242] However, given the nature of the study's aims to produce the first rigorous estimate of preventable death in England, it seemed more appropriate to use external reviewers to ensure the independence of the findings. I felt that the workload for NHS consultants could restrict their ability to take part in this study, and chose instead to recruit reviewers who were recently retired NHS hospital consultants. A limit of five years post retirement was set, to ensure that these doctors remained relatively up to date with current NHS practice.

Recruitment was undertaken through the Royal College of Physicians and Royal College of Surgeons, via adverts in College newsletters. I vetted CVs for relevant experience, and conducted informal interviews over the phone. The pilot study had indicated that most problems in care would be of a generalist nature and would not need a reviewer to have particular specialist knowledge. Only two surgeons were eventually recruited, the rest being physicians with a generalist background. I felt that the surgical resource could be most usefully employed by providing advice to the physicians on problems in care directly related

to surgery. Across the reviewer base, many different specialties were represented, which opened up the possibility of dealing with specialist-type queries generated by the reviewers within the group.

I designed a one day training programme in collaboration with Dr Graham Neale. After an initial introduction to the study, the protocol, definitions and review form, there was opportunity for the participants to practise making judgements on short and long cases. Clinical discussions based on the cases led by Dr Neale covered areas including avoidance of hindsight bias, and clarified the procedure in order to develop a consistency of approach amongst the reviewers. Dr Neale and I provided ongoing support, once the reviewers were out in the field, by email and telephone. In addition, a secure email list was set up to allow discussion of difficult cases, and a monthly newsletter sent round with a list of Frequently Asked Questions.

4.3.5 The Structured Medical Review Form

As part of the London RCRR study, Professor Vincent and Dr Neale had modified the original HMPS review form, to ensure better alignment with the typical phases of care within an NHS hospital admission.[235] The categories for harm events were also changed to reflect differences in nomenclature between the UK and the US. The London form drew on Professor Vincent's approach to incident investigations, by incorporating additional checklists for organisational and environmental contributory factors to those already in place (patient, individual, and team factors) having been identified as important underlying contributors to harm events.[243] The London form had undergone piloting across eight international clinical teams experienced in RCRR before implementation. Subsequently, two further UK-based RCRR studies and the Dutch Adverse Event study have used this form.[230 242]

Following changes made to the London form after the pilot, it was tested by Dr Neale, a third experienced reviewer, and myself, each of us taking ten records. It was also used by two trainee surgeons undertaking their own record review studies, both of whom provided valuable feedback. Finally, the reviewers themselves were invited to give feedback on the form during the training session. The form was structured to allow reviewers to progress through the review in a structured way. General demographic information and information on the patient's condition at admission were gathered in the first stage. If a problem in healthcare was found to have contributed to a patient's death, the form then led the reviewer through a series of questions documenting its type and timing, causative and contributory factors, and the preventability of the death. Information was also gathered on quality of care, the completeness of the records, the nature of missing information, and whether this impeded the review process.

4.3.6 The Review Process

Reviews were conducted on site, either in Medical Records or Clinical Governance Departments. Two reviewers were allocated to each site, and reviewed 50 case records each. The hospital lead was responsible for introducing the reviewers to the hospital site, ensuring that they were familiar with its organisational structure and had access to laboratory results and imaging if these were stored electronically.

Reviewers were required to complete a log, tracking which medical records had been reviewed and dates of transfer back to Medical Records Department (Access log). They also completed a key code that linked the patient's hospital number with a unique study number taken from a prepared list. Before commencing a review, reviewers added the patient's unique study number to the review form, along with their own reviewer ID code. No patient identifiable data were transferred onto the form. The key code document and completed forms were stored in separate locked cupboards in the review room.

Traditional RCRR is normally a two stage process. In the first stage it is usual for a nurse to screen the case record, using a defined list of sentinel events that indicate an adverse event may have occurred. Screen-positive records are forwarded to a physician for implicit review. As death is one of the screening criteria, I did not employ this initial screening step. All 1000 records underwent a complete implicit review.

Before the review commenced, reviewers were asked to check that the record was complete, that the death occurred during the index year (2009), and that the patient was not admitted for obstetric, psychiatric or paediatric care. If a post mortem report was found in the medical records, this was not to be read until the end of the review. The review focused on the admission in which the patient's death occurred. Reviewers could track back to previous admissions to uncover the origin of any harm discovered in the index admission, if this was not clear from that admission. Reviewers were asked to approach each review in a similar way:

1. Review the initial presentation with special attention to the general practitioner's referral letter, recent outpatient care, the need for admission, timeliness of initial assessment, diagnostic evaluation and management plan.

2. Review the rest of the doctor's record to determine if appropriate and timely care was given, and to evaluate the reasons for continued hospitalisation, testing and treatment. This step included review of the laboratory and radiology records to determine if important abnormalities were reported and acted on, and whether appropriate / inappropriate testing was performed.

3. Review the nursing notes and monitoring charts to determine if the management plan was adhered to and that new patient signs and symptoms were dealt with appropriately.

4. Review the medication record to determine if appropriate / inappropriate medicines were given.

Instructions were given to mark any causes for concern found in the initial read through with a sticky label to indicate the need to return to that point for a more detailed assessment later. It was recommended that the form be completed at the end of the review.

The reviewers were asked to imagine 'walking through' the case with the patient's clinical team and to ask themselves the question, 'Would this issue still be a problem if the patient had not died?' in an effort to prevent second guessing and hindsight bias. Each reviewer was provided with a manual to help guide the review process.

The record reviews were initially conducted independently. If, after full review, a reviewer was uncertain as to whether a death was caused by a problem in care, then a conversation with their reviewing partner could take place and a consensus reached. If judgement was hampered by a specialty specific question, it could be discussed with another reviewer in the group who was a specialist in this area. If there were no specialists in a particular area, then an opinion from a specialist outside the group was sought by me. Reviewers were asked to note down on the form if such advice had been sought.

On completion of all reviews at each site, reviewers discussed each case found to be a preventable death with Dr Neale to ensure any residual questions were dealt with and all the required information had been gathered from the case record.

4.3.7 Approaches to Reduction of Bias

A number of design features were built into the study that have previously been shown to improve the reliability of judgements made by reviewers. The study employed experienced physicians, reviewer training, written guidance, ongoing reviewer support and the use of a structured data collection form.[12 16 184 244] In anticipation that some reviewers might be harsher judges of the acceptability of healthcare processes than others even after training, two reviewers were sent to each site. 25% of each reviewer's case records were randomly

selected for double review, and the Kappa Coefficient calculated to provide an indicator of inter-rater reliability of judgements around life expectancy, the presence of a problem in care, and of the preventability of deaths.

4.3.8 Ethical Issues

The study was granted ethics approval by The Royal National Hospital for Neurological Diseases, and the Institute of Neurology Research Ethics Committee. One of the acute hospitals was nominated as the lead for research governance purposes and liaised with Research Governance Departments in the other nine sites to ensure all necessary permissions were sought and that 'letters of access' for the reviewers were issued.

Research based on medical records has a proven track record in providing health gains for the population, but any potential for such research to benefit society does have to be balanced against ethical concerns related to consent and confidentiality at an individual patient level. This imperative applies to the medical records of patients who have died, but for whom there may be particular challenges related to obtaining consent for record review from next of kin. For this study, tracing the next of kin of the 1000 patients in the sample who died at least one year before the study commenced was deemed impractical. Moreover, previous research led me to believe that approaching relatives for consent, following what may have been a disturbing and painful experience, would cause further and unnecessary distress.[74][245] For these reasons, I sought and was granted exemption from seeking consent from next of kin under Section 251 of the NHS Act 2006.

Measures were put in place to ensure patient and hospital confidentiality. The study complied with guidance set out in the NHS Code of Confidentiality, the London School of Hygiene and Tropical Medicine's Information Security Management Policy, and the Code of Practice for Higher and Further Education Sectors on the Data Protection Act 1998. During

reviewer training, issues of confidentiality were discussed, and the 'letter of access' issued to each reviewer contained a confidentiality clause outlining the consequences of any breach in confidentiality.

Each case study number was linked to the patient's hospital number in a key code document completed by the reviewer and kept on the hospital site. If reviewers were aware of a serious breach in standards of care and felt that this should be reported to the hospital, then it was possible for clinical governance staff to use the key code to obtain the patient's hospital number and retrieve the records for further examination. The mechanism for reporting such occurrences was agreed with each hospital before the study started, and a key contact in clinical governance or patient safety nominated. Hospitals were then able to arrange for an investigation in accordance with their normal procedures for such events.

4.3.9 Data Entry

Quantitative data from the review forms was entered into an Epidata Version 3 database before export to Stata Version 12 for analysis. Built in validation occurred on data entry. Qualitative data were entered into a Microsoft Access 2011 database. I clarified medical jargon, when necessary, for data entry personnel. The dataset included information on patient demographics and characteristics (age, sex, co-morbidity, functional status, life expectancy), admission characteristics (degree of urgency, length of admission, specialist team, ward placements, diagnosis), quality of care, and adequacy and completeness of the record. For patients experiencing problems in care, further information gathered included time and place of occurrence, categorisation (diagnosis and assessment, clinical monitoring, drugs and fluids, infections, procedures, resuscitation and other), sub-categorisation, contributory factors, and preventability ratings via the six point Likert scale. Narrative descriptions of the admission history and any problem(s) in care contributing to death were also collected. For each preventable death reviewers were asked to explain why the judgement had been made,

and to make recommendations for prevention under the categories of human, technical or organisational actions.

4.3.10 Analysis of Quantitative Data

The proportion of preventable deaths (deaths scoring 4 or above on the Likert scale) was calculated along with a sensitivity analysis, taking different thresholds of 3 (preventability not likely, less than 50:50 but close call) and 5 (strong evidence for preventability). Comparisons were made for age, sex, co-morbidities, functional ability, life expectancy and quality of care in patients with a) no problems in care, b) problems in care, and c) problems in care and preventable deaths. Comparisons of the types of problems in care associated with different phases of care, and in patients who did and did not experience a preventable death, were made. The median numbers of years of life lost as a result of a preventable death was calculated. Comparisons were made of the proportions of preventable deaths across hospitals. Descriptive statistics were used to derive means and percentages. Tests for the comparison of proportions in two independent groups corrected for binomial distribution were used, with two-sided estimation of significance and the significance level set at 0.05.

The degree of correlation between preventable death rates across the ten hospitals, and eight publicly available safety measures associated with the structure of healthcare (% staff indicating patient safety incidents were not reported as a proxy for safety culture, staff sickness absence), process (patient safety incidents, % patients reporting the hospital as not very clean, % patients reporting nurses not washing hands), and outcomes (HSMR, MRSA bacteraemia rates, emergency readmissions) were tested using Spearman's Rank Correlation Coefficient. Spearman's test is based on the correlation of the rank of values rather than the values themselves, and, unlike the Pearson Product Moment test, it avoids making assumptions about the distribution of the indicators or the linearity of the association between them.[246] It is generally regarded as more conservative than the Pearson test.

4.3.11 Content Analysis of Narrative Accounts

As a tool for understanding the epidemiology of patient harm, RCRR has generally been oriented to the quantitative analysis of the proportion of patient harm and underlying causes. However, it has been recognised for some time that the evolution of harm is often a consequence of a build up of multiple small problems, mainly omissions. These can combine across an admission to cause harm, especially in the frail and sick whose defences against such small insults are not as robust as they would be in a younger, fitter patient.[182] Approaches that can expose these threats to patient safety as efficiently as possible are needed to complement the traditional quantitative approach.

Case note review methodology has benefitted from the introduction of methods of incident analysis which derive from James Reason's organisational accident model, and which highlight both the chains of small harm events and the wider organisational and other factors which contribute to such events.[59 247-248] Such approaches, based on in-depth analysis of problems with the aim of discovering the root causes underlying their occurrence, have been most commonly employed in safety dependent industries.[249] Tools used for this purpose, particularly those developed for analysing problems and their contributory factors, might be usefully applied to narrative sections of mortality reviews to increase the utility of this information in the systematic analysis of patient harm.

In collaboration with Dr Frances Healey, Associate Director of Patient Safety at NHS Commissioning Board Authority, I developed a novel approach to content analysis, drawing on the methodology of root cause analysis,[247] and applied it to the case narratives of preventable deaths in order to explore the problems in care and contributory factors underlying such deaths. A number of different root cause analysis tools exist; some, such as nominal group technique or brainstorming, are designed to help uncover the underlying causes of harm within groups and others including 'fishbone analysis' and 'five whys' can

be used on case records, but may require the availability of more information than is available in the narratives.[250] Having noted the overall length and depth of the accounts (generally one to two paragraphs), a small number of tools were chosen which seemed feasible to use in this context.

Change analysis was used to specify problems in care by prompting consideration of what 'problem free' care might look like for a particular patient and using constant comparisons between this fictional 'problem free' care and what actually happened, as described in the narrative.[250] The problems identified in this way were categorised by type into clinical monitoring, diagnosis and assessment, drugs and fluids, technological, infection and 'other' based on definitions used by Woloshynowych *et al*.[235] The Contributory Factor Classification Framework,[243] developed by Professor Charles Vincent as part of the approach to patient incident analysis, was then used to assess the role of nine separate categories of contributory factors including patient, staff, task, communication, equipment, work environment, organisational, education and training, and team factors, and a series of sub-factors underlying each main heading.

Dr Healey and I applied the analysis to an initial five cases to test feasibility. We then met to discuss any discrepancies in our findings and make adjustments to the process. I went on to apply the method to the rest of the preventable deaths (n=47) and Dr Healey double reviewed a third of these. Levels of agreement in identifying problems in care and contributory factors were calculated using the Kappa Coefficient.

4.4 Conclusion

This chapter began by outlining the pilot study which confirmed the feasibility of the PRISM, and went on to describe the study methods in detail and actions taken to ensure both

the generalisability of findings and the reduction of known biases. The results from the study

are presented in Chapters 5, 6, and 7.

Chapter 5 Research Paper 2

5.1 Introduction to Paper 2

This paper was published in *BMJ Quality and Safety* and represents the main findings from the research included in this PhD thesis. It describes the study method and the headline findings, including, for the first time in England, an estimation of the proportion of hospital preventable deaths.

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

COVER SHEET FOR EACH 'RESEARCH PAPER' INCLUDED IN A RESEARCH THESIS

Please be aware that one cover sheet must be completed <u>for each</u> 'Research Paper' included in a thesis.

1. **For a 'research paper' already published**
   Title: Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study
   Helen Hogan, Frances Healey, Graham Neale, Richard Thomson, Charles Vincent, Nick Black

   1.1. Where was the work published? ....BMJ Quality and Safety

   1.2. When was the work published? 2012

      1.2.1. If the work was published prior to registration for your research degree, give a brief rationale for its inclusion
      ……
      ……
      ……

   1.3. Was the work subject to academic peer review? ....Yes

   1.4. Have you retained the copyright for the work? No but this paper was an open publication
      If yes, please attach evidence of retention.
      If no, or if the work is being included in its published format, please attach evidence of permission from copyright holder (publisher or other author) to include work

2. **For a 'research paper' prepared for publication but not yet published**

   2.1. Where is the work intended to be published? ....

   2.2. Please list the paper's authors in the intended authorship order

      ……

   2.3. Stage of publication – Not yet submitted / Submitted / Undergoing revision from peer reviewers' comments / In press

3. **For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)** See attached
   Professor Richard Thomson, as Head of Research and Epidemiology at the NPSA, recognised the need for a study to determine the proportion of preventable deaths in acute hospitals in England. I led on the design of the study with input from all co-authors, but particularly, Professor Charles Vincent and Dr Graham Neale, leading experts in retrospective case record review methodology. I secured grant funding for the study and managed its implementation. I was assisted by Dr Graham Neale in the design and conduct of the training for case note reviewers. Dr Neale and I also discussed all preventable deaths identified during the study. I supervised data entry and undertook data analysis and all co-authors contributed to interpretation. Dr Jenny Neuburger provided statistical advice. Under the supervision of Professor Nick Black, I was

responsible for production of the first draft of the paper. All co-authors made comments on successive drafts and approved the final version before journal submission. I acted as a guarantor of the final published version.

.

NAME IN FULL (Block Capitals) ... Dr Helen Hogan

STUDENT ID NO: 10586

CANDIDATE'S SIGNATURE .... Date 1st May 2014

SUPERVISOR/SENIOR AUTHOR'S SIGNATURE (3 above) ....

Original research

Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study

Helen Hogan,[1] Frances Healey,[2] Graham Neale,[3] Richard Thomson,[4] Charles Vincent,[3] Nick Black[1]

[1]Department of Health Services Research & Policy, London School of Hygiene & Tropical Medicine, London, UK
[2]National Patient Safety Agency, London, UK
[3]Clinical Safety Research Unit, Imperial College, London, UK
[4]Institute of Health and Society, University of Newcastle, Newcastle upon Tyne, UK

Correspondence to
Dr Helen Hogan, Clinical Lecturer in UK Public Health, Department of Health Services Research & Policy, London School of Hygiene & Tropical Medicine, 15-17 Tavistock Place, London WC1H 9SH, UK; helen.hogan@lshtm.ac.uk

Accepted 21 May 2012

ABSTRACT

Introduction: Monitoring hospital mortality rates is widely recommended. However, the number of preventable deaths remains uncertain with estimates in England ranging from 840 to 40 000 per year, these being derived from studies that identified adverse events but not whether events contributed to death or shortened life expectancy of those affected.

Methods: Retrospective case record reviews of 1000 adults who died in 2009 in 10 acute hospitals in England were undertaken. Trained physician reviewers estimated life expectancy on admission, to identified problems in care contributing to death and judged if deaths were preventable taking into account patients' overall condition at that time.

Results: Reviewers judged 5.2% (95% CI 3.8% to 6.6%) of deaths as having a 50% or greater chance of being preventable. The principal problems associated with preventable deaths were poor clinical monitoring (31.3%; 95% CI 23.9 to 39.7), diagnostic errors (29.7%; 95% CI 22.5% to 38.1%), and inadequate drug or fluid management (21.1%; 95% CI 14.9 to 29.0). Extrapolating from these figures suggests there would have been 11 859 (95% CI 8712 to 14 983) adult preventable deaths in hospitals in England. Most preventable deaths (60%) occurred in elderly, frail patients with multiple comorbidities judged to have had less than 1 year of life left to live.

Conclusions: The incidence of preventable hospital deaths is much lower than previous estimates. The burden of harm from preventable problems in care is still substantial. A focus on deaths may not be the most efficient approach to identify opportunities for improvement given the low proportion of deaths due to problems with healthcare.

BACKGROUND

Following the US Institute of Medicine's report *To err is human*,[1] the Chief Medical Officer for England estimated that 60 000 to 255 000 NHS patients each year suffer serious disability or death as a result of healthcare interventions.[2] This estimate was derived from retrospective case record review (RCRR) studies conducted in USA in the 1980s and 90s.[3][4] These and other national studies using comparable methods were not designed to establish the proportion of deaths that were preventable.[5−8]

Two smaller studies have specifically assessed the degree to which problems in care contributed to death. In one study of 111 deaths in US hospitals, reviewers judged 6% as either probably or definitely preventable.[9] A study from New Zealand concluded that 3.4% of 118 deaths were related to preventable errors in healthcare.[10] More recently, a large RCRR study in the Netherlands reported a figure of 4.1%,[11] which would be consistent with a more modest estimate of 9000 such deaths annually in England. These findings suggest that existing estimates in England based on extrapolations from studies with small numbers of deaths have overestimated preventable deaths.[10−12]

Given the considerable attention paid to hospital mortality as an indicator of quality of care,[13][14] we aimed to estimate more accurately the number of preventable deaths among hospitalised patients in England, to describe the problems in care that are responsible (type, phase of care) and to estimate the life expectancy of those affected.

METHODS

Design

RCRR is a method based on experts' retrospective reviews of healthcare records,

assessing the quality and safety of care provided during an index admission. It is the most sensitive approach in determining the proportion of hospital deaths that are preventable.[15–17] Our study design was adapted from previous RCRRs in the UK and the Netherlands,[7,11] which in turn, were based on the Harvard Medical Practice Study.[3] It also drew on a study of deaths by Hayward and Hofer.[9]

Sampling strategy

Deceased patients were identified at 10 randomly selected English acute hospital Trusts. To increase generalisability, we stratified our sampling on the basis of region (London, South, Midlands and North); teaching status; and bed size (<500, 500–700, >700) before random selection of the 10 sites from across these strata.

We estimated that 6% of deaths would be judged preventable.[9] A simple random sample would require 347 deaths to yield a 95% CI with a width of 2.5% on each side. Taking into account the two-stage sampling strategy and clustering effects at the hospital level increased the required sample size to 1000 cases. (This estimation used an intracluster correlation of 0.037 derived from the Dutch Adverse Event Study[11]).

One hundred case records of patients who had died in hospital during 2009 were randomly selected using the hospital administration system in each Trust. As in previous studies, obstetric, psychiatric and paediatric patients (who in total accounted for less than 5% of all hospital deaths in England and Wales in 2009[18]) were excluded. Of the 1000 randomly selected patients, 13 patients were admitted explicitly for planned palliative care and, therefore, were excluded and replaced.

Judgements of preventable deaths

The judgement of preventable deaths was undertaken in two stages. First reviewers were asked to judge whether there had been any problem in care that had contributed to the patient's death. Problems in care were defined as patient harm resulting from acts of omission (inactions), such as failure to diagnose and treat, or from acts of commission (affirmative actions) such as incorrect treatment or management, or harm as a result of unintended complications of healthcare. This definition was seen as more helpful than adverse event, patient safety incident, or error (box 1) because it extends beyond single discrete incidents to take a wider view of the overall quality of care provided and its contribution to a patient's death. The definition was also more likely to ensure that deaths related to failure to act (omissions) were recognised, particularly if these occurred over days or weeks.

Then, for each case where a problem in care that had contributed to death had been identified, reviewers judged the preventability of death. This two-stage

> **Box 1** Definitions previously used to describe harm due to care
>
> **Adverse event**
> An injury related to medical management, in contrast to complications of disease. Medical management includes all aspects of care, including diagnosis and treatment, failure to diagnose or treat, and the systems and equipment used to deliver care. Adverse events may be preventable or non-preventable.[19]
>
> **Patient safety incident**
> Any unintended or unexpected incident that could have or did lead to harm for one or more persons receiving healthcare.[20]
>
> **Error**
> The failure of a planned action to be completed as intended (ie, error of execution) or the use of a wrong plan to achieve an aim (ie, error of planning). Errors may be errors of commission or omission, and usually reflect deficiencies in the systems of care.[19]

approach was adopted because some problems in care contributing to death are not the result of poor practice (eg, a patient experiencing an intracerebral bleed after appropriate administration of a thrombolytic drug following myocardial infarction). Neither the problem nor the death would be regarded as preventable. In other cases where a problem in care had contributed to death, the problem may have been preventable but the patient's concurrent illness was so complex or severe that, the death itself was not judged preventable during that admission. Reviews focused on the admissions during which death occurred, but reviewers identified problems that occurred prior to that admission if these appeared to have contributed to a patient's death.

In line with previous RCRRs, reviewers assessed preventability on a 6-point Likert scale (box 2) which reflects the probabilistic nature of reviewers' decision making more closely than requiring a simple 'yes' or 'no' response.[4] The validity of this approach was demonstrated in the Harvard Medical Practice Study. Deaths were judged preventable if reviewers felt that there was more than a 50% chance the death was preventable (4–6 on the scale). This included all deaths in which reviewers judged the death was 'definitely preventable', 'strong evidence it was preventable' and 'probably preventable'. It excluded those deemed 'definitely not preventable', 'slight evidence of preventability' and 'possibly preventable but not very likely'.

The review process

The reliability of the reviews was maximised by: the use of experienced medical reviewers; providing reviewer training and written guidance; ongoing support from

> **Box 2** Scale used to judge preventability of death
>
> 1. Definitely not preventable.
> 2. Slight evidence for preventability.
> 3. Possibly preventable but not very likely, less than 50–50 but close call.
> 4. Probably preventable, more than 50-50 but close call.
> 5. Strong evidence for preventability.
> 6. Definitely preventable.

the research team with the opportunity to raise and discuss questions within the reviewer group; and the use of a structured data collection form.[4 11 21 22]

Given that previous studies have found that many problems in care are related to general clinical care processes (including omissions of care) rather than specialist technical processes and that our sample would include many patients with multiple pathologies, we required reviewers with a generalist orientation. To achieve this, 17 recently retired doctors all of whom had extensive experience as generalists (15 internal medicine and 2 general surgeons) were recruited through the Royal College of Physicians and other contacts. When necessary, specialist medical advice was available either from other reviewers within the group or from outside. This was most often used to obtain a surgical opinion.

Reviewers underwent one day of training in the review technique and could contact the principal investigator (HH) with any queries during the reviewing period. In addition, each case that was considered to be a preventable death was discussed with the principal investigator and an expert reviewer (GN). Two reviewers were allocated to each site and each reviewed 50 records to make a total of 100 for that site. Reviewers had previously had no connection with their allocated site. As reviews took place on site, they were able to request additional materials such as laboratory reports stored on computer, if these were missing from the clinical record. To determine inter-rater reliability, 25% of the records were re-reviewed by another reviewer.

Medical review form
Reviewers were asked to consider all aspects of patient care and review the entire record for the index admission, including nurses' and allied health professionals' notes, drug charts and diagnostic test results. Information was recorded by hand on a structured Medical Review Form. Demographic and clinical information on each patient included age, sex, admitting specialty (medical; surgical), type of admission (elective; emergency), comorbidity (number of conditions), and functional impairment based on the Karnofsky Performance Status Scale (none; mild; moderate; severe).[23] In all cases where a problem in care was judged to have

contributed to death, reviewers reported on the type of problem, its timing and any associated causative or contributory factors before making a judgement as to whether the death was preventable.

Reviewers estimated life expectancy on admission taking into account admitting diagnosis, functional state and degree of urgency of the admission. The use of a prognostic epidemiological tool based on survival analysis was rejected as it requires information that may not be present in case records. A similar approach to the one we adopted was used both in empirical studies of adverse events,[9 11] and in the development of tools to assess quality of care.[24]

Reviewers also rated overall quality of care by first rating each phase of care (initial assessment, treatment plan, ongoing monitoring and preparation for discharge) and then the overall quality of care on a scale from very poor to excellent, using a validated method.[25]

Analyses
Anonymised data were entered onto EpiData 3.1 and Microsoft Access databases and analysed using STATA (version 11.2) software. Demographic and health service utilisation data for the 10 hospital Trusts and for England were obtained from Hospital Episode Statistics. Summary statistics included proportions, means and medians. For all comparisons of rates, descriptive statistics and frequency tables were used, and tests for comparison of proportions in two independent groups corrected for binomial distribution.

RESULTS

Study sample characteristics
The study sample was representative of patients who die in hospital in England as regards age, admitting specialty and type of admission (table 1). Reviewers made the 'determination of a problem in care' (κ 0.54; 95% CI 0.37 to 0.71) and 'preventable death' (κ 0.49; 95% CI 0.2 to 0.8) with moderate inter-rater reliability. The wider CIs for preventable deaths reflect the fact that there were 17 preventable deaths among the 250 charts randomly selected for double review. There was substantial intra-rater agreement in assessing 'life expectancy' (weighted κ 0.66; 95% CI 0.53 to 0.79).

Patients experiencing a problem in care
In the first stage of review 131 (13.1%; 95% CI 10.9 to 15.1) patients were identified as having a problem in care that contributed to their death. There were no statistically significant differences (at $p<0.05$) in the characteristics of patients who experienced a problem in care and those that did not (n=869) as regards age, sex or comorbidity (table 2). However, on admission, patients who experienced a problem in care were more

Original research

Table 1 Comparison of study sample and all National Health Service hospital deaths in England (2009)

Characteristic	NHS population*	Study sample n=1000	p Value
Age at death %			
40–59	7.4	7.0	0.63
60–79	35.8	34.9	0.55
80–89	40.1	43.2	0.16
90–99	14.9	14.2	0.54
>100	0.5	0.4	0.15
Males %	48.7	46.3	0.13
Admitting specialty %			
Medical specialties	85.0	85.8	0.48
Surgical specialties	15	14.2	0.48
Type of admission %			
Elective	4.6	5.6	0.13
Emergency	95.4	94.4	0.13
Admission duration in days			
Median	7.5	9	0.009

*Source: Aggregated Hospital Episode Statistics Data, 2011, derived by Clinical Effectiveness Unit, The Royal College of Surgeons of England. Copyright (c) 2011. Data used with the permission of The Health and Social Care Information Centre. All rights reserved.

likely to be admitted under surgical specialties (23.6% vs 12.7%, p<0.005), as an elective admission (9.4% vs 4.5%, p<0.05) and be less severely impaired (46.3% vs 71.7%, p<0.001) than those in whom no problem was identified. Fifty-five (45.5%) of the former group were judged to have a life expectancy of more than 1-year compared with 86 (10.7%) of the latter (p<0.001).

Reviewers rated the overall quality of care received by patients to be excellent or good for 726 (73.8%, 95% CI 70.9 to 76.4) patients. Inevitably, the proportion was lower for the 131 patients deemed to have experienced a problem in care, 47 (37.0%; 95% CI 29.1 to 45.7; p<0.001) (table 3).

Preventable hospital deaths

Fifty-two deaths (5.2%; 95% CI 3.8 to 6.6) were judged to be preventable (ie, received a rating of 4–6), that is 39.7% of the 131 cases found to have a problem in care contributing to death. There were no significant differences between the proportions of preventable deaths found at each hospital. Only four (8%; 95% CI 3.1 to 18.8) of these patients were judged to have received excellent or good quality care overall and in these cases the problem in care occurred before the final hospital admission. In 33 (66.0%; 95% CI 52.2 to 77.6) patients suffering a preventable hospital death the overall quality of care was rated poor or very poor compared to 25 (2.9%; 95% CI 1.78 to 4.0) (p<0.001) in those with no problem in care.

Patients with preventable deaths were more likely to be admitted under surgical specialities (30.8% vs 13.3%,

p=0.0004). While problems occurred in all phases of care (table 4), 37 (44.0%; 95% CI 33.9 to 54.7) of the problems that contributed to a preventable death occurred during ward care. Of the rest, 13 (15.5%; 95% CI 9.3 to 24.7) patients experienced problems in care before admission (of whom five experienced no further problems in care after admission).

A wide range of types of problems were identified in patients whose death was judged to be preventable (table 5). In 73.1% (95% CI 59.7% to 83.2%) of preventable deaths more than one problem in care was identified. The most frequent problems related to clinical monitoring (31.3%; 95% CI 23.9 to 39.7), diagnosis (29.7%; 95% CI 22.5 to 38.1) and drugs or fluid management (21.1%; 95% CI 14.9 to 29.0). Clinical monitoring problems included failure to act upon results of tests or clinical findings, to set up monitoring systems, to respond to such systems or to increase the intensity of care when required. Problems with diagnosis occurred at all steps in the diagnostic process from physical examination to seeking specialist help if necessary. Examples of cases are provided in box 3.

Impact of preventable hospital deaths

If 5.2% of deaths in hospital are preventable, there would be 11 859 (95% CI 8712 to 14 983) adult preventable hospital deaths in English National Health Service (NHS) acute hospitals each year (based on 228 065 adult deaths in acute hospitals in England in 2009).[26] If a more demanding definition of preventable is employed (scores of 5 and 6 only on the Likert scales) our estimate of preventable deaths falls from 5.2% to 2.3% (5245 deaths), though this excludes deaths that reviewers thought were 'probably preventable'. Using a more relaxed definition (scores of 3 to 6 on the Likert scale, thus including 'possibly preventable but not very likely) the proportion rises from 5.2% to 8.5% (19 385 deaths).

The median estimated life expectancy of those suffering a preventable death in hospital was 6 months (IQR 4 months to 2 years) with 60% of cases having a life expectancy of <1 year.

DISCUSSION

Main findings

Among 1000 adult patients dying in acute hospitals in England, death was considered preventable in 5.2% of cases (95% CI 3.8% to 6.6%). Preventable deaths were more common among surgical admissions. The problems associated with preventable deaths occurred in all phases of hospital care but were most likely in wards (44%) and involved poor clinical monitoring (31%), diagnostic errors (30%), or inadequate drug or fluid management (21%).

Table 2 Comparison of the characteristics of patients who died having experienced a problem in care that contributed to their death with those that did not

Characteristic	Patients with problem/s in care contributing to death n=131	Patients with no problems in care contributing to death n=869	p Value
Age			
Mean (SD)	76.7 (13.4)	78.8 (12.4)	0.07
Median (IQR)	80 (50.5−83.0)	82 (73.0−87.0)	0.16
Male (%)	54 (41.2)	409 (47.0)	0.21
Admitting specialty (%)			
Medical specialties	97 (76.4))	715 (87.3)	0.01
Surgical specialties	30 (23.6)	104 (12.7)	0.01
Not known	4	50	
Comorbid conditions			
Mean (SD)	4.2 (2.4)	3.8 (2.5)	0.09
Type of admission (%)			
Elective	12 (9.4)	42 (4.5)	0.02
Emergency	116 (90.6)	795 (95.5)	0.02
Not known	3	32	
Functional impairment on admission (%)			
None*	6 (4.9)	8 (1.1)	0.02
Mild impairment†	35 (28.5)	77 (10.9)	<0.0001
Moderate impairment‡	25 (20.3)	116 (16.4)	0.28
Severe impairment§	57 (46.3)	508 (71.7)	<0.0001
Not known	8	160	
Estimated life expectancy on admission (%)			
<24 h	4 (3.3)	49 (6.1)	0.22
1−7 days	6 (5.0)	257 (32.0)	<0.0001
1−4 weeks	9 (7.4)	185 (22.9)	0.001
1−5 months	20 (16.5)	138 (17.2)	0.85
6−12 months	27 (22.3)	89 (11.1)	<0.0001
1−4 years	35 (28.9)	75 (9.3)	<0.0001
5−9 years	14 (11.6)	8 (1.0)	<0.0001
10−19 years	4 (3.3)	3 (0.4)	0.0007
>20 years	2 (1.6)	0 (0)	0.0003
Missing	10	65	
Estimated life expectancy in years on admission			
Mean (SD)	2.1 (4.3)	0.35 (1.0)	<0.0001
Median (IQR)	0.5 (0.25−2.0)	0.05 (0.01−0.25)	<0.0001

*Normal, no complaints or evidence of disease.
†Able to perform normal activity; minor signs and symptoms of disease/able to perform normal activity with effort; some signs and symptoms of disease.
‡Cares for self, unable to perform normal activity or to do active work/requires occasional assistance but is able to care for most of own needs.
§Requires considerable assistance and frequent medical care/requires special care and assistance; disabled.

Our best estimate of preventable deaths is based on a midpoint threshold on the Likert scale (categories 4−6), which may over or underestimate the actual proportion. Adopting a stricter definition in which deaths that reviewers judged to be 'probably preventable—more than 50−50 but close call' were excluded resulted in 2.3% defined as preventable but this would have excluded some preventable deaths. In contrast,

Table 3 Reviewers rating of the overall quality of care received by patients

Overall quality of care (%)	Patients with problem in care contributing to death n=131	Patients with no problems in care contributing to death n=869
Excellent	16 (12.6)	211 (24.6)
Good	31 (24.4)	468 (54.6)
Adequate	35 (27.5)	153 (17.9)
Poor	41 (32.3)	19 (2.2)
Very poor	4 (3.1)	6 (0.7)
Not known	4	12

Original research

Table 4 Phases of care during which problem in care that contributed to death occurred. (More than one option may apply for each patient)

Phase of care (%)	Preventable deaths n=52	Non-preventable deaths n=79
Before admission*	13 (15.5)	20 (19.2)
Early in admission†	19 (22.6)	14 (13.5)
Care during a procedure	8 (9.5)	21 (20.2)
Postoperative/ procedure care‡	7 (8.3)	8 (7.7)
General ward care	37 (44.0)	41 (39.4)

*General practitioner, outpatient clinic, previous admission.
†Includes assessment in the emergency department, emergency care before full assessment, admission ward, and preoperative assessment.
‡Includes high dependency or intensive care unit care.

a more relaxed definition which includes 'possibly preventable, but not very likely' resulted in 8.5% preventable though this would include deaths which are unlikely to be preventable.

These findings suggest there would have been 11 859 preventable deaths among adults in acute hospitals in England in 2009. Many of these deaths occured in elderly, frail patients with multiple comorbidities, with 60% judged to have had less than 1-year of life left to live.

Strengths and limitations of the study

Our study has a number of strengths: the large, representative sample drawn from Trusts in different regions and of different size and teaching status; our use of 'problem in care' rather than the commonly used 'adverse event' to minimise the risk of overlooking errors of omission; and the various measures to standardise data collection and ensure high quality record review.

Nonetheless, several limitations need to be considered. First, medical records may not document all problems in care, though this limitation applies to all RCRR studies, including ones that have generated previous estimates of preventable hospital deaths. Second, the estimates of life expectancy were dependent on reviewers' judgement, a notoriously difficult task. Third, RCRR studies are often criticised because of the

poor reliability of the reviewers' judgements. We used a number of approaches to improve reliability and obtained a moderately strong inter-rater agreement that compared favourably with previous studies. Some researchers have advocated using two reviewers for each case but this has not been shown to significantly improve reliability compared to employing a single reviewer.[27] Moreover, had we required agreement between two reviewers to count a case as a preventable death our estimate would have fallen to 2.8%. Thus, any problem with reliability is likely to have led to overestimating preventable deaths, not underestimating them. Another problem in RCRRs is hindsight bias, in which knowing the outcome and its severity influence the judgement of causation and preventability.[28] However, this problem would also be expected to overestimate preventable deaths, not underestimate them.

We chose to use experienced generalist reviewers rather than specialist reviewers, the majority of whom were physicians rather than surgeons. We thus ran the risk of biasing the judgement of the technical aspects of surgical care. This might have led to an underestimation of the number of preventable deaths if errors in these processes were not spotted. In fact, we found a higher proportion of both problems in care and preventable deaths among surgical patients than medical patients,

Table 5 Types of problems in care that contribute to patient death (More than one option may apply for each patient).

Type of problem in care (%)	Preventable deaths n=52	Non-preventable deaths n=79
Clinical Monitoring*	40 (31.3)	25 (18.0)
Diagnosis†	38 (29.7)	30 (21.6)
Drug or fluid related‡	27 (21.1)	30 (21.6)
Technical problem§	8 (6.3)	26 (18.7)
Infection related	9 (7.0)	22 (15.8)
Resuscitation	0 (0)	3 (2.2)
Other	6 (4.7)	3 (2.2)

*Failure to act upon results of tests or clinical findings, set up monitoring systems or respond to such systems or increase intensity of care when required.
†Missed, delayed or inappropriate diagnosis as a result of failure to perform an adequate assessment of patient's overall condition including appropriate tests or lack of focused assessment when required.
‡Side effects, inappropriate use, failure to give prophylactic care, anaphylaxis, etc.
§Related to an operation or procedure whether on ward, in a diagnostic suite or in theatre and including inappropriate or unnecessary procedures.

Box 3 Examples to illustrate nature of deaths judged to be preventable

A female patient in her early 80s presenting with watery diarrhoea where the diagnosis of inflammatory bowel disease took 18 days despite a past history of the disease. The patient had deteriorated significantly before appropriate treatment was commenced and failed to respond.

A middle aged male patient who developed infection at the site of a pharyngeal pouch excision. Antibiotic treatment was continued despite a failure to improve and subsequent open drainage proved too late.

A male patient in his 60s with previous history of ischaemic heart disease and treated carcinoma of the bladder (with no evidence of progression/ recurrence) underwent an unnecessary therapeutic ascitic tap when misdiagnosed as recurrent cancer when the actual diagnosis was congestive cardiac failure. He suffered a myocardial infarction after the procedure and went into multi-organ failure.

An obese woman in her 40s who presented with malaise, vomiting, anorexia, weight loss, early saity and night sweats. The diagnosis of ovarian malignancy took 21 days to confirm. On day 19 the patient's breathlessness and tachycardia were treated as a chest infection. Two days later she collapsed and subsequently died from pulmonary embolism. No risk assessment undertaken or thromboprophylaxis prescribed during stay.

A 30 year old man with a history of drug and alcohol use admitted with worsening shortness of breath and green sputum. Initially condition treated as a community acquired pneumonia until CT scan showed possible lung abscess or empyema. Patient developed clostridium difficile diarrhoea which delayed chest drainage and then went on to have a cardiac arrest when an attempt at drain insertion was subsequently made on the ward. Following transfer to the intensive care unit and drain insertion he continued to deteriorate and died.

A female patient in her 80s on warfarin for atrial fibrillation and admitted with an infected finger which had been treated with a combination of antibiotics by her general practitioner. Despite daily warfarin at a dose of 1mg being continued, the international normalised ratio (INR) was not checked until day 3, 1-day after blood was first noted in her stools. When the INR was found to be well above therapeutic levels at 10, vitamin K and fresh frozen plasma were administered with the clinical team commenting that a preferred treatment was not available at the time. Despite ongoing resuscitation she continued to deteriorate and died.

the majority of these being related to ward care rather than technical care. Our findings do resonate with previous reports that highlight that surgical patients do not always receive optimal management of their medical conditions.[29–31] However, it is possible that the greater risk for preventable deaths among surgical patients in our study reflects the impact of prognosis on reviewers' judgements. Reviewers typically did not judge deaths as preventable in the setting of imminent death or short life expectancy due to comorbid conditions. Patients with very short life expectancies due to underlying conditions are probably less likely to be admitted to surgical services. Consequently, surgical services have a greater proportion of patients for whom reviewers might judge problems in care judged as directly contributing to death.

Comparison with existing evidence

Our estimate of 11 859 preventable hospital deaths is similar to an estimate from the Netherlands which was based on 3983 patients dying in 25 Dutch hospitals in 2005.[11] However, our estimate is much lower than that suggested in 2000 by the Chief Medical Officer (60 000 to 255 000 serious disability or death),[2] derived from studies in USA which not only included relatively small numbers of deaths but did not examine the relationship between problems in care and death.[3 4] Our estimate is also inconsistent with suggestions of 25 000 deaths in England from venous thromboembolism,[32] if most of those are considered preventable. The difference from previous estimates is all the more surprising for two reasons: our more inclusive definition would have identified more 'problems in care' and, therefore, more preventable deaths; and the methodological limitations of this study outlined above suggest we probably overestimated the number of preventable deaths. The difference from earlier estimates appears to have arisen because these estimates were based on unjustified extrapolations.

The observation that patients were more likely to experience a problem in care if they were less functionally impaired, were elective admissions and had a longer life expectancy on admission was inconsistent with studies in other countries and might reflect a bias among reviewers towards discounting problems in the most frail, sick patients. We tried to avoid this bias by requiring reviewers to examine the entire record to the same depth and in the same structured way for all patients. Instead, we feel this finding may reflect a greater willingness in England than in some other countries to limit the extent of interventions in frail patients which would put them at less risk of experiencing a problem in care. This is inevitably speculative and would be worthy of further investigation in an international comparative study.

Implications for practice, policy and research

Although the quality of care that three-quarters of patients received was judged to be good or excellent, there is clearly plenty of scope for improvement in clinical practice. The principal area of concern is clinical monitoring on the ward. This finding is consistent with

previously voiced concerns and has already prompted various quality improvement initiatives. These include Early Warning Score Systems to avoid delay in identifying deteriorating patients,[33] explicit handover procedures to ensure vital clinical information is passed between clinicians,[34] and critical care outreach services.[35]

There are also implications for policy. While the spectre of preventable hospital deaths may prove helpful in raising interest in patient safety and a commitment to improvement, overestimating the size of the problem and the risk to patients may induce unjustified levels of anxiety and fear among the public. In addition, confirmation of the relatively small proportion of deaths that appear to be preventable provides further evidence that overall hospital mortality rates are a poor indicator of quality of care.[14]

This does not mean that preventable deaths should be ignored and no attempt made to improve our understanding of their causes. Indeed, this is one of the key areas for further research and we shall report on more detailed analyses of the type, place and timing of problems in care. Analyses will focus on clinical monitoring problems to ascertain if areas such as the early identification of deteriorating patients continue to threaten patient safety. If so, this will raise questions as to why the impact of existing initiatives has not been greater.

Mortality reviews have been adopted as a tool to identify serious harm arising from healthcare. Furthermore, given that many patients who die in hospital have been subjected to a complex series of medical interventions, studying deaths is likely to help identify a wide range of problems in care. However, it would be unwise to limit safety and quality monitoring to this relatively small proportion of patients, when the majority of problems in care may result in morbidity and disability rather than death. Further research needs to adopt a wider perspective of outcomes. There is also a need to consider other areas of secondary care, in particular preadmission care in ambulances and accident and emergency departments, and primary care where little is known about problems in care leading to serious morbidity and preventable deaths. And finally, research is required into the ways in which feedback of information on hospital mortality can be used effectively to reduce the occurrence of problems in care.

REFERENCES

1. Institute of Medicine. *To Err Is Human: Building A Safer Health Care System*. Washington DC: Institute of Medicine and National Academy Press, 1999.
2. Department of Health. *An Organisation with a Memory: Report of an Expert Group on Learning from Adverse Events in the NHS. Chaired by the Chief Medical Officer*. London: Department of Health, 2000.
3. Leape LL, Brennan TA, Laird N, et al. The nature of adverse events in hospitalized patients. Results of the Harvard Medical Practice Study I. *N Engl J Med* 1991;324:377—84.
4. Thomas EJ, Studdert DM, Burstin HR, et al. Incidence and types of adverse events and negligent care in Utah and Colorado. *Med Care* 2000;38:261—71.
5. Wilson R, Runciman W, Gibberd R, et al. Quality in Australian Health Care Study. *Med J Aust* 1995;163:472—5.
6. Davis P, Lay-Yee R, Briant R, et al. Adverse events in New Zealand public hospitals I: occurance and impact. *N Z Med J* 2002;115:U268.
7. Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001;322:517—19.
8. Baker G, Norton P, Flintoft V, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *CMAJ* 2004;170:1678—86.
9. Hayward R, Hofer T. Estimating hospital deaths due to medical error—preventability is in the eye of the reviewer. *JAMA* 2001;286:415—20.
10. Briant R, Buchanan J, Lay-Yee R, et al. Representative case series from New Zealand public hospital admissions in 1998—III: adverse events and death. *N Z Med J* 2006;119:U1909.
11. Zegers M, de Bruijne MC, Wagner C, et al. Adverse events and potentially preventable deaths in Dutch hospitals: results of a retrospective patient record review study. *Qual Saf Health Care* 2009;18:297—302.
12. McDonald CJ, Weiner M, Hui SL. Deaths due to medical errors are exaggerated in Institute of Medicine report. *JAMA* 2000;284:93—5.
13. Canadian Institute for Health Information. *Hospital Standardised Mortality: a New Approach For Measuring Hospital Mortality Trends in Canada*. Ottawa: Canadian Institute for Health Information, 2007.
14. Lilford R, Pronovost P. Using hospital mortality rates to judge hospital performance: a bad idea that just won't go away. *BMJ* 2010;340: c2016.
15. Baba-Akbari Sari A, Sheldon T, Cracknell A, et al. Sensitivity of routine system for reporting patient safety incidents in an NHS hospital: retrospective case note review. *BMJ* 2007;334:79.
16. Michel P, Quenon J, de Saraqueta A, et al. Comparison of three methods for estimating rates of preventable adverse events in acute care hospitals. *BMJ* 2004;328:199.

17. Brennan TA, Localio AR, Leape LL, *et al.* Identification of adverse events occurring during hospitalization. A cross-sectional study of litigation, quality assurance, and medical records at two teaching hospitals. *Ann Intern Med* 1990;112:221−6.

18. Office of National Statistics. *Mortality Statistics. Review of the Registrar General on Deaths by Cause, Sex and Age in England and Wales 2005*. London: HMSO, 2006.

19. World Alliance For Patient Safety. *WHO Draft Guidelines for Adverse Event Reporting and Learning Systems. From Information to Action.* Geneva: World Health Organization, 2005.

20. National Patient Safety Agency. *Seven Steps to Patient to Patient Safety*. London: National Patient Safety Agency, 2004.

21. Brennan TA, Localio AR, Laird N. Reliability and validity of judgements concerning adverse events suffered by hospitalized patients. *Med Care* 1989;27:1148−58.

22. Davis P, Lay-Yee R, Schug S, *et al.* Adverse events regional feasibility study: methodological results. *N Z Med J* 2001;114:200−2.

23. Mor V, Laliberte L, Morris JN, *et al.* The Karnofsky Performance Status Scale. An examination of its reliability and validity in a research setting. *Cancer* 1984;53:2002−7.

24. Rubenstein L, Kahn K, Harris E, *et al. Structured Implicit Review of the Medical Record: A Method for Measuring the Quality of In-hospital Medical Care and a Summary of Quality Changes Following Implementation of the Medicare Prospective Payments System. Prepared for the Health Care Financing Administration, US Department of Health and Human Services*. Santa Monica: RAND Corporation, 1991.

25. Rubenstein L, Kahn K, Harris E, *et al. Structured Implicit Review of the Medical Record: A Method for Measuring the Quality of In-hospital Medical Care and a Summary of Quality Changes Following Implemetation of the Medicare Prospective Payments System*. Santa Monica: RAND Corporation, 1991.

26. Neuberger J, Copley L. Aggregated Hospital Episode Statistics data. Clinical Effectiveness Unit, the Royal College of Surgeons of England. Copyright (c) 2011. Re used with the permission of The Health and Social Care Information Centre. All rights reserved. 2011.

27. Zegers M, de Bruijne MC, Wagner C, *et al.* The inter-rater agreement of retrospective assessments of adverse events does not improve with two reviewers per patient record. *J Clin Epidemiol* 2010;63:94−102.

28. Lilford R, Edwards A, Girling A, *et al.* Inter-rater reliability of case-note audit: a systematic review. *J Health Serv Res Policy* 2007;12:173−80.

29. The Royal College of Surgeons of England/Department of Health. *The Higher Risk General Surgical Patient: Towards Improved Care for a Forgotten Group*. London: RCSENG - Professional Standards and Regulation, 2011.

30. National Confidential Enquiry into Patient Outcome and Death. *Deaths in Acute Hospitals: Caring to the End?* London: National Confidential Enquiry into Patient Outcome and Death, 2009.

31. National Confidential Enquiry into Patient Outcome and Death. *An Age Old Problem. A Review of the Care Received by Elderly Patients Undergoing Surgery*. London: National Confidential Enquiry into Patient Outcome and Death, 2011.

32. House of Commons Health Committee. *The Prevention of Venous Thromboembolism in Hospitalsied Patients. Second Report of Session 2004-2005*. London: The Stationary Office, 2005.

33. Robb G, Seddon M. A multi-faceted approach to the physiologically unstable patient. *Qual Saf Health Care* 2010;19:e47.

34. Raduma-Tomas MA, Flin R, Yule S, *et al.* Doctors' handovers in hospitals: a literature review. *BMJ Qual Saf* 2011;20:128−33.

35. Godhill D, McNarry A. Intensive care outreach services. *Curr Anaesth Crit Care* 2002;13:346−61.

Chapter 6 Research paper 3

6.1 Introduction to Paper 3

This paper presents details of the problems in care, and contributory factors underlying preventable death. The paper has been submitted for publication in the Journal of the Royal Society of Medicine. Together with Dr Frances Healey, I developed a novel content analysis approach, which was applied to the case narratives of preventable deaths found in the Medical Review Forms. The paper highlights the role of both multiple problems in care, particularly omissions, in the generation of harm. It reveals the persistence of particular problems in care related to poor assessment, missed diagnosis, poor monitoring of warfarin and fluids, and failures to give venous thromboprophylaxis despite interventions aimed at reducing their occurrence introduced over the last decade

LONDON
SCHOOL *of*
HYGIENE
&TROPICAL
MEDICINE

COVER SHEET FOR EACH 'RESEARCH PAPER' INCLUDED IN A RESEARCH THESIS

Please be aware that one cover sheet must be completed <u>for each</u> 'Research Paper' included in a thesis.

1. **For a 'research paper' already published**

    1.1. Where was the work published? ....

    1.2. When was the work published? ....

        1.2.1. If the work was published prior to registration for your research degree, give a brief rationale for its inclusion
        ....

    1.3. Was the work subject to academic peer review? ....

    1.4. Have you retained the copyright for the work? **Yes / No**
    If yes, please attach evidence of retention.
    If no, or if the work is being included in its published format, please attach evidence of permission from copyright holder (publisher or other author) to include work

2. **For a 'research paper' prepared for publication but not yet published**

    Title: Learning from preventable deaths through narrative analysis

    2.1. Where is the work intended to be published? .... Journal of the Royal Society of Medicine

    2.2. Please list the paper's authors in the intended authorship order

    Hogan, H.; Healey, F.; Neale, G.; Thomson, R.; Black, N; Vincent, C.

    2.3. Stage of publication – Not yet submitted / Submitted / Undergoing revision from peer reviewers' comments / In press   In press

3. **For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)**
    With Dr Healey and Professor Vincent, I was responsible for the original study idea. All authors contributed to the design of the study. I was responsible for the content analysis of all cases and Dr Healey for double review of 30%. All co-authors contributed to data interpretation. I was responsible for production of the first draft of the paper. All co-authors made comments on successive drafts and approved the final version before journal submission. I acted as a guarantor of the final published version.

.

NAME IN FULL (Block Capitals) ...... Dr Helen Hogan

STUDENT ID NO: 10586

CANDIDATE'S SIGNATURE .... Date 1st May 2014

SUPERVISOR/SENIOR AUTHOR'S SIGNATURE (3 above) ....

Learning from preventable deaths through narrative analysis

Helen Hogan Clinical Lecturer in UK Public Health

Frances Healey Senior Head of Patient Safety Intelligence, NHS England, London

Graham Neale Visiting Professor, Clinical Safety Research Unit, Imperial College, London

Richard Thomson Professor of Epidemiology and Public Health, Institute of Health and Society, University of Newcastle, Newcastle upon Tyne

Nick Black Professor of Health Services Research, Department of Health Services Research & Policy, London School of Hygiene & Tropical Medicine

Charles Vincent Professor of Clinical Safety Research, Clinical Safety Research Unit, Imperial College, London

Abstract

Objectives

To explore problems in care and contributory factors in 52 patients whose death has been judged to be preventable using a new content analysis approach based on root cause analysis.

Design

A retrospective case record review of 1000 adults who died as in-patients in 2009 in ten acute hospitals in England provided 52 cases of preventable death. Change analysis, problem categorisation and the Contributory Factor Classification Framework were applied to narrative accounts of case histories written as part of the review process.

Results

An average of 3.0 (SD 1.5) problems in care and 5.2 (SD 2.5) contributory factors were identified. Almost three quarters of problems in care in both medical (71.4%) and surgical patients (72.1%) were due to omissions. Poor monitoring of laboratory tests and drugs and fluids, inadequate patient assessment, failures of the diagnostic process and drug omissions were the most common problems in care. Task and education and training factors together accounted for nearly half of the contributory factors underlying these problems.

Conclusions

Exploration of narrative accounts written as part of mortality reviews was used to systematically capture a multiplicity of problems in care and their contributory factors. Given that problems in care span initial assessment through to complex treatment,

128

improvements made as a result of such reviews are likely to provide safer environments for the majority of inpatients.

Introduction

Over the last decade, there has been a movement towards developing a more systematic understanding of causes of hospital mortality as part of a range of approaches that can be used to identify preventable harm, and so focus improvement efforts.[1] The importance of a focus on mortality been cemented in the minds of clinicians, the public and politicians following the well-publicised investigations at Bristol Royal Infirmary and Mid Staffordshire NHS Foundation Trust, both prompted by standardised hospital death rates found to be outside the expected range.[2-3] The Modernisation Agency[4] and the NHS Institute for Innovation and Improvement,[5] drawing upon the work of the US Institute for Healthcare Improvement,[6] have advocated the use of retrospective case record review (RCRR) for this purpose. The approach is also recommended by NHS national safety campaigns in both England and Wales.[7-8]

RCRR can be explicit whereby healthcare professionals assess the quality of processes of care using a set of pre-determined criteria or implicit, allowing clinicians to make judgements using their knowledge and experience. Enhancements to implicit review, such as the use of a structured review form and formal training, have been introduced over time in an effort to increase its reliability. Within the research sphere RCRR, both implicit and explicit, has usually been orientated towards quantitative analyses of either the prevalence of patient harm and underlying causes or of the percentage of patients in which a particular process was undertaken. However, it has been recognised for some time that preventable deaths are often a consequence of the build-up of multiple small problems, mainly omissions, which can combine to cause major harm especially in the frail and sick whose defences against such small insults are not as robust as those of a younger, fitter patient.[9] Approaches that can expose these threats to patient safety as efficiently as possible can complement the traditional quantitative approach.

Case note review methodology has benefitted from the introduction of methods of incident analysis which derive from James Reason's organisational accident model and are able to highlight both the chains of small events at the clinician/ patient interface and wider organisational factors which contribute to adverse events.[10-12] These approaches involve in-depth analysis of patient harm and aim to discover the root causes underlying its occurrence. We recognised that these tools, particularly those used for understanding problems in care and their contributory factors, might be usefully applied to narrative reports made by reviewers of hospital deaths to increase the utility of this information in the systematic analysis of patient harm.

A large scale retrospective case record review of 1000 acute hospital deaths has recently been conducted to provide a robust estimate of the proportion of preventable deaths in England. This provided the opportunity to develop and use a novel approach to analysis of case narratives of the 52 preventable deaths identified during the case record review, in order to determine the problems in care and contributory factors behind these deaths.

Method

Details of a retrospective case record review of 1000 hospital deaths in 2009 in ten randomly selected acute hospitals have been described elsewhere.[13] The method was based on previous similar studies.[14-18] The reviews were undertaken by 17 recently retired physicians, all of whom had extensive experience as generalists, supported by training and expert reviewer advice. For each case, reviewers were asked to provide a brief narrative account (up to one A4 page) of the circumstances.

The narrative accounts from the 52 deaths judged to be preventable were transcribed from the Review Form. A range of root cause analysis tools are available for qualitative analysis of causes and contributory factors underlying harm. We excluded techniques such as

multidisciplinary review meetings or nominal group technique that could not be applied. Two tools seemed suitable for use with written case histories. The first, change analysis asks a reviewer to specify problems in care by prompting consideration of what 'problem free' care might look like for a particular patient and then using constant comparisons between this fictional 'problem free' care and what actually happened during the admission. Any problems identified by change analysis were categorised by type (clinical monitoring, diagnosis and assessment, drugs and fluids, technological, infection and 'other') based on definitions used by Woloshynowych et al,[19] and into care delivery problems (problems that arise in the process of care at clinician/ patient level) or service delivery problems (problems related to decisions, procedures and systems at organisational level) based on National Patient Safety Agency (NPSA) definitions.[20] Charles Vincent and colleagues developed the Contributory Factor Classification Framework as a guide to determining the nature of underlying factors that enable problems in care to occur.[12] This second tool was used to categorise and subcategorise contributory factors into nine major groups, namely patient, staff, task, communication, equipment, work environment, organisational, education and training and team.

The method was applied to the first five cases to test feasibility. Two independent reviewers (HH and FH) met to discuss any discrepancies in their findings and make adjustments to the process. A third of all preventable death cases were then double reviewed by the same two investigators to test inter-rater reliability. The frequency of problems, contributory factors and their categories was calculated and descriptive statistics used to derive means and percentages. Tests for the comparison of proportions in two independent groups corrected for binomial distribution were used.

Results

Across the 52 preventable deaths, a mean of 3.0 (SD 1.5) problems in care and 5.2 (SD 2.5) contributory factors were identified. Reviewer agreement on problems in care was found in

132

71% of cases (Kappa coefficient = 0.64 indicating substantial agreement) and for contributory factors in 64% (Kappa coefficient= 0.56% indicating moderate agreement).

Problems in care causing patient deaths

71.4% of problems in care in medical patients and 72.1% in surgical patients were due to omissions such as inadequate assessment or failure to give an indicated drug. Care delivery problems (88.5%) were more frequently identified from the narratives than service delivery problems (11.5%), 61% of the latter being due to delays in tests or procedures.

Clinical monitoring problems (45) were the most frequent problems (Table 6.1). Just over a third were due to poor monitoring of laboratory tests. Poor monitoring of warfarin with subsequent bleeding was the most common monitoring problem originating prior to admission (7). Two thirds of these patients had ongoing anticoagulation problems after admission which contributed to their deaths. Inadequate monitoring of fluid balance accounted for almost a quarter of problems in this category, fluid overload in nine patients, and dehydration in two, contributing to deaths. These problems were more common in surgical patients (35.7% of all surgical monitoring problems). A combination of failure to monitor pre-operative clinical observations and pre-operative fluids led to a third of preventable surgical deaths. Figure 6.1 illustrates the distribution of problems in care across medical and surgical patients.

Table 6.1 Comparison of problem types amongst medical and surgical patients

Characteristic	Medical n=36	Surgical n=16	P values
Type of problems in care			
Clinical Monitoring[a]	31 (29.0)	14 (27.5)	0.8
Assessment and Diagnosis[b]	30(26.7)	12 (23.5)	0.3
Drugs and fluids[c]	20(21.9)	7 (13.7)	0.1
Technical[d]	11(10.5)	8 (15.7)	0.2
Infection[e]	6(5.7)	7(13.7)	0.04
Resuscitation[f]	0(0)	0 (0)	0
Other[g]	9(8.4)	3 (5.9)	0.6
	107 (100)	51(100)	

[a]Clinical monitoring: failure to set up monitoring systems, act upon changes in observations, results of tests or clinical findings or provide increased intensity of monitoring when required
[b]Assessment and Diagnosis: missed, delayed or inappropriate diagnosis due to premature closure of decision making or failure to base diagnosis on available or necessary information, failure or delay to perform an adequate assessment of the patient's overall condition including history, physical examination, appropriate tests or seeking specialist advice
[c]Drugs and fluids: omissions, inappropriate use, failure to give prophylactic care, drug interactions
[d]Technical: related to an operation or procedure whether on the ward, in a diagnostic suite or in theatre and including delayed, inappropriate or unnecessary procedures
[e]Infection: any form of hospital acquired infection
[f]Resuscitation: A problem occurring at any stage of the resuscitation process
[g]Other: problems in care not captured by any of the above definition

Assessment and diagnosis were the other frequent category (42) with 38.9% occurring early

in admission. Inadequate assessment leading to missed diagnoses, failures to secure

appropriate specialist opinions or to appreciate the risks of chosen treatments, most often

underpinned these problems. In wrong diagnoses, the case history or clinical examination

often pointed to alternative diagnoses. A cancer diagnosis led to an increased risk of

misdiagnosis. Two wrong diagnoses and two premature closures of investigation contributing to avoidable deaths in this patient subgroup. Failure or delay in diagnosing sepsis or in recognition of its extent was the predominant underlying problem in 7 of preventable deaths.

Drugs and fluids accounted for 27 problems overall with failure to prescribe an indicated drug making up 40.7%. Venous thromboembolism (VTE) prophylaxis in patients at high risk of thrombosis was the most commonly omitted drug and occurred predominantly in the early phase of admissions. Five patients subsequently died of a pulmonary embolus

Amongst the 19 problems associated with technical procedures (11 surgical procedures and 6 non-surgical ward-based procedures), organ damage as a result a poorly performed procedure occurred in 7 patients and procedure delays led to poor outcomes in 5. Thirteen problems were related to hospital acquired infection of which 6 were post-surgical or after a ward procedure. Surgical patients had significantly more infections than medical patients (13.7% vs 5.7%, p=0.04). Of the four cases of MRSA infection uncovered, two were as a result of a previous hospital procedure (hip prosthesis and biliary stent) and two arose from infected pressure ulcers during the final admission. Three of these four patients went on to die from their infection.

Serious falls were the most prevalent problems in the "other" category. Of the 6 patients who fell, two sustained a fractured neck of femur and two, significant brain injuries, contributing to their deaths. Other problems in this category were related to inappropriate admissions or ward placements compromising clinical care.

The most common problem types to cluster together were clinical monitoring, diagnosis and assessment and drugs and fluid problems (Box 6.1).

Figure 6.1 Pie Charts showing the distribution of problem in care subtypes across medical and surgical preventable deaths

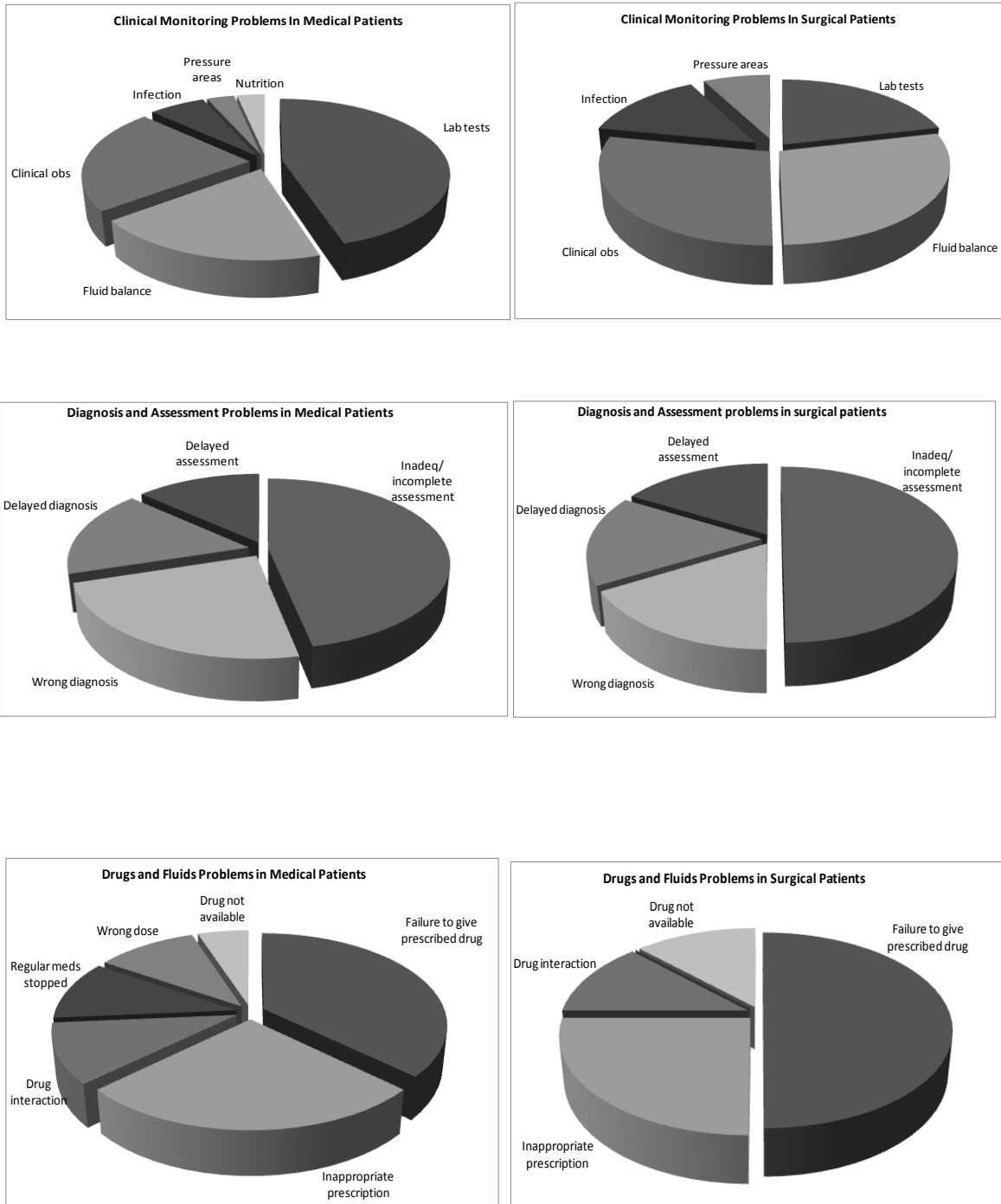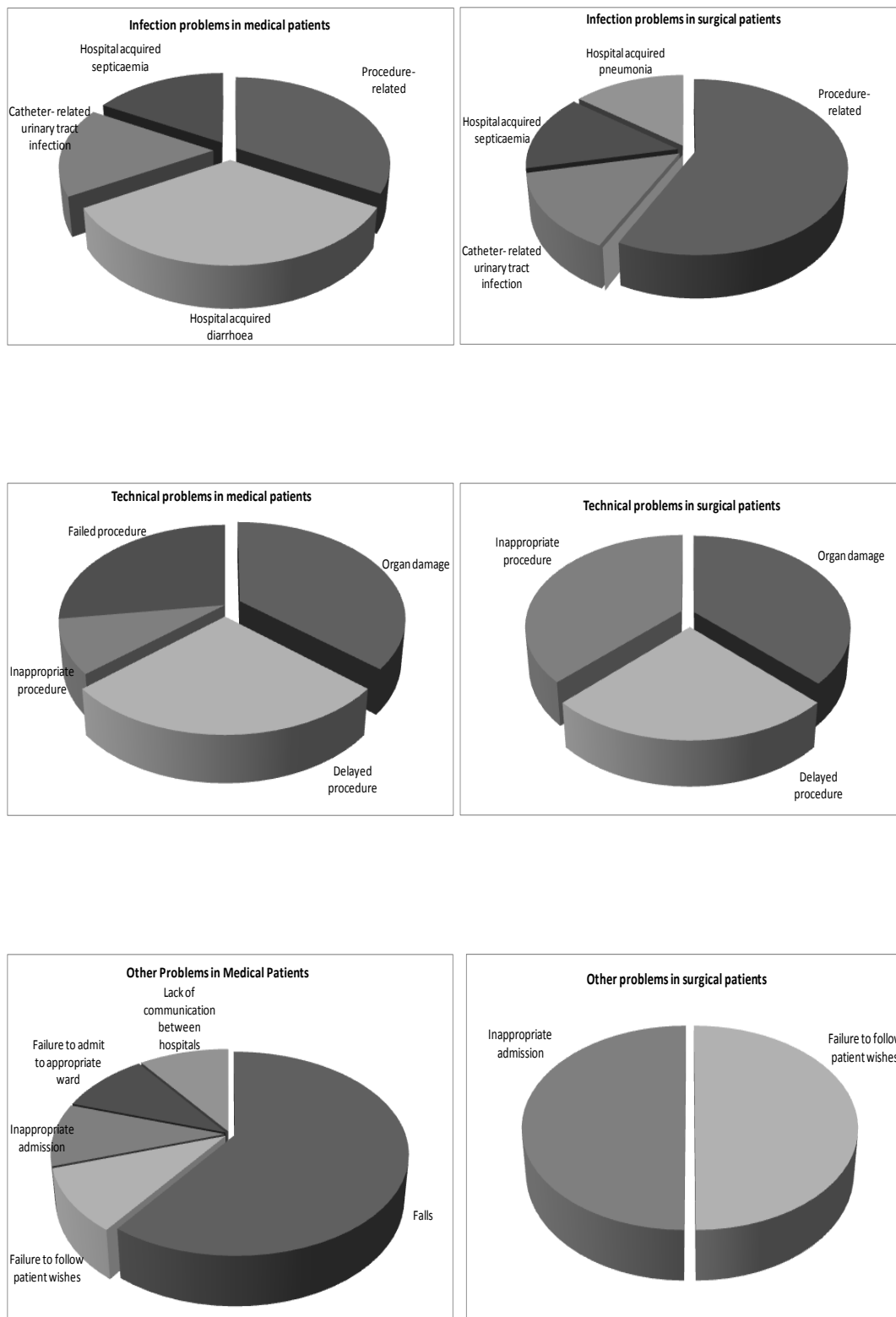**Figure 6.1 (continued) Pie Charts showing the distribution of problem in care subtypes across medical and**

Infection problems in medical patients

Infection problems in surgical patients

Technical problems in medical patients

Technical problems in surgical patients

Other Problems in Medical Patients

Other problems in surgical patients

Box 6.1 Examples of how different problems in care cluster together

An 87 year old female with a past history of stroke was admitted with a chest infection. An early CT scan showed a dilated oesophagus with food residue. She was kept nil by mouth for 5 days waiting for a swallowing assessment (**problem 1/diagnosis and assessment**). Fluid balance during that period was poorly charted (**problem 2/clinical monitoring**) but laboratory tests indicated developing dehydration. No changes in fluid regime were made in response (**problem 3/drugs and fluids**). On day 5, a trip over the drip stand (**problem 4/other**) led to a fractured femur. The patient died from post -operative renal failure, to which her poor preoperative state had contributed.

An 82 year old female on regular warfarin developed an infected finger and was prescribed two antibiotics (flucloxacillin and sodium fusidate) (**problem 1/drugs and fluids**) by her GP, leading to an increase in the coagulant effect of warfarin. On admission the patient was commenced on intravenous antibiotic treatment for osteomyelitis. Two days passed without an assessment of clotting status (**problem 2/clinical monitoring**) whist warfarin was continued at her standard dose. On day 3, the patient developed gastrointestinal bleeding and her level of anticoagulation was found to be well above the therapeutic range. The preferred treatment to reverse the effect of warfarin was not available on the ward overnight (**problem 3/drugs & fluids**) and the patient was given a second line alternative. Despite treatment, including transfusion of blood, she continued to bleed and died.

Contributory factors

Task factors, in particular task design, were the most frequent contributory factors identified in clinical monitoring problems (Figure 6.2). A typical task design factor would be the use of monitoring systems that allowed staff to enter information indicating gross fluid imbalance without automatically triggering action. Education and training factors contributed to 22.5% of monitoring problems, mainly a lack of knowledge, which often manifest as a lack of action in response to changes in a patient's condition.

Contributory factors in assessment and diagnostic problems were spread more evenly across five main factors: communication (19), staff (16), education and training (15), team (13), and task (12). Almost half (47.4%) of communication factors were ineffective flows of information impeding transfer of important clinical information within and between teams which, in turn, caused delays in arranging investigations or receiving results. Staff confirmatory/ expectation bias resulted in premature closure of the search for alternative diagnoses and failures to base the diagnosis on all the available evidence. Education and training factors particularly featured amongst surgical problems, a lack of knowledge underlying unsuccessful assessment of the severity of a patient's condition or a lack of an appropriate investigative approach. Team factors arose when two or more specialist teams were involved in the process of diagnosis and assessment and there was a lack of clarity around roles and responsibilities and task factors included decision making based on incomplete information.

For drug and fluid problems, education and training factors (21) were dominant. Lack of individual knowledge was associated with the prescription of contraindicated drugs, drugs given at the wrong dose and a lack of awareness of the impact of potential drug interactions. Task factors accounted for 26.0% of the contributory factors, with 61.5% of these being related to staff failing to follow evidence-based guidelines.

Education and Training factors (8), team factors (7) and task factors (7) were the main contributory factors associated with technical procedural problems. Lack of skills was most often associated with procedures that failed or caused organ damage when undertaken by junior staff. Almost 60% of team factors reflected issues of shared understanding and lack of leadership between different specialist teams resulting in procedure delays. Poor task design reflected cases where procedures were carried out in environments without adequate monitoring facilities or without the use of equipment which might aid better visualisation. Task factors were also the dominant contributory factor for infection.

Communication factors made the largest contribution to the "other" problem category (7). These factors featured in breakdowns of communication with external agencies such as social services leading to unnecessary admissions, to failures in coordination of cancer investigations as an outpatient or to ensuring that patients were admitted to wards best able to meet their nursing needs.

Seven patients who experienced a preventable death were known to have dementia on admission. In two of these patients, the assessment and management of their underlying conditions was not adequate, in another two, an indicated procedure were delayed. The final three patients experienced a fractured neck of femur from a fall, a pulmonary embolus after inappropriate sedation and an MRSA infection of a pressure ulcer. Three other patients who developed acute confusion during admission experienced delays in diagnosis and in receiving indicated drugs. Box 6.2 provides case examples of contributory problems.

Box 6.2   Examples of contributory factors

Monitoring: poor task design

A fluid balance chart had separate sheets for each day's fluid balance which made the accumulated overload less immediately obvious

Diagnosis and assessment: written communication

 A note of a previous significant diagnosis was not clearly visible from the case records. The fact that it was missed led to a delayed diagnosis

Drugs and Fluids: competence- lack of knowledge

Lack of knowledge of the consequences of administration of high flow oxygen to a patient with chronic obstructive pulmonary disease resulted in the patient developing respiratory failure.

Technical: competence- lack of skills

A junior doctor made several unsuccessful attempts at aspiration of a pleural effusion. The patient developed a pneumothorax, followed by a chest infection, after a chest drain was later inserted.

Infection: task

Poor management of fluid balance led to three separate insertions of a urinary catheter for output monitoring in a patient who went on to develop a urinary tract infection.

Other: communication management

A lack of communication with social services led to an unnecessary admission for a man with dementia and no indication of any other significant underlying condition. He went on to develop a pulmonary embolus after 2 weeks sedation and bed rest

Figure 6.2 Pie charts showing the distribution of contributory factors across different categories of problems in care

Distribution of contributory factors in clinical monitoring problems

Distribution of contributory factors in diagnosis and assessment problems

Distribution of Contributory Factors in Drugs and Fluids problems

Distribution of contributory factors in technical problems

Distribution of contributory factors in infection problems

Distribution of contributory factors in problems in "Other" category

Discussion

We developed a new approach for systematically analysing case narratives based on root cause analysis methods. The approach was designed to identify multiple problems care across a single admission and found an average of 3 problems and 5 contributory factors per patient, with just under three-quarters related to omissions.

Main Findings

Problems generated when processes of care go wrong were found to accumulate across the admission. Notably, problems related to anticoagulant management were the most common preadmission problems and continued to be associated with ongoing difficulties during admission. Diagnostic errors made early in admission were compounded by later failures to monitor the side effects of inappropriate medication and delays in identifying poor responses to treatment. Problems related to clinical monitoring, assessment and diagnosis and drugs and fluid problems were the major causes of preventable deaths, with half of these problems occurring during ward care. Specifically, poor monitoring of laboratory results, fluids or clinical observations, inadequate or delayed assessments and missed diagnoses were core to the genesis of almost half of all serious harm. Amongst the third largest category, drug and fluid problems, drug omissions, principally, the failure to prescribe and administer venous thromboembolism prophylaxis were the most important contributors to harm. We also found that patients with a previous diagnosis of dementia or cancer were particularly vulnerable and that failure to diagnose infection continued to put patients at risk of serious harm.

Contributory factors shed light on the issues underlying the problems in care found in preventable deaths and varied in distribution across the different problem categories. Task factors accounted for a quarter of all factors and were found most commonly amongst monitoring problems. Education and Training factors also accounted for around a quarter of all the contributory factors with a lack of knowledge being the most frequent factor for assessment and diagnosis and drug and fluid problems.

143

Lack of skill was the dominant factor amongst technical and procedural problems, contributing to organ damage caused by junior doctor operators. A combination of team, staff, communication, task and work environment factors underpinned assessment and diagnosis problems. Our findings capture the fact that clinical errors are generated within a system where there are poor methods for communication amongst and across clinical teams, inadequate levels of supervision and inadequate preparation of staff to deal with the complexity of clinical problems

Relationship to previous research and policy

Our findings confirm those from previous large RCRR studies, both in the predominance of omissions as a major factor in serious harm and the nature of the problems in care underpinning preventable deaths.[15][18][21] The problems we found are all well-known and particularly persistent. It is now two decades since the publication of a National Confidential Enquiry into Perioperative Deaths report highlighted fluid imbalance as a leading cause of serious postoperative morbidity and mortality and called for fluids to be given the same status as the prescription of other drugs.[22] As a result of the continuing poor quality of fluid management, the National Institute for Health and Care Excellence is preparing a national guideline for publication in late 2013.[23] Previous studies have also confirmed that warfarin is one of the most common causes of drug-related admissions.[24] Our observations correlate with these findings, and those from the NPSA, which found warfarin to be the third most commonly reported cause of drug-related severe harm or death.[25] On the other hand, the NPSA has also found that drug omissions form the largest category of drug-related patient safety incidents.[26] Omissions of indicated drugs such as venous thromboembolism prophylaxis continue to lead to devastating consequences despite repeated national guidance since 2007.[27]

Systematic assessment is the cornerstone of reaching the correct diagnosis, identifying if and why a patient has not responded to treatment, assessing the risk from a treatment or procedure and identifying deterioration early. There have been multiple reports implicating poor assessment in the generation of patient harm, both the lack of timely assessment and ineffectual assessments that do not

lead to appropriate responses.[28-30] Oliver recognised the problem of older patients being managed

without proper assessment and diagnosis, resulting in treatable conditions going untreated, and

labelled the term "therapeutic nihilism".[31] Failure to recognise sepsis is already a well-known cause

for concern because of its links with serious harm and preventable death.[32] The nature of diagnostic

problems are complex, our findings concur with the work of Graber *et al* who highlight the

contribution of both clinical cognitive error and system-wide factors.[33]

Across all contributory factors we found that 33.6% were related to individuals' clinical technical

performance (lack of knowledge or skills, staff cognitive processes) and 34.3% to non-technical issues

(communication or teamwork) confirming the findings of previous research which emphasised the

continued significant threats to patient safety from failures of these non-technical aspects of care.[34]

Strengths and Limitations

Examining the narratives of deaths judged to be preventable allowed a deeper understanding of the

nature of problems in care underlying such deaths and was particularly good at identifying multiple

omissions across the care pathway. However, these case stories were limited in length, ranging from

one paragraph to single sheet of A4. Missing detail is likely to have lead to a failure to identify some

problems and their contributory factors. Even with the availability of the full admission record it is

unlikely that retrospective review can find the full spectrum of hospital-related patient harm.

Clinicians are more likely to record clinical details than factors related to organisational policies and

processes, therefore reviews of records will be biased to the identification of problems with clinical-

technical aspects of care, especially those related to human error, rather than system-wide issues.[35-36]

Only 11.5% of the problems we identified were classified as service delivery problems. For the same

reasons, contributory factors are often not explicitly recorded and factors such as a lack of knowledge

have to be inferred from the nature of the problem itself. Another valid criticism would be that the

reviewers already knew they were reviewing narratives of patients who had experienced a preventable

death and with this hindsight bias, may have identified problems, even if the evidence for these was

scant. However, we attempted to decrease this bias by comparing reviewer performance early on to increase consistency and undertaking double review in a third of all cases.

Implications for practice

As the vast majority of patients who die in acute hospitals today are elderly and frail with multiple co-morbidities, hospital death reviews provide a window on how well healthcare is delivered to those with complex conditions. Such care tests the reliability of hospital systems, with fragmented and poorly co-ordinated care increasing the opportunity for omissions and ensuing harm, especially in those with fragile health states.[3] Mortality reviews can highlight key areas of risk as patients move through the care process thus allowing more focused targeting of resources to reduce these risks. We found over 70% of problems in care to be due to omissions which would suggest that one focus for improvement should be improving the reliability of delivery of care processes. Although initiatives such as national audits have stimulated improvements in this domain, within certain specified areas, a move beyond a patchwork approach to a "whole system" approach would be more likely to have a bigger impact. Human failure remains an important contributor to monitoring, assessment and diagnosis, drug and technical procedure-related problems. Urgent consideration is needed of how adequate senior supervision and support for junior staff can be provided, in increasingly cost constrained environments, to reduce such harm.

Conclusion

Mortality reviews highlight the continued harm generated by failures of early assessment and diagnosis, poor clinical monitoring of warfarin and intravenous fluids and failures to give indicated drugs. Such reviews offer the opportunity to identify high risk areas and target appropriate interventions. Given that problems in care span initial assessment through to complex treatment,

improvements made as a result of mortality reviews are likely to provide safer environments for all patients.

Competing Interests

None declared.

Funding

Ethical Approval

Ethics approval was received from the National Hospital for Neurology and Neurosurgery and the Institute of Neurology joint multi-centre research ethics committee and research governance approval was granted by each participating Trust.

Guarantor

HH is guarantor.

Contributorship

HH, FH and CV were responsible for the original study idea. All authors contributed to the design of the study. HH was responsible for the content analysis of all cases and FH for double review of 30%.

HH and FH were responsible for data interpretation. HH drafted the manuscript and all authors contributed to its revision.

Acknowledgements

References

1. Sari AB, Sheldon TA, Cracknell A, Turnbull A, Dobson Y, Grant C, et al. Extent, nature and consequences of adverse events: results of a retrospective casenote review in a large NHS hospital. *Qual Saf Health Care* 2007;16(6):434-9.

2. Department of Health. Learning from Bristol: the report of the public enquiry into children's heart surgery at the Bristol Royal Infirmary 1984-1995. London: HMSO, 2001.

3. Francis R. Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry. London: The Stationary Office, 2013.

4. Modernisation Agency. A matter of life and death. Increasing reliability and quality to reduce hospital mortality and improve end of life care. London: Department of Health, 2004.

5. NHS Institute for Innovation and Improvement. Reducing avoidable deaths (Medical Directors): Department of Health, 2007.

6. Institute for Healthcare Improvement. Move Your Dot. Measuring, evaluating and reducing hospital mortality rates. Innovation Series 2003. Massachusetts, USA: Institute for Healthcare Improvement, 2003.

7. 1000 Lives Plus. A guide to measuring mortality. Cardiff: 1000 Lives Plus, 2010.

8. Patient Safety First Campaign. Patient Safety First 2008 to 2010. The campaign review. London: National Patient Safety Agency, 2011.

9. Mold JW, Stein HF. The cascade effect in the clinical care of patients. *N Engl J Med* 1986;314(8):512-4.

10. Spath P, editor. *Error reduction in healthcare: a systems approach to improving patient safety*. Washington: AHA Press, 1999.

11. Reason JT. *Human Error*. New York: Cambridge University Press, 1990.

12. Adams S, Vincent C. Systems analysis of clinical incidents: the London protocol. Available at: http://www1.imperial.ac.uk/resources/C85B6574-7E28-4BE6-BE61-E94C3F6243CE/londonprotocol_e.pdf. Accessed 20th January 2013.

13. Hogan H, Healey F, Neale G, Thomson R, Vincent C, Black N. Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study. *BMJ Qual Saf* 2012;21(9):737-45.

14. Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. 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.

15. Wilson R, Runciman W, Gibberd R, Harrison B, Hamilton J. Quality in Australian Health Care Study. *Med J Aust* 1995;163(9):472-5.

16. Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001;322:517-9.

17. Hayward R, Hofer T. Estimating hospital deaths due to medical error- preventability is in the eye of the reviewer. *JAMA* 2001;286(4):415-20.

18. Zegers M, de Bruijne MC, Wagner C, Hoonhout LH, Waaijman R, Smits M, et al. Adverse events and potentially preventable deaths in Dutch hospitals: results of a retrospective patient record review study. *Qual Saf Health Care* 2009;18(4):297-302.

19. Woloshynowych M, Neale G, Vincent C. Case record review of adverse events: a new approach. *Qual Saf Health Care* 2003;12:411-5.

20. National Patient Safety Agency: National Reporting and Learning System. Root Cause Analysis Tools: Guide to report writing following Root Cause Analysis of Patient Safety Incidents. London: National Patient Safety Agency, 2008.

21. Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, et al. 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.

22. National Confidential Enquiry into Perioperative Deaths. Extremes of age. London: NCEPOD, 1999.

23. National Institute of Health and Clinical Excellence: Centre for Clinical Practice. Scoping document: Intravenous fluid therapy in adults in hospital. London: NICE, 2013.

24. Howard RL, Avery AJ, Slavenburg S, Royal S, Pipe G, Lucassen P, et al. Which drugs cause preventable admissions to hospital? A systematic review. *Br J Clin Pharmacol* 2007;63(2):136-47.

25. Cousins D, Gerrett D, Warner B. A review of medication incidents reported to the National Reporting and Learning System in England and Wales over six years (2005 – 2010). *Br J Clin Pharmacol* 2012;74(4):597-604.

26. Cousins DH, Gerrett D, Warner B. A review of medication incidents reported to the National Reporting and Learning System in England and Wales over 6 years (2005-2010). *Br J Clin Pharmacol* 2012;74(4):597-604.

27. Department of Health. Report of an independent expert working group on the prevention of venous thromboembolism in hospitalised patients. A report to Sir Liam Donaldson, Chief Medical Officer. London: Department of Health, 2007.

28. Seward E, Greig E, Preston S, Harris RA, Borrill Z, Wardle TD, et al. A confidential study of deaths after emergency medical admission: issues relating to quality of care. *Clin Med* 2003;3(5):425-34.

29. National Confidential Enquiry into Patient Outcome and Death. An acute problem? London: NCEPOD, 2005.

30. National Confidential Equiry into Patient Outcome and Death. Time to Intervene? London: NCEPOD, 2012.

31. Oliver D. 'Acopia' and 'social admission' are not diagnoses: why older people deserve better. *J R Soc Med* 2008;101(4):168-74.

32. Slade E, Tamber PS, Vincent JL. The Surviving Sepsis Campaign: raising awareness to reduce mortality. *Crit Care* 2003;7(1):1-2.

33. Graber ML. Expanding the Goals of Peer Review to Detect Both Practititoner and System Error. *Journal on Quality Improvement* 1999;28(8):396-407.

34. Scanlon MC, Karsh BT. Value of human factors to medication and patient safety in the intensive care unit. *Crit Care Med* 2010;38(6 Suppl):S90-6.

35. Smits M, Zegers M, Groenewegen PP, Timmermans DR, Zwaan L, van der Wal G, et al. Exploring the causes of adverse events in hospitals and potential prevention strategies. *Qual Saf Health Care* 2010;19(5):e5.

36. Bates D, Cullen D, Laird N. Incidence of adverse drug events and potential adverse drug events. Implications for prevention. *JAMA* 1995;274:29-34.

Chapter 7 Research Paper 4

7.1 Introduction to Paper 4

This paper explores the association between preventable deaths identified by RCRR and a range of publicly available safety indicators. It has been submitted for publication in the *International Journal of Quality in Health Care.* My conceptual framework identifies different measures of patient safety according to Donebedian's distinction between structure (inputs), process or outcomes.[32] This study was undertaken to illuminate the degree of correlation between a range of measures of safety, and the proportion of preventable hospital deaths. The hypothesis was that there should be an association. A lack of association would reflect that these measures assess distinctly different aspects of hospital safety, and indicates that caution should be applied when using a single primary measure to reflect the whole patient safety domain. This study was only exploratory, given a sample size of only ten hospitals.

Using Spearman's rank correlation coefficient to test these associations, only MRSA bacteraemia rate was significantly associated with the proportion of preventable hospital deaths (r=0.73, p=0.02), emphasising the prominent role played by infection in the causation of serious harm. There was a suggestion that the strength of association declined from outcomes (MRSA) to process (hand hygiene 0.51) to structure (patient safety culture 0.26). These findings require confirmation in a larger study. No significant correlation was found with HSMR, which is consistent with other studies.

COVER SHEET FOR EACH 'RESEARCH PAPER' INCLUDED IN A RESEARCH THESIS

Please be aware that one cover sheet must be completed <u>for each</u> 'Research Paper' included in a thesis.

1. **For a 'research paper' already published**

   1.1. Where was the work published? ....

   1.2. When was the work published? ....

       1.2.1. If the work was published prior to registration for your research degree, give a brief rationale for its inclusion

   ....

   ....

   ....

   1.3. Was the work subject to academic peer review? ....

   1.4. Have you retained the copyright for the work? **Yes / No**
   If yes, please attach evidence of retention.
   If no, or if the work is being included in its published format, please attach evidence of permission from copyright holder (publisher or other author) to include work

2. **For a 'research paper' prepared for publication but not yet published**

   Title: Relationship between preventable hospital deaths and other measures of safety: an exploratory study

   2.1. Where is the work intended to be published? .... International Journal of Quality in Healthcare

   2.2. Please list the paper's authors in the intended authorship order

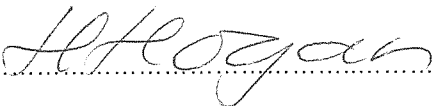   Hogan, H.; Healey, F.; Neale, G.; Thomson, R.; Vincent, C; Black, N.

   2.3. Stage of publication – Not yet submitted / Submitted / Undergoing revision from peer reviewers' comments / In press   In press

3. **For multi-authored work, give full details of your role in the research included in the paper and in the preparation of the paper. (Attach a further sheet if necessary)**

   Professor Nick Black and I had the original idea for the analysis. I identified the safety measures for the correlation with the approval of the co-authors. I was responsible for gathering safety measure values for each hospital and undertaking the correlation. Dr Jenny Neuburger and Mr Andrew Hutchings provided statistical advice. With the support of Professor Black, I produced the first draft of the paper. All co-authors made comments on successive drafts and approved the final version before journal submission. I acted as a guarantor of the final published version.

NAME IN FULL (Block Capitals) ......Dr Helen Hogan

STUDENT ID NO: 10586

CANDIDATE'S SIGNATURE .... Date 1st May 2014

SUPERVISOR/SENIOR AUTHOR'S SIGNATURE (3 above) ....

Relationship between preventable hospital deaths and other measures of safety: an exploratory study

Helen Hogan Clinical Lecturer in UK Public Health

Frances Healey Joint Head of Clinical Review and Response, NHS England, London

Graham Neale Visiting Professor, Clinical Safety Research Unit, Imperial College, London

Richard Thomson Professor of Epidemiology and Public Health, Institute of Health and Society, University of Newcastle, Newcastle upon Tyne

Charles Vincent Professor of Clinical Safety Research, Clinical Safety Research Unit, Imperial College, London

Nick Black Professor of Health Services Research, Department of Health Services Research & Policy, London School of Hygiene & Tropical Medicine

Abstract

Objectives

To explore associations between the proportion of hospital deaths that are preventable and other measures of safety. Our hypothesis was that the association would be strongest with other measures of outcome, but progressively less strong with process and structure measures.

Design

A retrospective case record review of 1000 adults who died as in-patients in 2009 in ten acute hospitals in England provided estimates of preventable death proportions. The relationship with eight other measures of patient safety was explored by estimating simple monotonic correlations using Spearman's rank correlation coefficient.

Results

The proportion of preventable deaths varied between hospitals (3% - 8%) but was not statistically significant ($p=0.94$). Only one of the eight measures of safety (MRSA bacteraemia rate) was clinically and statistically significantly associated with preventable death proportion ($r=0.73$; $p<0.02$). There were no significant associations with the other measures including hospital standardised mortality ratios ($r = -0.01$). Despite this, there was a suggestion that preventable deaths may be more strongly associated with some other measures of outcome than with process or with structure measures

Conclusions

The exploratory nature of this study inevitably limited its power to provide definitive results. The observed relationships between safety measures suggest a larger more powerful study is needed to establish the inter-relationship of different measures of safety (structure, process, outcome), in particular the widely used standardised mortality ratios

Key words: preventable death, patient safety measures, HSMRs

Introduction

A wide variety of measures are used to assess the safety of hospitals.[1] They can be grouped into three broad categories reflecting Donabedian's typology of outcomes, processes and structures (sometimes referred to as inputs).[2] Outcomes include preventable mortality, hospital acquired infections and emergency readmissions. Processes include events that might result in an adverse outcome such as patient safety incidents (e.g. falls and medication errors) and structure, aspects such as the safety culture of a hospital and the attitudes of staff towards safety. It is unclear whether or not these different dimensions of poor safety are associated with one another, an association that would suggest common cause. This is of considerable policy importance as if there is little or no association then the choice of a primary or leading measure of safety of a hospital will influence judgments about its performance.

Over the past decade, the most commonly used primary measure in many countries has been the standardised mortality ratio (SMR) for the entire hospital. In England the most commonly used versions have been the Hospital Standardised Mortality Ratio (HSMR) and, more recently, the Standardised Hospital-Level Mortality Indicator (SHMI). Their use has had and continues to have an enormous influence on health care policy, despite the lack of any evidence of their validity either as accurate indicators of safety or as a screening tool to raise suspicions of poor safety. [3-5]

Arguably, from the point of view of patients, the public, politicians and staff, the most important and credible indicator of hospital safety is preventable mortality, when preventability is determined by trained reviewers undertaking retrospective case record review. It has two important attributes: clinical credibility by taking account of the complexity of patients' conditions and care, and it can indicate whether or not poor care was responsible for any death. In addition clinicians identify the nature of any poor care and this starts to instil improvements in clinical practice.

Only four studies, all in North America, have looked at the association of preventable death proportions with other frequently used measures of safety. All four studies focused on the relationship with SMRs either for selected specialties,[6-7] specific diseases,[8] or a specific intervention,[9] rather than for hospital-wide deaths. The studies were limited to considering aggregated data from groups of high and of low SMR hospitals due to the small samples from individual hospitals (less than 50). Three of the studies either found no correlation,[6-7] or a non-significant negative correlation.[9] The fourth study, which reviewed patients with one of three medical conditions also found no association for two conditions (stroke and myocardial infarction) but did find a positive association in patients with pneumonia. [8]

To date, comparing the proportion of preventable deaths with other measures of safety in England has been limited by the lack of an accurate estimate for the proportion of deaths in hospitals that were preventable. An opportunity to carry out an initial exploration of this key issue has arisen with the availability of data collected in a recent large retrospective case record review.[10] The rigour with which the measurements were made suggests that they provide a valid and credible indication of safety.

Our aim was to carry out an exploratory study of the associations between the proportion of deaths in a hospital that are deemed preventable and other measures of safety. Our hypothesis was that the association would be strongest with other measures of outcome (HSMR, hospital acquired infections, emergency readmissions), less strong with measures of process (patient safety incidents; hospital cleanliness; staff hand hygiene) and weakest with structure measures (safety culture, staff sickness absence).

Method

Preventable deaths in hospital

Details of the retrospective case record review of 1000 hospital deaths in 2009 carried out in ten randomly selected acute hospitals have been described elsewhere.[10] The method was based on previous similar studies.[11-14] Record reviews were undertaken by 17 recently retired doctors, all of whom had extensive experience as generalists and received training for the task.

The judgement of preventable deaths was undertaken in two stages. First, reviewers were asked to determine whether there had been any 'problems in care' that had contributed to the patient's death. Problems in care were defined as patient harm resulting either from acts of omission (inactions), such as failure to diagnose and treat, or from acts of commission (affirmative actions) such as incorrect treatment or management. Then, for each case in which a problem in care that had contributed to death was identified, reviewers judged the preventability of death. Preventability was assessed on a six point Likert scale.[15] Deaths were deemed preventable if it was judged that there was more than a 50% chance that the death was preventable (scored from 4 to 6 on a 1 to 6 Likert scale). For deaths judged to be preventable, reviewers reported the type of problem, its timing and any associated causative or contributory factors.

Other patient safety measures

A priori, we selected measures of safety that were publicly available and reflected safety across entire hospitals rather than restricted to specific departments. The measures were:

- Outcomes: HSMR, obtained from Dr Foster Intelligence; MRSA bacteraemia reports, from the Health Protection Agency; emergency readmissions within 28 days of discharge, from Hospital Episode Statistics

- Processes: patient safety incidents, reported to the National Reporting and Learning System; patients' views of hospital cleanliness and of nurses' hand cleaning, obtained from the NHS Inpatient Survey

- Structure: staff views of safety culture, obtained from the NHS Staff Survey; staff sickness absence rates, from the NHS Staff Sickness Reports

Details of these data sources and the validity of the measures are shown in Table 7.1 and how they varied across the hospitals in Table 7.2. It is important to recognise that all eight measures are inevitably subject to chance variation in addition to the specific limitations mentioned in the table. All of these data were publicly available at hospital-level on the internet.

Analyses

Median and inter-quartile ranges for each patient safety measure were calculated to show the distribution of values across the hospitals. Simple monotonic correlations between preventable death proportions and each of the other safety measures for the ten hospitals were examined using Spearman's rank correlation coefficient. The level of what would be deemed a clinically significant association was set as a correlation coefficient of at least 0.3. Tests for significance were two-sided and the significance level set at 0.02, given the multiple comparisons being tested.

Table 7.1: Sources and description of eight safety measures

	Source	Description of measure	Collection	Threats to Validity
Outcomes				
Hospital Standardised Mortality Ratio (HSMR)	Dr Foster 2009/10 (http://www.drfosterhealth.co.uk/docs/hospital-guide-2010.pdf)	HMSRs are calculated using Hospital Episode Statistics (for 56 conditions known to be related to 80% of hospital mortality) by calculating the ratio of observed deaths to expected deaths with adjustment for case mix.	Hospital Episode Statistics are derived from Patient Administration Systems. Expected death rates are calculated using national data. Adjustments are made for age, co-morbidity, number of previous admissions and sociodemographic factors	Low sensitivity for measuring quality of care (most quality problems do not occur in patients who die) Low specificity for measuring quality of care (most deaths don't reflect poor quality) Artifactual variation for a range of reasons including coding depth and patient exclusions Structural factors e.g. local services available, admission thresholds and technology/ treatments available cause variation
Methicillin Resistant Staph Aureus (MRSA) bacteraemia rates	Health Protection Agency: MRSA surveillance programme 2009/10 (http://www.hpa.org.uk/web/HPAweb&HPAwebStandard/HPAweb_C/1233906818165)	Hospital apportioned MRSA bacteraemia reports per 100,000 admissions. A positive blood culture on or after the third day of admission is classified as hospital apportioned.	Mandatory surveillance of MRSA bacteraemia conducted by Health Protection Agency based on hospital submitted reports of positive blood cultures and accompanying demographic and clinical data. Data submitted via web-based system.	Some MRSA infections occurring prior to admission or recurrent infections may be included Rates are not adjusted for hospital demographics or case mix
Emergency readmissions	The NHS Information Centre 2009/10 (https://mqi.ic.nhs.uk/PerformanceIndicatorChapter.aspx?number=1.01)	Percentage of emergency admissions to hospital in England for adults that occur within 28 days of the previous discharge (indirectly standardised)	Hospital Episode Statistics derived from Patient Administration Systems.	Some readmissions result from of lack of primary/ community services Quality of coding Admission decision dependent on variation in clinical judgment of the admitting doctor

	Source	Description of measure	Collection	Threats to Validity
Processes				
Patient safety incident reports	National Reporting and Learning System 2009/10 (www.nrls.npsa.nhs.uk)	Overall reporting rate per 100 admissions.	Voluntary self reports are received by the NRLS via downloads from local risk management systems or web-based e- forms (including open access e-forms).	Individual reports are not investigated or verified by NPSA Quality/ volume of data variable depending on reporting system used and reporter Counts based on incidents reported: known under reporting Variability in reporting rates may not reflect safety but organisational culture Some reports are not PSIs e.g. misreporting of harm to staff or lost patient property Rates are not adjusted for hospital demographics or case mix
Patients' views of hospital cleanliness	Acute Hospitals Adult Inpatient Survey 2009, Economic and Social Data Service, http://www.esds.ac.uk/findingData/snDescription.asp?sn=7034&key=Acute+Trusts:+Adult+Inpatients+Survey,+2010	Proportion of patients at each hospital giving negative responses to question related to general cleanliness of the hospital	Annual Survey of patient experience commissioned by the Care Quality Commission. Each hospital identifies a random sample of 850 adult and psychiatry patients who at least one night.	Response rate around 50%. Patient exclusions include maternity

|	Source	Description of measure	Collection	Threats to Validity
Patients' views of staff hand hygiene	Acute Hospitals Adult Inpatient Survey 2009, Economic and Social Data Service, http://www.esds.ac.uk/findingData/snDescription.asp?sn=7034&key=Acute+Trusts:+Adult+Inpatients+Survey,+2010	Proportion of patients at each hospital giving negative responses to question related to whether nurses washed their hands between patients	Annual Survey of patient experience commissioned by the Care Quality Commission. Each hospital identifies a random sample of 850 adult and psychiatry patients who at least one night.	Response rate around 50%. Patient exclusions include maternity
Structures				
Staff view of reporting of patient safety incidents	The Staff Survey Coordination Centre, Picker Institute Europe: NHS Staff Survey 2009 (http://www.NHSStaffSurveys.com)	Proportion of staff at each hospital giving a negative response to question: "The last time that you saw an error, near miss or incident that could have hurt patients / service users, did you or a colleague report it?"	Annual survey of staff views. Random sample of staff based on the size of the institution. Analysed by Staff Survey Coordination Centre. Hospital management do not see individuals' completed surveys but are sent amalgamated results for their hospital following analysis.	Wording of some questions is ambiguous Response rates range from 40%-65%
Staff sickness absence	The NHS Information Centre: NHS Electronic Staff Records 2009/10)(http://www.ic.nhs.uk/statistics-and-data-collections/workforce/sickness-absence)	Annual staff sickness absence rate	Data collected monthly from the Electronic Staff Record System which links to the payroll and human resource systems within hospitals and contains records for the majority of NHS staff. The rates are calculated using Full Time Equivalent (FTE) days lost to sickness divided by the FTE days available	Gives overall measure of sickness absence amongst NHS staff but does not indicate which staff and in which roles

Table 7.2: Hospital characteristics and patient safety indicator values, 2009

PRISM Trusts	A	B	C	D	E	F	G	H	I	J
Bed Numbers	871	393	1449	998	1108	693	999	628	483	417
Annual Admissions	100,828	37,345	171,954	111,003	141,166	94,961	117,727	76,873	55,238	51,756
No. Adult Coronary Care Unit beds	32	7	62	67	58	10	24	12	10	7
Hospital type	Large acute	Small acute	Acute teaching	Acute teaching	Large acute	Large acute	Large acute	Medium acute	Small acute	Small acute
Preventable deaths (%)	0.04	0.06	0.05	0.06	0.05	0.03	0.08	0.04	0.06	0.05
Hospital Standardised Mortality Ratio	107.6	97.8	79.6	96.8	96.0	102.1	107.4	89.4	90.3	112.0
MRSA bacteraemia rates per 100,000 admissions	1.0	4.2	3.4	6.2	2.9	0.9	3.3	1.6	4.2	0.7
Emergency readmissions within 28 days of discharge (%)	12.4	11.5	13.2	12.0	13.0	10.1	9.5	10.2	12.7	9.6
Patient safety incidents per 100,000 admissions	6298.8	4236.2	4869.3	6316.9	3536.3	5255.8	5971.4	3903.8	4764.8	4134.8
Patients reporting hospital 'not very clean'/ 'not clean at all' (%)	4.1	4.2	5.2	2.4	4.6	2.9	2.9	1.7	2.5	5.3
Patients reporting hospital not cleaning their hands between patients (%)	2.2	2.6	4.9	4.2	2.1	2.6	2.8	1.5	3.6	2.8
Staff indicating that patient safety incidents were not reported (%)	33	34	34	40	36	34	35	36	34	35
Staff sickness absence rate (%)	3.2	4.4	2.8	3.5	4.7	4.5	4	3.1	3.7	4.1

Results

Among 1000 adult patients dying in acute hospitals in England, death was considered preventable in 5.2% of cases (95% CI 3.8% - 6.6%). The proportion varied between hospitals from 3% to 8% but these differences were not statistically significant (p= 0.94) (Figure 7.1). Table 7.3 shows the distribution of the safety measures across the ten hospitals.

Figure 7.1: Proportion of preventable deaths across ten English acute hospitals
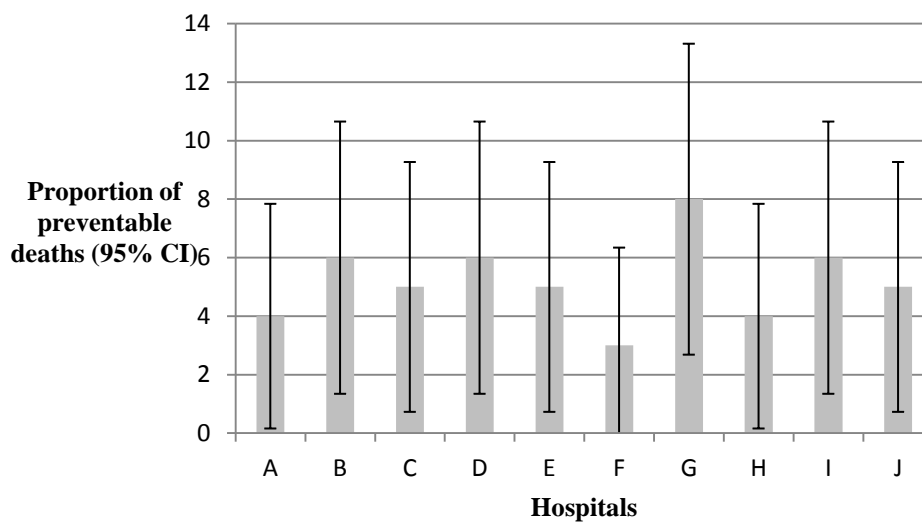
The relationships between preventable deaths and the other measures of safety are shown in Table 7.4. Only one association was clinically and statistically significant: there was a positive correlation between preventable death proportion and MRSA bacteraemia rate (r = 0.73; p = 0.02). (Figure 7.2) Although a positive association was also observed with one other measure (nurses not cleaning their hands between patients r = 0.51) and a weak positive relationship with two other measures (staff indicating that adverse events were not reported r = 0.26; patient safety incidents r = 0.23) none were statistically significant. As regards the other four measures (HSMR, emergency readmission, hospital cleanliness, staff sickness absence), there was no evidence of an association with preventable deaths. Given that previous studies have compared groups of hospitals with high SMRs with those with low SMRs, we did the same by aggregating data from the hospitals with the three highest

and the three lowest SMRs. There was no significant difference: 5.6% v 5.0% respectively (p=0.74).

As regards our hypothesis based on Donabedian's categories of outcome, process and structure, there was a suggestion that the strength of association declined from 0.73 with another outcome (MRSA bacteraemia), to 0.51 with a process measure (staff hand hygiene), to 0.26 with a structure measure (poor safety culture) but the lack of statistical significance of the latter two correlations means such an observation must be treated cautiously.

Table 7.3: Distribution of patient safety measure values across ten acute hospitals

Safety measure	Median	Inter quartile Range
Preventable deaths (%)	5.00	4.00-6.00
Outcomes Hospital standardised mortality ratio	97.30	91.73-106.10
MRSA bacteraemia rates per 100,000 admissions	3.10	1.15-4.0
Emergency readmissions within 28 days of discharge (%)	11.74	10.15-12.59
Processes		
Patient safety incidents per 100,000 admissions	4317	4160-5793
Patients reporting hospital 'not very clean' or 'not at all clean' (%)	3.53	2.58-4.50
Patients reporting nurses did not clean their hands between patients (%)	2.67	2.30-3.37
Structures		
Staff indicating that patient safety incidents were not reported (%)	34.68	33.84-35.85
Staff sickness absence rate (%)	3.85	3.28-4.33

Table 7.4: Correlations between preventable deaths and other patient safety measures

Safety measure	Spearman correlation coefficient	Lower confidence limit	Upper confidence limit	P value
Outcomes				
Hospital standardised mortality ratio	-0.012	-0.64	0.62	0.97
MRSA bacteraemia rates per 100,000 admissions	0.73	0.19	0.93	0.02
Emergency readmissions within 28 days of discharge (%)	-0.06	-0.66	0.59	0.86
Processes				
Patient safety incidents per 100,000 admissions	0.23	-0.47	0.75	0.52
Patients reporting hospital 'not very clean'/'not at all clean' (%)	-0.08	-0.68	0.58	0.80
Patients reporting nurses not cleaning hands between patients (%)	0.51	-0.17	0.86	0.12
Structures				
Staff indicating patient safety incidents were not reported (%)	0.26	-0.44	0.76	0.47
Staff sickness absence rate	0.06	-0.59	0.66	0.86

Figure 7.2: Scatter plot of hospital apportioned MRSA bacteraemia rates and hospital preventable death proportion
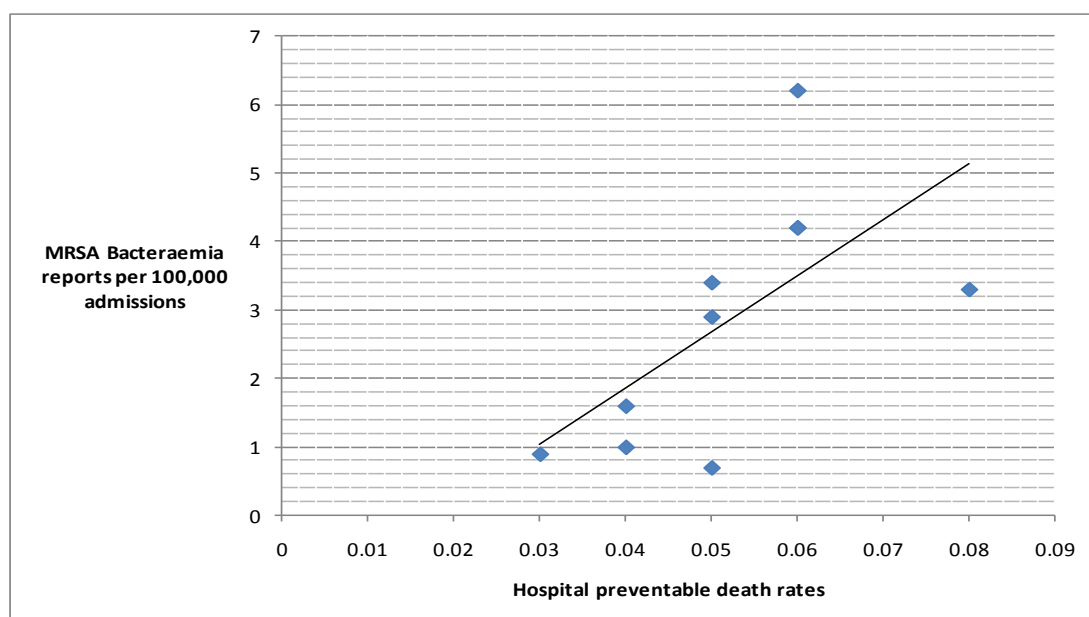
Discussion

Main findings

This exploratory study has found that only one of the eight measures of safety (MRSA bacteraemia rate) was significantly associated with the proportion of preventable deaths. In contrast, for four of the other measures, there appeared to be no association with preventable deaths (HSMR, emergency readmissions, hospital cleanliness, staff sickness absence).

Comparison with other studies

Due to the lack of previous studies on most of the measures of safety we considered, we can only compare our results for the association between preventable deaths and standardised mortality ratios. Our results are consistent with all the published evidence (except for pneumonia deaths in one study[8])[6-7 9] which also found no significant positive correlation.

One study from England of 173 acute hospitals considered associations between several patient safety measures but did not include preventable death proportion.[16] It found that a process measure (patient safety incident rates) had no association with several outcome measures (MRSA bacteraemia rates, SMRs, incidence of decubitus ulcers, post-operative sepsis rates). There were, however, positive associations with some structure measures (staff views of the safety culture in their hospital, risk management ratings).

Interpretation of findings

Our overall finding of no or only a weak association between preventable deaths and most other measures of safety suggests that each measure has different underlying causes. The one exception was a moderately strong correlation with a hospital acquired infection (MRSA) which gives credence to the importance given to the latter in many countries, including England.[17] In the study from which our data were drawn, 7% of preventable deaths were associated with hospital acquired infection and 3.8% with MRSA septicaemia.[18]

Our findings do provide support to policies in UK, USA and France aimed at reducing MRSA bacteraemia rates in order to reduce preventable deaths.[19] Although our study was insufficiently powered to detect a statistically significant association with one of the cornerstones of infection control, hand washing by healthcare staff,[20] the findings do suggest a link with hospital acquired infection-related deaths. Measuring this activity may be an effective way to assess and drive improvements in safety.[21]

The lack of correlation of preventable deaths with rates of patient safety incidents may reflect a true lack of association or incomplete reporting to the NPSA. In addition, there is some uncertainty as to whether a high rate of reported incidents reflects poor safety or the opposite, acting as an indication of a greater propensity to address safety issues. Staff

concerns about unfair blame or fear of litigation, particularly in organisations with a poor safety culture, generally discourage reporting.[22-23]

The lack of correlation with HSMR is consistent with findings from other studies.[24] After taking into account artefactual and structural factors, it remains unclear how much of the residual variation in HSMRs represents differences in safety between organisations.[4-5] Given that this study (and others) suggests that only 5% of deaths are preventable, the signal to noise ratio of preventability would preclude the ability of HSMRs to be valid measures of safe care.

Strengths and limitations

The principal strengths of this study are the methodological rigour of determining preventable deaths and the first time the relationships with a wide-range of other safety measures at the level of individual hospitals has been undertaken outside the USA.

However, with only ten hospitals, the principal limitation is its power to detect associations between safety measures. The small sample size may have led to the undue influence of outliers on the value of the correlation coefficients, though this was addressed by using the more conservative Spearman Rank Correlation test.

Whilst retrospective case record review provides a comprehensive picture of patient care, it is inevitably limited to what is written in the record and is vulnerable to the risk of hindsight bias.[25] Our method maximised the validity and reliability of the judgment of preventability and was fit for the purposes to which the data have been put in this analysis.

Our choice of safety measures was guided by the desire to look at safety from multiple perspectives and to focus on entire hospitals. Use of other, not yet publicly available

measures, such as in-hospital cardiac arrest rates currently being collected as part of a national clinical audit in England, would have strengthened the study.[26] Despite the selection of safety measures being limited to those with reasonable coverage and measurement properties, several of these are of uncertain accuracy (such as patient incident reports, patients' views, staff views).

Although our study of preventable deaths was conducted during calendar year 2009, the majority of safety measures used in the correlation analyses were from the financial year 2009/10. We have no reason to believe that this minor lack of concordance of data collection periods would have introduced any significant bias.

Implications

Given the lack of association between preventable death proportion and some widely used measures of safety (HSMRs, emergency readmissions), a larger study is needed to establish whether this reflects the limited power of this study or real relationships. In autumn 2013 there are plans to study an additional 24 hospitals in England which will enable this to be resolved.[27]

This study also underlines the need for governments and others responsible for health care systems to consider a portfolio of measures of safety when assessing a hospital, given the limited inter-relationships between the various options. This, in turn, reflects the diverse aspects of safety that each measure detects.

Acknowledgements

Hutchings for statistical advice, and the National Institute of Health Research, Research for Patient Benefit Programme for funding.

Funding

References

1. Hogan H, Olsen S, Scobie S, Chapman E, Sachs R, McKee M, et al. What can we learn about patient safety from information sources within an acute hospital: a step on the ladder of integrated risk management? *Qual Saf Health Care* 2008;17(3):209-15.

2. Donabedian A. Methods for deriving criteria for assessing the quality of medical care. *Med Care Rev* 1980;37(7):653-98.

3. Shojania KG, Forster AJ. Hospital mortality: when failure is not a good measure of success. *CMAJ* 2008;179(2):153-7.

4. Scott IA, Brand CA, Phelps GE, Barker AL, Cameron PA. Using hospital standardised mortality ratios to assess quality of care--proceed with extreme caution. *Med J Aust* 2011;194(12):645-8.

5. van Gestel YR, Lemmens VE, Lingsma HF, de Hingh IH, Rutten HJ, Coebergh JW. The hospital standardized mortality ratio fallacy: a narrative review. *Med Care* 2012;50(8):662-7.

6. Best WR, Cowper DC. The ratio of observed-to-expected mortality as a quality of care indicator in non-surgical VA patients. *Med Care* 1994;32(4):390-400.

7. Gibbs J, Clark K, Khuri S, Henderson W, Hur K, Daley J. Validating risk-adjusted surgical outcomes: chart review of process of care. *Int J Qual Health Care* 2001;13(3):187-96.

8. Dubois RW, Rogers WH, Moxley JH, 3rd, Draper D, Brook RH. Hospital inpatient mortality. Is it a predictor of quality? *N Engl J Med* 1987;317(26):1674-80.

9. Guru V, Tu JV, Etchells E, Anderson GM, Naylor CD, Novick RJ, et al. Relationship between preventability of death after coronary artery bypass graft surgery and all-cause risk-adjusted mortality rates. *Circulation* 2008;117(23):2969-76.

10. Hogan H, Healey F, Neale G, Thomson R, Vincent C, Black N. Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study. *BMJ Qual Saf* 2012;21(9):737-45.

11. Brennan T, Leape L, Laird N, Herbert L, Localio R, Lawthers A, et al. Incidence of adverse events and negligence in hospitalised patients. Results of the Harvard Medical Practice Study 1. *N Eng J Med* 1991;324:370-6.

12. Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001;322:517-9.

13. Hayward R, Hofer T. Estimating hospital deaths due to medical error- preventability is in the eye of the reviewer. *JAMA* 2001;286(4):415-20.

14. Zegers M, de Bruijne MC, Wagner C, Hoonhout LH, Waaijman R, Smits M, et al. Adverse events and potentially preventable deaths in Dutch hospitals: results of a retrospective patient record review study. *Qual Saf Health Care* 2009;18(4):297-302.

15. Brennan TA, Localio RJ, Laird NL. Reliability and validity of judgments concerning adverse events suffered by hospitalized patients. *Med Care* 1989;27(12):1148-58.

16. Hutchinson A, Young TA, Cooper KL, McIntosh A, Karnon JD, Scobie S, et al. Trends in healthcare incident reporting and relationship to safety and quality data in acute hospitals: results from the National Reporting and Learning System. *Qual Saf Health Care* 2009;18(1):5-10.

17. National Audit Office. The management and control of hospital acquired infection in acute Trusts in England. London: National Audit Office, 2008.

18. Hogan H, Healey F, Neale G, Thomson R, Vincent C, Black N. Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study. *BMJ Qual Saf* 2012.

19. Haustein T, Gastmeier P, Holmes A, Lucet JC, Shannon RP, Pittet D, et al. Use of benchmarking and public reporting for infection control in four high-income countries. *Lancet Infect Dis* 2011;11(6):471-81.

20. World Health Organisation/ Patient Safety Alliance. WHO guidelines on hand hygiene in healthcare: First global patient safety challenge: Clean care is safe care. Geneva: WHO, 2009.

21. Kirkland KB, Homa KA, Lasky RA, Ptak JA, Taylor EA, Splaine ME. Impact of a hospital-wide hand hygiene initiative on healthcare-associated infections: results of an interrupted time series. *BMJ Qual Saf* 2012;21(12):1019-26.

22. Firth-Cozens J. Barriers to incident reporting. *Qual Saf Health Care* 2002;11(1):7.

23. Stanhope C, Crowley-Murphey M, Vincent C, et al. An evaluation of adverse incident reporting. *J Eval Clin Pract* 1995;5:1-4.

24. Pitches DW, Mohammed MA, Lilford RJ. What is the empirical evidence that hospitals with higher-risk adjusted mortality rates provide poorer quality care? A systematic review of the literature. *BMC Health Serv Res* 2007;7:91.

25. Caplan R, Posner K, Cheney F. Effect of outcome on physicians' judgements of appropriateness of care. *JAMA* 1991;265:1957-60.

26. Nolan J, Gallegher L, Llyod-Scott L, Rowan K. National Cardiac Arrest Audit (NCAA). *JICS* 2009;10:313-15.

27. Keogh B. Review into the quality of care and treatment provided by 14 hospital trusts in England: overview report. London: Department of Health, 2013.

Chapter 8 Discussion

8.1 Introduction

The aim of this research was to identify a method most suited to the identification of preventable hospital deaths, and to apply it to determine the proportion of preventable deaths in acute NHS hospitals in England and their underlying causes and contributory factors. The extent of correlation between the proportion of preventable deaths across hospitals and other safety indicators was explored. This chapter places my research findings within the body of current knowledge and highlights my contribution to its expansion. I also reflect on the implications of this research for clinical practice, NHS policy and future research.

8.2 Key Findings

8.2.1 Objective 1 and 2: To describe the strengths and weaknesses of current measures of patient safety for identifying harm in hospitals and compare the scale and scope of hospital harm identified by different measures of patient safety

A number of methods for measuring harm exist. Some are more frequently used in the service setting (Routine Data, Incident Reporting, Global Trigger Tool, Morbidity and Mortality Meetings, Confidential Enquiries), others in research settings (Direct Observation, Case Control, Patient Reported Harm) and some in either setting (Claims / Inquest Records, RCRR, Prospective Surveillance). By undertaking a narrative literature review of harm measurement, and exploring the range of measures and their performance in a single acute hospital, I was able to determine that common measures differ in the scope and scale of

harm they identify, with relatively little overlap between sources. Each measure is subject to a number of biases that threaten validity and reliability when measuring severe harm including preventable death. Compared to other harm measures, both RCRR and prospective surveillance identify the broadest spectrum of harm including that at the severe end of the spectrum and due to multiple small omissions. Both methods also capture adequate contextual information to allow an assessment of the preventability of harm to be made. RCRR also enables the examination of deaths for preventability on a larger scale with fewer resources, and makes comparisons with previous epidemiological harm studies that traditionally use this approach, possible.

8.2.2 Objective 3: To determine the proportion of preventable deaths, causes, contributory factors, subpopulations affected, and years of life lost in acute hospitals in England

The main study in this thesis, PRISM, has provided for the first time in England, a robust estimate of the proportion of preventable hospital deaths. This estimate of 5.2% (95% CI 3.8% - 6.6%) equates to approximately 11,859 such deaths annually across the NHS in England. I found that preventable deaths were more common amongst surgical than medical patients, with problems in care for both groups occurring most frequently during ward care (44%). The majority of problems in care were related to clinical monitoring (31%), diagnostic errors (30%), or inadequate drug or fluid management (21%). Most preventable deaths occurred in elderly and frail patients with multiple co-morbidities; 60% were judged to have less than one year of life left to live on admission.

Content analysis of the case narratives of the preventable deaths using a novel approach based on root cause analysis sought to understand in more depth the nature of these deaths. On average, three healthcare problems and five contributory factors were found per patient, with around three quarters related to omissions. Problems were found to accumulate across

the patient's hospital stay, as demonstrated by five out of seven patients admitted with warfarin-induced bleeding who went on to have further management problems related to anticoagulation during hospitalisation. Failures of early assessment and diagnosis (39% of all assessment and diagnosis problems), poor clinical monitoring of intravenous fluids (24% of all monitoring problems) and warfarin (20% of all monitoring problems), and failures to give indicated drugs (particularly venous thromboembolism prophylaxis) (41% of all drug and fluid problems) were the most frequent problem subtypes. Factors related to clinical tasks, e.g. design or guidance, and education and training made up 50% of contributory factors. Across all contributory factors, 34% were related to individuals' clinical technical performance (lack of knowledge or skills, staff cognitive processes) and 34% to non-technical issues (communication or teamwork), continuing to emphasise the interplay between clinical and system factors in the generation of harm.

8.2.3 Objective 4: To determine whether the proportion of preventable death across hospitals correlates with other patient safety indicators

Correlation of the proportion of preventable deaths with eight other patient safety indicators, chosen as measures linked to the structure, process and outcomes of healthcare, found that only one, the MRSA bacteraemia rate, had a statistically significant correlation ($r=0.73$, $p=0.02$). This study was necessarily exploratory due to the limited number of data points (ten hospitals), but there was an indication that the strength of association decreased across some measures from outcome (MRSA bacteraemia 0.73), through to process (staff hand hygiene 0.51) and to structure (poor safety culture 0.26) where such an association would seem feasible. This observation would require testing a in a larger study. There was no significant correlation between preventable deaths and five other measures of safety: HSMR, emergency readmissions, hospital cleanliness, and staff sickness absence.

8.3 Setting Thesis Findings in Context

There are only two previous studies with a significant focus on measuring preventable death by RCRR. Hayward and Hofer found 6% of deaths to be preventable, in a review of 111 deaths from across seven Veterans Affairs hospitals in the US.[17] More recently, a review of 3,983 deaths in 25 Dutch hospitals in 2005 found that a preventable adverse event contributed to death in 4.1% of cases.[174] My finding of 5.2% preventable deaths is compatible with this previous work. The Dutch study also estimated that around half of dying patients, having experienced a preventable adverse event, had a potential life expectancy of less than one year.[174]

My estimate is much lower than the 60,000 to 255,000 episodes of healthcare-related serious disability or death in England suggested in 2000 by the Chief Medical Officer, these figures being derived from extrapolations of the findings of the HMPS, and the Utah and Colorado studies in the US.[9] The robustness of extrapolations from these studies is limited by their small sample sizes for deaths, and the fact that reviewers were not required to explicitly examine the relationship between an adverse event occurrence and the preventability of the death.

Despite some issues with methodological differences when making direct comparisons between my findings and the large RCRRs of the past, it is reassuring that these studies also found similar causes of serious harm, namely clinical monitoring, assessment and diagnosis, and drugs and fluids problems, with over half of this harm occurring during ward care.[170 173-174] Additionally, my finding that a relatively small number of preventable deaths are associated with technical problems (6.3%), compared to diagnosis and monitoring problems, is consistent with the findings from these previous RCRR studies. The HMPS found that the balance in frequency between the usually more common technical and less common diagnostic problems altered in patients suffering severe harm or death. These patients were

more likely to have experienced diagnostic mishaps than problems related to surgery or procedures.[170] This finding was mirrored in the QAHCS where surgical adverse events occurred in only 3% of patients who died having experienced an adverse event, and diagnostic, system and therapeutic errors accounted for 64% of harm in this group.[173] More recently, the Dutch RCRR reviewing 3,983 deaths found that twice as many adverse events were related to the diagnostic process in this group than in patients who were discharged alive (14.8% vs. 6.3%).[174]

My finding that omissions in care, such as the failure to monitor or give an indicated drug, are twice as common as problems related to acts of commission (giving the wrong drug or an error during a procedure), is consistent with findings from previous studies. In the QAHCS, omissions were found to be twice as common as errors of commission and the HMPS found that, not only were omissions more common, but they were more likely to be preventable.[170 173] Hutchinson and colleagues also identified the greater impact on harm generation of multiple omissions along the care pathway rather than single acts of commission.[171]

Comparison of the findings from the correlation study are limited by its exploratory nature and the lack of similar such studies. My finding that there is no significant association between the proportion of preventable deaths identified by case record review and HSMR, is consistent with three out of the four published studies in this area.[21-22 24] This lack of correlation with HSMR is not surprising, and it agrees with findings from other studies that report a weak and inconsistent relationship between HSMR and other measures of the safety of hospital care.[251] After taking into account artifactual and structural factors, it remains unclear how much of the residual variation in HSMR represents differences in safety between organisations.[53 252] Modelling studies suggest that the proportion of preventable death would need to be significantly higher than 5.2% to show any clear association between these measures.[234] One caveat must be that although I can rule out any medium to strong

association with HSMR, smaller associations will be masked by the low power of the study, due to the relatively small number of hospitals in the sample. The strongly positive correlation with MRSA is consistent with the challenges related to MRSA infections and associated deaths that the NHS was facing around 2009.[253] Only one other study from England has looked at associations between a range of patient safety indicators, but did not include preventable deaths identified through RCRR.[254] Hutchinson *et al* found no association between reported patient safety incident rates to the NPSA from 173 acute hospitals, and MRSA bacteraemia rates, HSMRs, incidence of decubitus ulcers or post-operative sepsis rates. There were, however, positive associations with structure measures (staff views of the safety culture in their hospital, risk management ratings).

8.4 Strengths and Limitations of the Research

8.4.1 Strengths

Several methodological approaches were considered before choosing RCRR. A prospective cohort study would have allowed determination of incidence and associated risk factors, but given that preventable death is a rare outcome, such a cohort would have needed to be unfeasibly large [17 255] A case control study would have doubled the cost and time for data collection and given that the focus of the study was not on identifying risk factors for preventable deaths, but rather on estimating the proportion of preventable deaths by close examination of the appropriateness of processes of care, the extra investment was not justifiable. RCRR compares favourably to other harm methods in the identification of harm at the severe end of the spectrum.[256-257]

The research design included a number of measures to enhance validity and reliability. The external validity was strengthened by the large random sample size, which was drawn from

acute hospitals in different regions and of different sizes and teaching status. It was the largest sample of deaths to be reviewed in England and second only to the Netherlands, internationally.[174] I was also able to confirm that the study sample was representative of the population of patients who die in English hospitals each year. Efforts were made to ensure that the participating hospitals undertook correct sampling procedures and worked hard to trace any missing records, particularly if these records were part of a medicolegal case. For those records that could not be traced, additional information on the age, sex, specialty and reason for loss were sought, in order to check for selection bias. In addition, reviewers were asked to rate the completeness of the case records and identify any where the amount of missing information prevented a review. Only 0.8% of records were rejected on this basis.

Using retired consultant physicians as reviewers was another study strength. They had the experience to make sense of complex clinical scenarios, and to take the necessary global overview required when making judgements of the preventability of deaths. An additional advantage of using retired doctors was that many had practised medicine at a time when there was less specialisation and were comfortable assessing the care of patients spanning a range of specialties. The study design could be criticised for not using specialist reviewers. However, findings from previous RCRR had highlighted that much of the harm we might find in PRISM was likely to be of a general nature related to diagnosis, assessment or monitoring, and would not require specialist reviewers. [15 37 169-170 214] Reviewers did have access to specialty expertise either from other members of the review group and via external sources if necessary.

The study used of a new definition of harm, 'problem(s) in care', rather than the more commonly used 'adverse event'. This change was designed to ensure omissions and multiple problems in care would be captured. Reviewers were asked not only to identify problems in care that had contributed to a death but also to answer a separate question on the preventability of the death using a six point Likert scale. This judgement step had previously

only been included in only one other RCRR study,[17] but was essential in enabling reviewers to take into account not only the appropriateness of care the patient received but also the impact of the patient's condition on admission, including co-morbidities, on the likelihood of death.

A key measure to improve reliability was training. The design of training programme was based on that used for three previous RCRR studies (HMPS, QAHCS, London). Reviewers were able to practise using the structured Medical Review form on real cases during the training, giving them the opportunity to ask questions and compare their performance with peers. A source of bias that is hard to eradicate from retrospective studies is hindsight bias. Reviewers were instructed on approaches to the review that helped to minimise this bias. Each reviewer was also given an instruction manual to remind them of key issues to look out for when completing the structured review form. Once reviews commenced, expert reviewers were available to answer reviewers' questions, and a group email facility allowed further discussions.

The qualitative analysis developed a novel content analysis approach based on root cause analysis tools in current use. The method enabled extraction of richer material on problems in care, and measurement of inter-rater reliability indicated a good level of agreement for problems in care, and moderate agreement for contributory factors.

8.4.2 Limitations

RCRR has been shown to more often identify harm related to the clinical aspects of care at the clinician-patient interface, especially if related to human error, than harm that results from underlying system problems.[37 120 218]   In my study, I found only 11.5% of the problems identified could be considered mainly healthcare system problems. This may be partly

because it is not always possible to discern contributory factors such as teamwork, leadership or communication issues because of lack of detail in the case record.

 Important information may be missing due to illegibility, poor recording practices or misfiling. In the QAHCS, over a quarter of the records did not have enough content to allow a judgement of whether harm had occurred.[169] However, despite 12% of records in this study having missing components, only 0.8% were rejected by the reviewers on the grounds of inadequate content.[258] The problem of missing content may lead to information bias. It is possible in hospitals with more rigorous record keeping standards and well maintained records, that more harm will be identified than in those with poor standards where crucial details may be missing. I guarded against this by the inclusion of an item in the review form that rated the quality of the records for each case, and was able to ensure that there were no serious outliers for quality of record keeping.

RCRR studies are often criticised because of the poor reliability of the reviewers' judgments. Most studies rarely achieve more than moderate reliability for judgements of harm and preventability.[211] The nature and complexity of judgements, given that the majority of patients under review are elderly and frail with multiple co-morbidities, limits the degree of reliability that can be achieved. Some researchers have advocated using more than one reviewer for each case, but it has been shown that improvement in reliability only occurs when five or more clinicians undertake the review.[207-208] That number of reviewers is beyond the means of most large scale studies. Hayward and Hofer, who have done much of the empirical research on inter-rater reliability in the judgement of harm, have concluded that having one or two reviewers per case record is permissible for studies comparing hospitals where there is a large number of cases being reviewed but would not be satisfactory for individual patient level decisions.[209] As the Dutch had recently found no worthwhile improvement in reliability using two reviewers,[210] I settled for one reviewer per case record, but with two reviewers visiting each site and a quarter of all reviewed records

undergoing double review. In addition, to account for error around judgements of preventability using the Likert scale, a sensitivity analysis was incorporated into the final analysis using different Likert scale thresholds (3 and 5) to give an indication of the potential range for the proportion of preventable deaths.

A further limitation is hindsight bias, in which knowing the outcome and its severity influences judgements of harm causation and preventability.[211 216] I did consider the feasibility of using copies of the admission record with masked outcomes in addition to reviewer training on the issue, but the resources required were beyond the means of this research. There was also a risk that altering the records in this way would compromise the ability of reviewers to make judgements.

Experienced generalist reviewers rather than specialist reviewers were used, the majority of whom were physicians rather than surgeons. I thus ran the risk of missing harm related to technical aspects of surgical care. This could have led to an under-estimation of the number of surgical preventable deaths. However, I did have the support of two general surgeons who were on hand to deal with any questions related to surgery raised by the reviewers. In the end, the study found a higher proportion of problems in care and preventable deaths amongst surgical patients than medical patients, but the majority of these were related to ward care rather than the surgical intervention itself.

Study findings that patients were more likely to experience harm if they were less functionally impaired, were elective admissions and had a longer life expectancy on admission were inconsistent with previous studies and could have reflected a bias amongst reviewers towards discounting problems in the most frail, sick patients. To reduce this tendency, reviewers were required to examine the entire record to the same depth and in the same structured way for all patients, regardless of their condition on admission. An alternative explanation might reflect the greater willingness in England, compared to the US,

to limit the extent of medical intervention in frail patients. This, in turn, would result in less exposure to potential high risk procedures or treatments, and reduce the opportunity for harm.

The estimates of life expectancy were dependent on reviewers' judgements, a notoriously difficult task especially when many patients were very elderly and frail. However, no alternative approach proved satisfactory given the nature of the patients in the sample.

The robustness of the correlation between the proportion of preventable death and other safety indicators was limited in its power to detect associations, because of the small number of hospitals in this study. This analysis should be regarded as exploratory and will benefit from testing in a larger study. However, while the lack of statistically significant correlations with some measures cannot rule out the existence of weak associations with preventable mortality, this study was still able to confirm the absence of strong associations. It cannot be ruled out that the lack of correlation with some measures may have been a reflection of the poor quality of the data.

The small sample size may have led to the undue influence of outliers on the value of the correlation coefficients, but this was addressed by using the more conservative Spearman Rank Correlation test. The choice of safety measures was guided by the desire to look at safety from multiple perspectives and to focus on entire hospitals. Use of other, not yet publicly available measures, such as in-hospital cardiac arrest rates being collected as part of a national clinical audit, would have strengthened this element of the study.[259] Furthermore, the proportion study was conducted during the calendar year 2009, and the majority of safety measures used in the correlation analyses were from the financial year 2009/10. It is unlikely that this lack of concordance of data collection periods would have introduced significant bias.

The content analysis was also weakened by the fact that it depended on the level of detail recorded by reviewers in the case narrative section of the review form, which varied in length from a single paragraph to several sides of A4. This was a more noticeable handicap for identifying contributory factors.

8.5 Implications for the NHS

8.5.1 Implications for Clinical Safety in the NHS

The revelation of the scale of the poor healthcare at the Mid Staffordshire NHS Foundation Trust has ensured that hospital-related harm has moved up the political agenda with reports of such harm rarely out of the media.[27] Today, reduction of harm, particularly serious harm, is a key focus for improvement in the NHS. My research highlights that whilst three-quarters of patients receive good quality of care, there is scope for improvement in the safety of healthcare provision. The findings show that well known and long standing causes of serious patient harm, such as inadequate assessment, poor fluid balance, failure to monitor warfarin or give thromboembolism prophylaxis, persist in the NHS.[123 260-263]

My analysis of contributory factors implicates both individuals' clinical technical performance (lack of knowledge or skills, staff cognitive processes) and non-technical factors (communication or teamwork). Case narratives capture the fact that individual human error at the patient-clinician interface occurs within a system in which poor communication, poor teamwork and inadequate levels of supervision are common and inexperienced clinicians continue to be exposed to complex clinical situations without the adequate knowledge and skills to deal with them. This situation very much reflects James Reason's 'Swiss Cheese' model of failed defences, demonstrating how risks to patient

safety, inherent in most healthcare systems, will reach the patient when all the holes in the cheese line up.[59]

Figure 8.1 James Reason's Swiss cheese model of harm generation
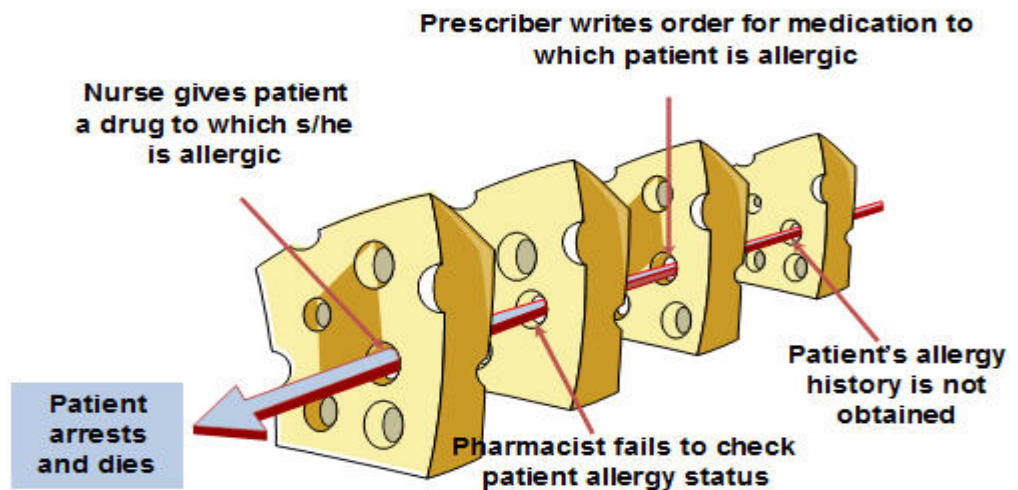
In the last decade a wide range of interventions have been introduced in the NHS to address the known underlying risks that lead to serious harm, including Early Warning Score Systems to avoid delay in identifying deteriorating patients,[264] explicit handover procedures to ensure vital clinical information is passed between clinicians,[265] and critical care outreach services to rescue deteriorating patients.[266] However, my findings that over 70% of problems in care are due to omissions would suggest that one focus for improvement should be on improving the reliability of the delivery of care overall. Although initiatives such as national audits have led to improvements in standardisation of care delivery within certain specified areas, a move beyond a patchwork approach to a 'whole system' approach would have a greater impact. Approaches aimed at achieving total system reliability, such as those based on Six Sigma principles (a process improvement strategy to identify and remedy defects and

variability) are gradually being introduced across some NHS wards, and evidence is emerging that there are improving outcomes.[267-268] Further research is needed to understand the key components of these programmes linked to success.[269] Improved reliability will not eradicate all human error. Training and supervision remain key defences to human failures. The link between preventable death and inadequate senior supervision and support for junior staff is becoming increasingly recognised, both through studies that show poor outcomes from out of hours care and via recent investigations of failing hospitals.[270-271]

With an aging population, nearly a quarter of all NHS hospital admissions are aged over 75 years.[272] By 2025 it is estimated that the over 65s will account for 60% of acute hospital admissions.[273] Older people who come into contact with increasingly complex and technological healthcare provision are vulnerable to higher rates of harm due to frailty and multiple co-morbidities, and are likely to suffer more serious consequences.[168] Their longer lengths of stay in hospital compared to younger patients', increase the opportunity for injury.[274] My findings suggest hospital provision still appears ill equipped to keep this vulnerable population safe. Several recent reports have highlighted failings in the standards of care for older people within hospitals, due to poorly co-ordinated care.[27 275] Within a system where care is fragmented, the opportunity for harm as a result of omissions is high. A recent King's Fund report corroborated my findings when highlighting how the breakdown of team structure as a consequence of shift work, the lack of opportunity for regular multi-professional team meetings to plan care, larger numbers of specialists involved across the care journey, and inadequate senior clinical supervision of junior staff are all contributing to fragmentation of care delivery.[273]

A consensus is beginning to emerge as to how acute care can be provided in hospitals to ensure both a better experience for patients, and protection from healthcare-related harm. Recent publications by the Royal Colleges of Physicians[276-277] have put forward a number of recommendations to this end, including:

- *Early comprehensive multi-disciplinary assessment*

- *Early establishment of treatment goals*

- *Consistent consultant cover and early consultant review*

- *Integrated services, with access to care of the elderly expertise*

- *Assessment, documentation and treatment of acute illness standardised across the NHS*

- *Standard handover procedures*

- *Education and training in order to better address the specific needs of an aging population*

- *Avoidance of unnecessary transfers*

- *Prompt consultant review after transfer*

In addition, the widespread adoption of approaches shown to improve outcomes in the elderly, such as regular multidisciplinary team meetings,[278] need to be combined with general improvements to the quality of healthcare provision through strengthened ward leadership, regular senior review of patient progress, electronic prescribing systems and clear pathways for communication to ensure appropriate care is delivered as intended and any negative impacts identified and dealt with rapidly. Effective approaches aimed at improving non-technical aspects of care, such as teamwork and communication, exist and are currently being tested in healthcare settings. A recent study showed team-based training programmes previously used to train airline cockpit staff, can lead to a reduction in error rates in surgical settings.[279]

8.5.2 Implications for Patient Safety Measurement

Recent media stories linking raised HSMR / SHMI to thousands of preventable deaths across NHS acute hospitals have caused alarm amongst the public and politicians alike.[280] While

the spectre of preventable hospital deaths may prove helpful in raising interest in patient safety and a commitment to improvement, over-estimating the size of the problem and the risk to patients may induce unjustified levels of anxiety and fear among the public, as well as dismay and despondency amongst the staff that care for them in resource stretched environments. Preventable deaths form a relatively small proportion of all deaths, and with their low signal to noise ratio, it is very unlikely that variation in HSMR and SHMI between hospitals can ever reflect variation in actual preventable deaths. The findings from the exploratory correlation study support the fact that any association is not likely to be strong. With serious concerns being voiced as to the validity of using HSMR and SHMI to identify excess mortality in acute hospitals, my study presents a feasible alternative measure.[51][281]

The two public enquiries undertaken by Robert Francis QC,[27][282] on the failings in standards of care at the Mid Staffordshire NHS Foundation Trust revealed a raft of failings at the Trust resulting from the lack of commitment to quality and safety at Board level, a culture that tolerated low standards of care compounded by understaffing and fears among some staff of raising awareness of these problems. He saw the situation compounded by the emphasis placed on the importance of financial balance by central supervisory NHS bodies and the failure of the regulator, the Care Quality Commission, to undertake sufficient scrutiny. Frances' recommendations called for the development and enforcement of fundamental standards which place the patient at their heart combined with a duty to be open and transparent when things go wrong, an emphasis on improving the culture amongst nursing staff, in particular, the quality of communication with patients and improvements in the measurement and dissemination of performance related information.

In the government's response,[283-284] Trusts would now be required to publish data on complaints and ward-based staffing and there was a proposal for new legislation related to candour and the criminalisation of patient neglect. The government set up a number of reviews to look at areas ranging from patient safety to the burden of NHS information

collection and monitoring on Trusts. One of these reviews undertaken in 2013 was led by Sir

Bruce Keogh, the Medical Director of the NHS, and investigated 14 NHS acute Trusts that

had persistently high mortality rates.[271] Although the review found evidence of differing

issues impacting on quality and safety at each Trust, common themes to emerge were the

failure of management to engage and support frontline staff, professional isolation and the

limited capacity within Trusts to utilise measurement for quality improvement. Sir Bruce

also noted that the different measures of hospital-wide mortality used to identify the 14

hospitals had generated two completely different lists of outliers, concluding that reliance on

these measures to identify failing hospitals was likely to lead to misleading results. He called

for an examination of the relationship between the measures and preventable deaths

identified via case record review. My colleagues and I have been commissioned to undertake

this study and will do so by extending the PRISM study to 24 more sites, providing a

statistically robust sample size to enable correlation with HSMR/SHMI measures. The

review became the blueprint for a new approach to hospital inspection for the Care Quality

Commission which takes account of a far wider range of intelligence, with the views of

patients and staff at all levels being seen as key.

During the period that the post-Francis reviews have been taking place, the Department of

Health and NHS England have been developing a new national indicator, 'hospital deaths

attributable to problems in care', based on my case record review methodology. This

indicator will be incorporated into the NHS Outcomes Framework and will be used to both

track the national proportion of preventable deaths over time, and encourage a systematic

approach to mortality review across NHS hospitals.[285] It is planned that hospitals will use a

national standard form to review random sample of records of deaths each year, with a

proportion of these records selected for external review. The extension of the PRISM study

to a further 24 hospitals will provide an opportunity to pilot the new form and establish a

baseline proportion of preventable deaths for 2012/13. Whether, NHS England, will want to

replace HSMR/SHMI measures with a mortality measure derived from case record review for benchmarking hospitals remains to be seen

Extension of my method beyond its research origins will raise a number of important issues for consideration by policy makers. These include:

1. *Scope of harm indicator*

    As death occurs relatively rarely in hospital (in approximately 2% of inpatients[64]) and preventable death is even rarer, it should be acknowledged that the greater burden of harm, even at the severe end, falls on those who are discharged from hospital alive. The profile of this harm may well be different from that found in patients who die. Moreover, some specialties such as ophthalmology will have relatively few deaths, which limits the utility of mortality review to improve safety in these areas. No single measure of harm is likely to capture the full picture, and a combination of approaches would seem the optimum way of finding the broadest range of harm.

    Many aspects of the poor care found at Mid Staffordshire and at other NHS acute Trusts since then are clearly related to organisational culture. These problems are likely to be compounded by the tough financial constraints that most NHS organisations are facing today. The NHS budget has been frozen in real terms since 2010/11 leading to staff reductions and pay restraints placed on many others.[286-287] The ability of case record review to be able to reveal links between these underlying factors and failures in the processes of care at the clinician/patient interface is limited by the fact that information on contributory factors is often poorly recorded. The first Francis investigation included a case record review of 60 deaths at Mid Staffordshire and revealed only one death thought to be avoidable.[27] Clearly this approach would not be an adequate safeguard of quality and safety on its own.

2. *Training for internal and external reviewers to increase reliability of judgements*

   Training is key to improving the reliability of the measurement of preventable

   deaths by case record review. Providing timely training to staff across English

   hospitals will be challenging, and consideration of the use of technologies such as

   such as *elearning* will be necessary to ensure the widest audience reach.

3. *Resource intensiveness*

   The average review time for my study was one hour. Death reviews will place a

   considerable resource burden on acute hospitals. Consideration of how to avoid

   over-burdening reviewers with unnecessary review questions, whilst at the same

   time not losing essential information will be important. The possibility of

   developing a screening tool, similar to GTT, which can be used to target reviews on

   high risk cases should be explored. Senior consultant time is an expensive resource

   within the NHS, and consideration might need to be given to substitution with

   middle grade doctors. In addition, nurses can bring a different perspective to case

   record review and they should be included in the process. To maximise the

   relevance of the reviews to local learning, they should be conducted as soon after the

   patient's death as possible. This will require administrative support to ensure the

   records for review are rapidly identified and forwarded to the reviewers.

4. *Local Learning*

   A key outcome for local review will be the learning that arises as a result of the

   findings. The new national indicator should include a way of tracking local

   improvements that are implemented following reviews. In addition, the national

   programme should be flexible enough to enable hospitals to collect additional

   information on key quality and safety areas of interest to them, such as the quality of

   end of life care.

5. *Hospital Comparisons*

There will always be a need to use measurement for national comparisons of hospital performance. However, two main barriers will limit the validity of using mortality reviews for this purpose. Firstly, sample sizes at each organisation would need to be large enough to rule out the impact of chance variation, given that only around 1 in 20 deaths are likely to be preventable. Secondly, the moderate reliability of the measure, whilst acceptable when used for internal monitoring purposes or when large numbers of cases are reviewed to establish the national indicator, would limit use for hospital-level comparisons.

Although measurement of preventable deaths and their underlying causes will provide a rich source of information that can lead to safety improvement, my conceptual framework indicates that other harm measures will shed light on different aspects of harm. In addition, there are a wide number of other safety measures apart from harm, such as safety culture or frequency of staff hand washing, which can provide a broader picture of safety related to the structures and processes of an organisation, and not to just the outcomes. A broader approach to safety measurement can give a clearer indication of organisational risk and where interventions might be targeted.

Harm measurement provides information on how safe an organisation was in the past, reflecting patient safety at the time of measurement. Although it can be helpful for identifying key risks in healthcare, and has previously been instrumental in raising the consciousness of the public, clinicians and policy makers and galvanising action, it is perhaps best seen as only part of the mechanism for safeguarding patient safety. Recently Professor Charles Vincent and colleagues have introduced a new framework for patient safety measurement (Figure 3).[288] Drawing on approaches to safety from other high risk industries such as aviation and the nuclear industry, and the fact that no single approach to

measurement can capture all facets of safety, the new framework attempts to capture its most
important dimensions. These dimensions are listed below and their interactions highlighted
in the following figure:

- *Measures of harm*

- *Measures of reliability*

- *Measures of day to day operational safety*

- *Measures of preparedness*

- *Measures of integration and learning*

Figure 8.2 A Framework for safety measurement and monitoring
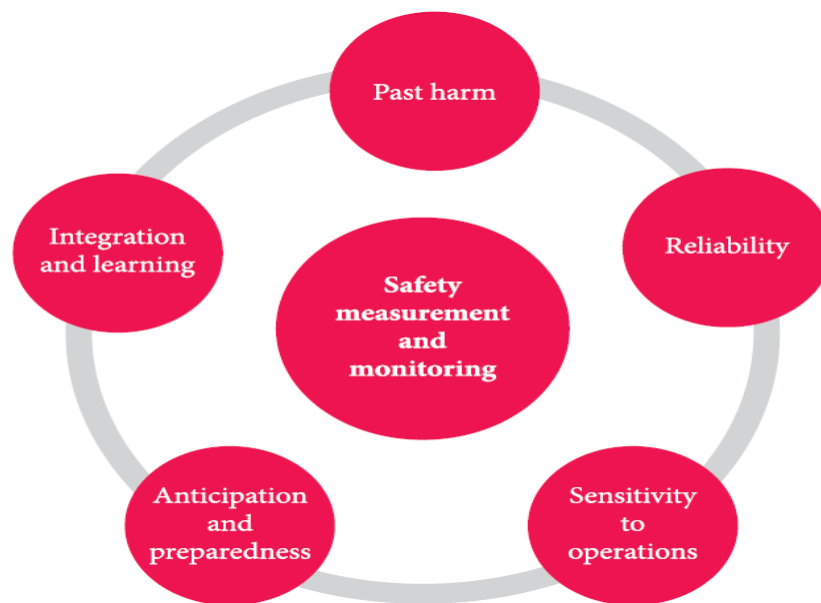
The authors indicated that this is an aspirational model, and that few NHS organisations

would currently be in a position to deliver measurements for each of its dimensions.

However, it would seem to address some of the disadvantages of harm measurement, in that

measures of day to day safety and potential risks bring safety measurement from the past

into the present. A focus on measures of reliability will help tackle the burden of harm from omissions in care. Integration of information and learning is vital to organisational quality and safety improvement.

Even though it is recognised that clinician involvement in improving the safety of the environment in which care is delivered is paramount, this has not been altogether successful to date. Donald Berwick, a physician who has spearheaded the championing of quality and safety improvement in US healthcare organisations, addressing the fact that relatively little progress had been made in tackling the common underlying problems that lead to serious patient harm, wrote that this was in part due to the fact that 'doctors fail to see the problem'.[289] In a recent article in the BMJ, Ian P Leistikow from the Dutch Healthcare Expectorate, echoed Berwick, identifying challenges to implementing patient safety initiatives which included invisibility, ambiguity and clinician autonomy.[290]

8.6 Areas for Future Research

A key question that remains to be addressed is the degree of association between HSMR / SHMI and the proportion of preventable deaths found from case record reviews. To date only a small number of published studies from the US,[21-24] and a modelling study from the UK,[234] have been able to shed light on the likely relationship, which does not appear to be strong. I did explore this relationship in my correlations between proportions of preventable deaths across the sampled hospitals and a number of safety indicators including HSMR. Again, no strong association was found. However, the small number of hospitals in the sample limited the robustness of this finding. I have recently been the recipient of a further grant from the Department of Health to extend the PRISM study to a further 24 acute hospitals, which will provide an adequate sample size to detect any statistical association. Findings from this study, which was the first recommendation of a recent review of 14 acute

hospitals with higher than expected HSMR or SHMI, undertaken by the Medical Director of the NHS,[271] will be available at the end of 2014.

As mortality review based on RCRR is likely to remain a key aspect of harm measurement, future research should focus on methods that might improve the efficiency of such reviews, be these screening tools to enable deaths at high risk of preventability to be more reliably identified or the development methods of communication of findings from death reviews that can effectively drive organisational learning. In addition, more needs to be understood about the factors that lead to the persistence of common causes of preventable death, such as missed diagnosis or poor clinical monitoring. Although it is unlikely that many more large-scale, broad-based RCRR studies will be commissioned, given the consensus on the nature of hospital harm that has emerged from these studies, the need for in-depth exploration of specific harms and their underlying causes will persist. National Confidential Enquiries and specialty specific investigations of high risk areas can make valuable contributions, but careful consideration of the methodology around such investigations is required to ensure that the findings are robust and generalisable enough to justify service change.

Little is known about the scale and scope of harm, including preventable death, in primary care. Although typical RCRRs may be difficult to conduct in this setting given the different structure of General Practitioner records and the need to capture information on care provided by the wider primary care team, suitable adaptations should make this possible.

8.7 Conclusion

The findings from my research have delivered the first accurate national estimate for the proportion of preventable deaths in English hospitals, based on a large random sample of deaths from within the NHS. This finding will be used as a national baseline against which future national estimates (derived from the new national indicator) will be compared. My work has also made an important contribution to the development of the approach to

mortality review, which has applications to both service and research settings. Its adaptation to become a new national indicator will see the introduction of mortality review based on this methodology across all NHS hospitals. For the first time, I have provided a robust estimate that can be used in international comparisons and it is somewhat reassuring to see that our proportion compares favourably with the estimation from the Netherlands. Much work remains to be done to reduce preventable deaths and serious patient harm, and it is somewhat discouraging that many of the underlying causes of the harm found in my study have been well known for many years. Given this fact, it is unlikely that progress will be easy. However, in the face of recent scandals in the NHS, hospitals managers and clinicians appear to be enthusiastic to meet this challenge. This thesis will form the basis of one tool that will support them in this endeavour.

References

1. Health and Social Care Information Centre. Key facts. Monthly HES data for admitted patient care, 2013.

2. Ipsos MORI Social Research Institute. Public perceptions of the NHS and social care survey. An ongoing tracking study conducted for the Department of Health, December 2011 Wave. London: Ipsos MORI/ Department of Health, 2012.

3. Schimmel EM. The Hazards of Hospitalization. *Ann Intern Med* 1964;60:100-10.

4. Barr DP. Hazards of modern diagnosis and therapy: the price we pay. *J Am Med Assoc* 1955;159(15):1452-6.

5. Mills D. Medical insurance feasibility study. A technical study. *West J Med* 1978;128:360-5.

6. Brennan T, Leape L, Laird N, Herbert L, Localio R, Lawthers A, et al. Incidence of adverse events and negligence in hospitalised patients. Results of the Harvard Medical Practice Study 1. *N Eng J Med* 1991;324:370-6.

7. Wilson RM, Michel P, Olsen S, Gibberd RW, Vincent C, El-Assady R, et al. Patient safety in developing countries: retrospective estimation of scale and nature of harm to patients in hospital. *BMJ* 2012;344:e832.

8. Institute of Medicine. To err is human: Building a safer healthcare system. Washington DC: National Academy Press, 1999.

9. Department of Health. An organisation with a memory: Report of an expert group on learning from adverse events in the NHS. Chaired by the Chief Medical Officer. London: Department of Health, 2000.

10. Department of Health. Learning from Bristol: the report of the public enquiry into children's heart surgery at the Bristol Royal Infirmary 1984-1995. London: HMSO, 2001.

11. Healthcare Commission. Investigation into outbreaks of Clostridium difficile at Stoke Mandeville Hospital, Buckinghamshire Hospitals NHS Trust. London: Healthcare Commission, 2006.

12. Thomas E, Studdert D, Bernstin H, Orav J, Zeena T, Williams E, et al. Incidents and types of adverse event and negligent care in Utah and Colorado. *Med Care* 2000;38:261-71.

13. Aylin P, Tanna S, Bottle A, Jarman B. How often are adverse events reported in English hospital statistics? *BMJ* 2004;329(7462):369.

14. National Audit Office. A safer place for patients: learning to improve patient safety. London: National Audit Office, 2005.

15. Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001;322:517-9.

16. Zegers M, de Bruijne C, Wagner C, Groenewegen P, Waaijman R, van der Wal G. Design of a retrospective patient record study on the occurrence of adverse events among patients in Dutch hospitals. *BMC Health Serv Res* 2007;7:27.

17. Hayward R, Hofer T. Estimating hospital deaths due to medical error- Preventability is in the eye of the reviewer. *JAMA* 2001;286(4):415-20.

18. National Reporting and Learning System. Quarterly National Reporting and Learning System data summary. London: National Patient Safety Agency, 2006.

19. Patient Safety Observatory. Safer care for the acutely ill patients: learning from serious incidents. The fifth report from the Patient Safety Observatory. London: National Patient Safety Agency, 2007.

20. Giles S, Fletcher M, Baker M, Thomson R, editors. *Incident Reporting and Analysis*. Buckingham: Open University, 2005.

21. Best WR, Cowper DC. The ratio of observed-to-expected mortality as a quality of care indicator in non-surgical VA patients. *Med Care* 1994;32(4):390-400.

22. Gibbs J, Clark K, Khuri S, Henderson W, Hur K, Daley J. Validating risk-adjusted surgical outcomes: chart review of process of care. *Int J Qual Health Care* 2001;13(3):187-96.

23. Dubois RW, Rogers WH, Moxley JH, 3rd, Draper D, Brook RH. Hospital inpatient mortality. Is it a predictor of quality? *N Engl J Med* 1987;317(26):1674-80.

24. Guru V, Tu JV, Etchells E, Anderson GM, Naylor CD, Novick RJ, et al. Relationship between preventability of death after coronary artery bypass graft surgery and all-cause risk-adjusted mortality rates. *Circulation* 2008;117(23):2969-76.

25. Committee of Public Accounts. A safer place for patients: learning to improve patient safety. Fifty-first report of session 2005-06, together with formal minutes, oral and written evidence. London: UK Parliament: House of Commons Committee of Public Accounts, 2006:p.5.

26. House of Commons Health Committee. Sixth report on patient safety In: 2008-2009 HOCS, editor. London: UK Parliment: House of Cmmons Health Committee, 2009.

27. Francis R. The Mid Staffordshire NHS Foundation Trust Inquiry. Independent inquiry into care provided by Mid Staffordshire NHS Foundation Trust. January 2005-March 2009. Volume 1: Stationary Office (HC375-1), 2010.

28. Darzi A. High quality care for all. NHS next stage review final report. London: Department of Health, 2008.

29. Soanes C, Hawker S, Elliott J. *Paperback Oxford English dictionary, Volume 10*. Oxford: Oxford University Press, 2006.

30. World Health Organization. World Alliance on Patient Safety: Conceptual Framework for international classification for patient safety, Version 1.1. Final technical report. Geneva: World Health Organisation, 2009:p.15.

31. National Patient Safety Agency. Seven steps to patient to patient safety. London: National Patient Safety Agency, 2004.

32. Donabedian A. Methods for deriving criteria for assessing the quality of medical care. *Med Care Rev* 1980;37(7):653-98.

33. Donabedian A. *Explorations in quality assessment. Volume 1 The definition of quality and approaches to its assessment*. Ann Arbour, MI: Health Administration Press, 1980.

34. 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(3):158-62.

35. Mardon RE, Khanna K, Sorra J, Dyer N, Famolaro T. Exploring relationships between hospital patient safety culture and adverse events. *J Patient Saf* 2010;6(4):226-32.

36. Reason J. *Managing the risks of organisational accidents*. Aldershot: Ashgate, 1997.

37. Thomas EJ, Petersen LA. Measuring errors and adverse events in health care. *J Gen Intern Med* 2003;18(1):61-7.

38. Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. *JAMA* 2002;288(16):1987-93.

39. Landrigan CP, Rothschild JM, Cronin JW, Kaushal R, Burdick E, Katz JT, et al. Effect of reducing interns' work hours on serious medical errors in intensive care units. *N Engl J Med* 2004;351(18):1838-48.

40. Groene O, Mora N, Thompson A, Saez M, Casas M, Sunol R. Is the maturity of hospitals' quality improvement systems associated with measures of quality and patient safety? *BMC Health Serv Res* 2011;11:344.

41. Gaba DM. Anaesthesiology as a model for patient safety in health care. *BMJ* 2000;320(7237):785-8.

42. ACSNI Study Group on Human Factors/ UK Health and Safety Commission. Third Report: organizing for safety. *HSMO*. London, 1993.

43. Singer SJ, Falwell A, Gaba DM, Meterko M, Rosen A, Hartmann CW, et al. Identifying organizational cultures that promote patient safety. *Health Care Manage Rev* 2009;34(4):300-11.

44. Hofman D, Mark B. An investigation between safety climate and medication errors and other nurse and patient outcomes. *Personnel Psychology* 2006;59:847-69.

45. Pronovost PJ, Weast B, Holzmueller CG, Rosenstein BJ, Kidwell RP, Haller KB, et al. Evaluation of the culture of safety: survey of clinicians and managers in an academic medical center. *Qual Saf Health Care* 2003;12(6):405-10.

46. Sherwood G, Thomas E, Bennett DS, Lewis P. A teamwork model to promote patient safety in critical care. *Crit Care Nurs Clin North Am* 2002;14(4):333-40.

47. Catchpole K, Mishra A, Handa A, McCulloch P. Teamwork and error in the operating room: analysis of skills and roles. *Ann Surg* 2008;247(4):699-706.

48. de Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care* 2008;17(3):216-23.

49. Burnett S, Franklin BD, Moorthy K, Cooke MW, Vincent C. How reliable are clinical systems in the UK NHS? A study of seven NHS organisations. *BMJ Qual Saf* 2012;21(6):466-72.

50. Tucker A, Edmondso A. Why hospitals don't learn from failures? Organisational and psychological dynamics that inhibit a system change. *California Management Review* 2003;45(2):1-18.

51. Lilford R, Pronovost P. Using hospital mortality rates to judge hospital performance: a bad idea that just won't go away. *BMJ* 2010;340:c2016.

52. Penfold R, Deans S, Flemons W, Mofat M. Do hospital standardized mortality ratios measure patient safety? HSMRs in the Winnepeg Regional Health Authority. *Healthcare Papers* 2008;8:8-24.

53. van Gestel YR, Lemmens VE, Lingsma HF, de Hingh IH, Rutten HJ, Coebergh JW. The hospital standardized mortality ratio fallacy: a narrative review. *Med Care* 2012;50(8):662-7.

54. 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.

55. Andrews LB, Stocking C, Krizek T, Gottlieb L, Krizek C, Vargish T, et al. An alternative strategy for studying adverse events in medical care. *Lancet* 1997;349(9048):309-13.

56. Weinger MB, Pantiskas C, Wiklund ME, Carstensen P. Incorporating human factors into the design of medical devices. *JAMA* 1998;280(17):1484.

57. Shojania KG, Forster AJ. Hospital mortality: when failure is not a good measure of success. *CMAJ* 2008;179(2):153-7.

58. Hollnagel E, Kaarstad M, Lee HC. Error mode prediction. *Ergonomics* 1999;42(11):1457-71.

59. Reason JT. *Human Error*. New York: Cambridge University Press, 1990.

60. Wears RL, Janiak B, Moorhead JC, Kellermann AL, Yeh CS, Rice MM, et al. Human error in medicine: promise and pitfalls, part 1. *Ann Emerg Med* 2000;36(1):58-60.

61. Leape LL, Bates DW, Cullen DJ, Cooper J, Demonaco HJ, Gallivan T, et al. Systems analysis of adverse drug events. ADE Prevention Study Group. *JAMA* 1995;274(1):35-43.

62. Berwick DM. Not again! *BMJ* 2001;322(7281):247-8.

63. 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.

64. Office of National Statistics. Mortality Statistics. Review of the Registrar General on deaths by cause, sex and age in England and Wales 2005. London: HMSO, 2006.

65. Garber J, Leadbeater C. Dying for change. London: DEMOS, 2010.

66. Metersky ML, Waterer G, Nsa W, Bratzler DW. Predictors of in-hospital vs postdischarge mortality in pneumonia. *Chest* 2012;142(2):476-81.

67. Department of Health. The NHS Outcomes Framework 2013/14. London: Department of Health, 2012:p.12.

68. Baglin TP, Cousins D, Keeling DM, Perry DJ, Watson HG. Safety indicators for inpatient and outpatient oral anticoagulant care: Recommendations from the British Committee for Standards in Haematology and National Patient Safety Agency. *Br J Haematol* 2007;136(1):26-9.

69. Patient Safety First Campaign. Patient Safety First 2008 to 2010. The campaign review. London: National Patient Safety Agency, 2011.

70. National Institute for Health and Care Excellence. Acutely ill patients in hospital: NICE guideline CG50. London: National Institute for Health and Care Excellence, 2007.

71. Keogh B. Review into the quality of care and treatment provided by 14 hospital trusts in England: overview report. London: Department of Health, 2013.

72. Institute of Medicine. Medicare: a strategy for quality assurance. Volume 1. Washington DC: National Academy Press, 1990.

73. World Health Organisation. WHO Performance Assessment Tool for Quality Improvement in Hospitals (PATH) Implementation Manual. Geneva: WHO, 2006.

74. Vincent C. *Patient Safety*. Edinburgh: Elsevier Churchill Livingstone, 2006.

75. Jost TS. Medicare peer review organizations. *Qual Assur Health Care* 1989;1(4):235-48.

76. Street A. The resurrection of hospital mortality statistics in England. *J Health Serv Res Policy* 2002;7(2):104-10.

77. Angelow A, Black N. The use and impact of national confidential enquiries in high-income countries. *BMJ Qual Saf* 2011;20(1):38-45.

78. Weisman CS, Morlock LL, Teitelbaum MA, Klassen AC, Celentano DD. Practice changes in response to the malpractice litigation climate. Results of a Maryland physician survey. *Med Care* 1989;27(1):16-24.

79. Barach P, Small SD. Reporting and preventing medical mishaps: lessons from non-medical near miss reporting systems. *BMJ* 2000;320(7237):759-63.

80. National Patient Safety Agency. Building a memory: preventing harm, reducing risks and improving patient safety. London: National Patient Safety Agency, 2005.

81. National Reporting and Learning System. Organisation Patient Safety Incident Reports. Ninth data release: 20 March 2013. London: NHS Commissioning Board, 2013.

82. Scott T, Mannion R, Davies H, Marshall M. The quantitative measurement of organizational culture in health care: a review of the available instruments. *Health Serv Res* 2003;38(3):923-45.

83. Public Health England. Results for Mandatory Surveillance Data. London: Public Health England, 2013.

84. Institute for Healthcare Improvement. Move Your Dot. Measuring, evaluating and reducing hospital mortality rates. Innovation Series 2003. Massachusetts, USA: Institute for Healthcare Improvement, 2003.

85. 1000 Lives Plus. A guide to measuring mortality. Cardiff: 1000 Lives Plus, 2010.

86. Dr Foster Intelligence. Building the future. Dr Foster Hospital Guide 2013. London: Dr Foster Intelligence, 2013.

87. NHS Confederation. Hospital standardised mortality ratios: their uses and abuses. Briefing. London: NHS Confederation, 2010.

88. Institute for Innovation and Improvement. Reducing avoidable mortality- Cheif Executives lead the way. Leeds: NHS Institute for Innovation and Improvement, 2007.

89. Modernisation Agency. A matter of life and death. Increasing reliability and quality to reduce hospital mortality and improve end of life care. London: Department of Health, 2004.

90. NHS Institute for Innovation and Improvement. Reducing avoidable deaths (Medical Directors): Department of Health, 2007.

91. Zhan C, Miller M. Administrative data based patient safety research: a clinical review. *Qual Saf Health Care* 2003;2003(12 (supp II)):ii158-63.

92. Australian Institute for Health and Welfare. Reporting of adverse event in routinely collected data sets in Australia: Australian Institute for Health and Welfare, 2001.

93. Jarman B, Gault S, Alves B, Hider A, Dolan S, Cook A, et al. Explaining differences in English hospital death rates using routinely collected data. *BMJ* 1999;318(7197):1515-20.

94. Marshall MN, Shekelle PG, Leatherman S, Brook RH. The public release of performance data: what do we expect to gain? A review of the evidence. *JAMA* 2000;283(14):1866-74.

95. Warden J. Hospital death rates to be published for England. *BMJ* 1998;316:1767.

96. Bridgewater B, Grayson AD, Brooks N, Grotte G, Fabri BM, Au J, et al. Has the publication of cardiac surgery outcome data been associated with changes in practice in northwest England: an analysis of 25,730 patients undergoing CABG surgery under 30 surgeons over eight years. *Heart* 2007;93(6):744-8.

97. Shekelle PG, Lim Y, Mattke S, Danberg C. Does public release of performance results improve quality of care? A systematic review. London: Health Foundation, 2008.

98. Smith P. On the unintended consequences of publishing performance data in the public sector. *Int J Pub Admin* 1995;18:277-31.

99. Bates DW, Evans RS, Murff H, Stetson PD, Pizziferri L, Hripcsak G. Detecting adverse events using information technology. *J Am Med Inform Assoc* 2003;10(2):115-28.

100. Roos LL, Brazauskas R. Outcomes and quality assurance: facilitating the use of administrative data. *Qual Assur Health Care* 1990;2(1):77-88.

101. Evans RS, Larsen RA, Burke JP, Gardner RM, Meier FA, Jacobson JA, et al. Computer surveillance of hospital-acquired infections and antibiotic use. *JAMA* 1986;256(8):1007-11.

102. Koch KE. Use of standardized screening procedures to identify adverse drug reactions. *Am J Hosp Pharm* 1990;47(6):1314-20.

103. Iezzoni L, Daley J, Heeren T, Foley S, Fisher E, Duncan C, et al. Identifying complications of care using administrative data. *Med Care* 1994;32:700-15.

104. McDonald K, Romano P, Geppert J, Davies S, Duncan B, Shojania K, et al. Evidence report for measure of patient safety based on hospital administrative data. Rockville, MD: Agency for Healthcare Research and Quality, 2002.

105. Raleigh VS, Cooper J, Bremner SA, Scobie S. Patient safety indicators for England from hospital administrative data: case-control analysis and comparison with US data. *BMJ* 2008;337:a1702.

106. Honigman B, Light P, Pulling RM, Bates DW. A computerized method for identifying incidents associated with adverse drug events in outpatients. *Int J Med Inform* 2001;61(1):21-32.

107. Battles JB, Kaplan HS, Van der Schaaf TW, Shea CE. The attributes of medical event-reporting systems: experience with a prototype medical event-reporting system for transfusion medicine. *Arch Pathol Lab Med* 1998;122(3):231-8.

108. Beckmann U, West LF, Groombridge GJ, Baldwin I, Hart GK, Clayton DG, et al. The Australian Incident Monitoring Study in Intensive Care: AIMS-ICU. The development and evaluation of an incident reporting system in intensive care. *Anaesth Intensive Care* 1996;24(3):314-9.

109. Kaplan HS, Rabin Fastman B. Organization of event reporting data for sense making and system improvement. *Qual Saf Health Care* 2003;12:ii68.

110. National Reporting and Learning System. Quarterly National Reporting and Learning System data summary. London: National Reporting and Learning System, 2013.

111. Firth-Cozens J. Barriers to incident reporting. *Qual Saf Health Care* 2002;11(1):7.

112. Busse DK, Wright DJ. Classification and analysis of incidents in complex medical environments. *Top Health Inf Manage* 2000;20(4):1-11.

113. Ricci M, Goldman A, de Leval M, et al. Pitfalls of adverse event reporting in paediatric intensive care. *Arch Dis Child* 2004;89:856-9.

114. Gordon L. *Gordon's Guide to the Surgical Morbidity and Mortality Conference* Philadelphia: Hanley and Belfus, 1994.

115. Travaglia J, Debono D. Mortality and morbidity reviews: a comprehensive review of the literature. Sydney: Centre for Clinical Governance Research in Health, Faculty of Medicine, University of New South Wales, 2009.

116. Shojania KG, Burton EC, McDonald KM, Goldman L. Changes in rates of autopsy-detected diagnostic errors over time: a systematic review. *JAMA* 2003;289(21):2849-56.

117. Sington J, Cottrell B. Medical error reporting must take necropsy data into account. *BMJ* 2001;323(7311):511.

118. Wachter R, Shojania K, Minichiello T. AHRQ Web M&M- Online medical error reporting and analysis, 2005.

119. Smits M, Zegers M, Groenewegen PP, Timmermans DR, Zwaan L, van der Wal G, et al. Exploring the causes of adverse events in hospitals and potential prevention strategies. *Qual Saf Health Care* 2010;19(5):e5.

120. Bates D, Cullen D, Laird N. Incidence of adverse drug events and potential adverse drug events. Implications for prevention. *JAMA* 1995;274:29-34.

121. National Confidential Enquiry into Patient Outcome and Death. Deaths in acute hospitals: Caring to the end? London: National Confidential Enquiry into Patient Outcome and Death, 2009.

122. National Confidential Enquiry into Patient Outcome and Death. AKI: Adding insult to injury. London: National Confidential Enquiry into Patient Outcome and Death, 2009.

123. National Confidential Enquiry into Patient Outcome and Death. Time to intervene. London: National Confidential Enquiry into Patient Outcome and Death, 2012.

124. Jick H. Drugs: remarkably nontoxic. *N Engl J Med* 1974;291(16):824-8.

125. Classen DC, Pestotnik SL, Evans RS, Burke JP. Computerized surveillance of adverse drug events in hospital patients. *JAMA* 1991;266(20):2847-51.

126. Rozich J, Haraden C, Resar R. Adverse drug event trigger tool: a practical methodology for measuring medication related harm. *Qual Saf Health Care* 2003;12:194-200.

127. Griffin F, Resar R. Innovation Series 2009. IHI Global Trigger Tool for measuring adverse events. Second edition: Institute of Healthcare Improvement, 2009.

128. 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.

129. Bruce J, Russell EM, Mollison J, Krukowski ZH. The measurement and monitoring of surgical adverse events. *Health Technol Assess* 2001;5(22):1-194.

130. Wilson B, Thornton JG, Hewison J, Lilford RJ, Watt I, Braunholtz D, et al. The Leeds University Maternity Audit Project. *Int J Qual Health Care* 2002;14(3):175-81.

131. Naessens JM, Campbell CR, Huddleston JM, Berg BP, Lefante JJ, Williams AR, et al. A comparison of hospital adverse events identified by three widely used detection methods. *Int J Qual Health Care* 2009;21(4):301-7.

132. Classen DC, Resar R, Griffin F, Federico F, Frankel T, Kimmel N, et al. 'Global trigger tool' shows that adverse events in hospitals may be ten times greater than previously measured. *Health Aff (Millwood)* 2011;30(4):581-9.

133. Cohen MM, Kimmel NL, Benage MK, Cox MJ, Sanders N, Spence D, et al. Medication safety program reduces adverse drug events in a community hospital. *Qual Saf Health Care* 2005;14(3):169-74.

134. Vincent CA. Research into medical accidents: a case of negligence? *BMJ* 1989;299(6708):1150-3.

135. Cheney FW, Posner K, Caplan RA, Ward RJ. Standard of care and anesthesia liability. *JAMA* 1989;261(11):1599-603.

136. Eichhorn JH, Cooper JB, Cullen DJ, Maier WR, Philip JH, Seeman RG. Standards for patient monitoring during anesthesia at Harvard Medical School. *JAMA* 1986;256(8):1017-20.

137. Rolph JE, Kravitz RL, McGuigan K. Malpractice claims data as a quality improvement tool. II. Is targeting effective? *JAMA* 1991;266(15):2093-7.

138. Vincent C, Davy C, Esmail A, Neale G, Elstein M, Cozens JF, et al. Learning from litigation. The role of claims analysis in patient safety. *J Eval Clin Pract* 2006;12(6):665-74.

139. Esmail A. Patient safety- what claims against the NHS can teach us. *J Health Serv Res Policy* 2010;15 Suppl 1:33-6.

140. Rothschild JM, Federico FA, Gandhi TK, Kaushal R, Williams DH, Bates DW. Analysis of medication-related malpractice claims: causes, preventability, and costs. *Arch Intern Med* 2002;162(21):2414-20.

141. Usha Kiran TS, Jayawickrama NS. Complaints and claims in the UK National Health Service. *J Eval Clin Pract* 2002;8(1):85-6.

142. Vincent C, Young M, Phillips A. Why do people sue doctors? A study of patients and relatives taking legal action. *Lancet* 1994;343(8913):1609-13.

143. 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.

144. Forster AJ, Worthington JR, Hawken S, Bourke M, Rubens F, Shojania K, et al. Using prospective clinical surveillance to identify adverse events in hospital. *BMJ Qual Saf* 2011;20(9):756-63.

145. O'Neil A, Peterson L, Cook F, et a. Physician reporting compared with medical record review to identify adverse medical events. *Ann Int Med* 1993;119:370-6.

146. Michel P, Quenon J, de Saraqueta A, Scemama O. Comparison of three methods for estimating rates of preventable adverse events in acute care hospitals. *BMJ* 2004;328:199 ; doi:10.1136/bmj.328.7433.199

147. Andrews L, Stocking C, Krizek T, et al. An alternative strategy for studying adverse events in medical care. *Lancet* 1997;349:309-13.

148. Flynn EA, Barker KN, Pepper GA, Bates DW, Mikeal RL. Comparison of methods for detecting medication errors in 36 hospitals and skilled-nursing facilities. *Am J Health Syst Pharm* 2002;59(5):436-46.

149. Donchin Y, Gopher D, Olin M, al e. A look into the nature and causes of human errors in the intensive care unit. *Crit Care Med* 1995;23:294–300.

150. Morey JC, Simon R, Jay GD, Wears RL, Salisbury M, Dukes KA, et al. Error reduction and performance improvement in the emergency department through formal teamwork training: evaluation results of the MedTeams project. *Health Serv Res* 2002;37(6):1553-81.

151. Catchpole K, Bell MD, Johnson S. Safety in anaesthesia: a study of 12,606 reported incidents from the UK National Reporting and Learning System. *Anaesthesia* 2008;63(4):340-6.

152. Carthey J. The role of structured observational research in health care. *Qual Saf Health Care* 2003;12 Suppl 2:ii13-6.

153. Blendon RJ, DesRoches CM, Brodie M, Benson JM, Rosen AB, Schneider E, et al. Views of practicing physicians and the public on medical errors. *N Engl J Med* 2002;347(24):1933-40.

154. Weingart SN, Pagovich O, Sands DZ, Li JM, Aronson MD, Davis RB, et al. What can hospitalized patients tell us about adverse events? Learning from patient-reported incidents. *J Gen Intern Med* 2005;20(9):830-6.

155. Davis RE, Sevdalis N, Neale G, Massey R, Vincent CA. Hospital patients' reports of medical errors and undesirable events in their health care. *J Eval Clin Pract* 2013;19(5):875-81.

156. Forster AJ, Murff HJ, Peterson JF, Gandhi TK, Bates DW. The incidence and severity of adverse events affecting patients after discharge from the hospital. *Ann Intern Med* 2003;138(3):161-7.

157. Agoritsas T, Bovier PA, Perneger TV. Patient reports of undesirable events during hospitalization. *J Gen Intern Med* 2005;20(10):922-8.

158. King A, Daniels J, Lim J, Cochrane DD, Taylor A, Ansermino JM. Time to listen: a review of methods to solicit patient reports of adverse events. *Qual Saf Health Care* 2010;19(2):148-57.

159. Waterman AD, Gallagher TH, Garbutt J, Waterman BM, Fraser V, Burroughs TE. Brief report: Hospitalized patients' attitudes about and participation in error prevention. *J Gen Intern Med* 2006;21(4):367-70.

160. Davis RE, Jacklin R, Sevdalis N, Vincent CA. Patient involvement in patient safety: what factors influence patient participation and engagement? *Health Expect* 2007;10(3):259-67.

161. Girou E, Stephan F, Novara A, Safar M, Fagon JY. Risk factors and outcome of nosocomial infections: results of a matched case-control study of ICU patients. *Am J Respir Crit Care Med* 1998;157(4 Pt 1):1151-8.

162. Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP. Adverse drug events in hospitalized patients. Excess length of stay, extra costs, and attributable mortality. *JAMA* 1997;277(4):301-6.

163. Gross PA, Van Antwerpen C. Nosocomial infections and hospital deaths. A case-control study. *Am J Med* 1983;75(4):658-62.

164. Garcia- Martin M, Lardelli- Claret P, Bueno- Cavanillas A, de Dios Luna-del-Castillo J, Espigares-Garcia M, Galvez-Vargas R. Proportion of hospital deaths associated with adverse events. *J Clin Epidemiol* 1997;50(12):1319-26.

165. Zegers M, de Bruijne M, Wagner C, P G, van der Wal G. Differences in adverse events between discharged and deceased patients. Poster Presentation. *International Forum on Quality and Safety*. Amsterdam, 2010.

166. Gawande AA, Studdert DM, Orav EJ, Brennan TA, Zinner MJ. Risk factors for retained instruments and sponges after surgery. *N Engl J Med* 2003;348(3):229-35.

167. Thomas EJ, Studdert DM, Runciman WB, Webb RK, Sexton EJ, Wilson RM, et al. A comparison of iatrogenic injury studies in Australia and the USA. I: Context, methods, casemix, population, patient and hospital characteristics. *Int J Qual Health Care* 2000;12(5):371-8.

168. Thomas EJ, Brennan TA. Incidence and types of preventable adverse events in elderly patients: population based review of medical records. *BMJ* 2000;320(7237):741-4.

169. Wilson R, Runciman W, Gibberd R, Harrison B, Hamilton J. Quality in Australian Health Care Study. *Med J Aust* 1995;163(9):472-5.

170. Leape LL, Brennan TA, Laird N, Lawthers AG, Localio AR, Barnes BA, et al. 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.

171. Hutchinson A, McIntosh A, Dean JE, Cooper KL, Bath PA, Walters SJ, et al. When is an event an event? Data from the quality and safety continuum. *From safe design to safe practice*. Cambridge: The Ergonomics Society, 2008.

172. Office of National Statistics. Mortality Statistics. Review of the Registrar General on deaths by cause, sex and age in England and Wales 2005. London: HMSO, 2006.

173. Wilson RM, Harrison BT, Gibberd RW, Hamilton JD. An analysis of the causes of adverse events from the Quality in Australian Health Care Study. *Med J Aust* 1999;170(9):411-5.

174. Zegers M, de Bruijne MC, Wagner C, Hoonhout LH, Waaijman R, Smits M, et al. Adverse events and potentially preventable deaths in Dutch hospitals: results of a

retrospective patient record review study. *Qual Saf Health Care* 2009;18(4):297-302.

175. Davis P, Lay-Yee R, Briant R, Ali W, Scott A, Schug S. Adverse events in New Zealand public hospitals I: occurrence and impact. *N Z Med J* 2002;115(1167):U271.

176. Baker R, Norton P, Flintoft V, Blais R, Brown A, Cox J, et al. The Canadian Adverse Events Study: the incidence of adverse events among hospital patients in Canada. *Canadian Medical Association* 2004;170(11):1678-86.

177. Schiøler T, Lipczak H, Pedersen B, Mogensen T, Bech K, Stockmarr A, et al. Danish Adverse Event Study *Ugeskrift for Laeger* 2001;163(39):5370-78.

178. Sari AB, Sheldon TA, Cracknell A, Turnbull A. Sensitivity of routine system for reporting patient safety incidents in an NHS hospital: retrospective patient case note review. *BMJ* 2007;334(7584):79.

179. Baines RJ, Langelaan M, de Bruijne MC, Asscheman H, Spreeuwenberg P, van de Steeg L, et al. Changes in adverse event rates in hospitals over time: a longitudinal retrospective patient record review study. *BMJ Qual Saf* 2013;22(4):290-8.

180. Cook R, Woods D. *Operating at the sharp end: the complexity of human error*. Hillsdale NJ:: Lawrence Erlbaum Associates, 1994.

181. Stanhope C, Crowley-Murphey M, Vincent C, et al. An evaluation of adverse incident reporting. *J Eval Clin Pract* 1995;5:1-4.

182. Mold JW, Stein HF. The cascade effect in the clinical care of patients. *N Engl J Med* 1986;314(8):512-4.

183. Thornlow DK, Anderson R, Oddone E. Cascade iatrogenesis: factors leading to the development of adverse events in hospitalized older adults. *Int J Nurs Stud* 2009;46(11):1528-35.

184. Brennan TA, Localio RJ, Laird NL. Reliability and validity of judgments concerning adverse events suffered by hospitalized patients. *Med Care* 1989;27(12):1148-58.

185. Michel P. Strengths and weaknesses of available methods for assessing the nature and scale of harm caused by the health system: literature review. Geneva: World Health Organisation/ Patient Safety World Alliance for Safer Health Care, 2003.

186. Kirshner B, Guyatt G. A methodological framework for assessing health indices. *J Chronic Dis* 1985;38(1):27-36.

187. Streiner D, Norman G. *Health measurement scales: a practical guide to their development and use. Fourth Edition*. Oxford: Oxford University Press, 2008.

188. Walshe K. The validity and reliability of of adverse event measures of quality of healthcare (PhD Thesis). University of Birmingham, 1998.

189. Robinson G. Providing assurance, driving improvement. Learning from mortality and harm reviews in NHS Wales. Improving Healthcare White Paper Series No. 10. Cardiff: 1000 Lives Plus.

190. Weingert S, Wilson R, Gibberd R, Harrison B. Epidemiology of medical error. *BMJ* 2000;320:774-7.

191. Zhan C, Miller MR. Excess length of stay, charges, and mortality attributable to medical injuries during hospitalization. *JAMA* 2003;290(14):1868-74.

192. McCarthy E, Iezzoni L, Davis R, et a. Does clinical evidence support ICD-9 CM diagnosis coding of complications. *Med Care* 2000;38:868-76.

193. Silber JH, Romano PS, Rosen AK, Wang Y, Even-Shoshan O, Volpp KG. Failure-to-rescue: comparing definitions to measure quality of care. *Med Care* 2007;45(10):918-25.

194. Isaac T, Jha AK. Are patient safety indicators related to widely used measures of hospital quality? *J Gen Intern Med* 2008;23(9):1373-8.

195. Mihrshahi S, Brand C, Ibrahim JE, Evans S, Jolley D, Cameron P. Validity of the indicator 'death in low-mortality diagnosis-related groups' for measuring patient safety and healthcare quality in hospitals. *Intern Med J* 2010;40(4):250-7.

196. Perkins GD, McAuley DF, Davies S, Gao F. Discrepancies between clinical and postmortem diagnoses in critically ill patients: an observational study. *Crit Care* 2003;7(6):R129-32.

197. Anderson RE, Hill RB, Key CR. The sensitivity and specificity of clinical diagnostics during five decades. Toward an understanding of necessary fallibility. *JAMA* 1989;261(11):1610-7.

198. Johnson CW. How will we get the data and what will we do with it then? Issues in the reporting of adverse healthcare events. *Qual Saf Health Care* 2003;12 Suppl 2:ii64-7.

199. Lawton R, Parker D. Barriers to incident reporting in a healthcare system. *Qual Saf Health Care* 2002;11(1):15-8.

200. Hutchinson A, Coster JE, Cooper KL, McIntosh A, Walters SJ, Bath PA, et al. Comparison of case note review methods for evaluating quality and safety in health care. *Health Technol Assess* 2010;14(10):iii-iv, ix-x, 1-144.

201. Ashton CM, Kuykendall DH, Johnson ML, Wray NP. An empirical assessment of the validity of explicit and implicit process-of-care criteria for quality assessment. *Med Care* 1999;37(8):798-808.

202. Weingart SN, Ship AN, Aronson MD. Confidential clinician-reported surveillance of adverse events among medical inpatients. *J Gen Intern Med* 2000;15(7):470-7.

203. Weingart SN, Callanan LD, Ship AN, Aronson MD. A physician-based voluntary reporting system for adverse events and medical errors. *J Gen Intern Med* 2001;16(12):809-14.

204. Bland M. An introduction to medical statistics. Oxford: Oxford University Press, 1995.

205. Thomas EJ, Lipsitz SR, Studdert DM, Brennan TA. The reliability of medical record review for estimating adverse event rates. *Ann Intern Med* 2002;136(11):812-6.

206. Dubois R, Brook R. Preventable deaths: who, how often and why. *Ann Intern Med* 1998;109:582-89.

207. Hofer TP, Bernstein SJ, DeMonner S, Hayward RA. Discussion between reviewers does not improve reliability of peer review of hospital quality. *Med Care* 2000;38(2):152-61.

208. Rubin HR, Rogers WH, Kahn KL, Rubenstein LV, Brook RH. Watching the doctor-watchers. How well do peer review organization methods detect hospital care quality problems? *JAMA* 1992;267(17):2349-54.

209. Hayward R, McMahon L, Bernard A. Evaluating the care of general medical inpatients: how good is structued implicit review? *Ann Intern Med* 1993;1993:550-56.

210. Zegers M, de Bruijne MC, Wagner C, Groenewegen PP, van der Wal G, de Vet HC. The inter-rater agreement of retrospective assessments of adverse events does not improve with two reviewers per patient record. *J Clin Epidemiol* 2010;63(1):94-102.

211. Liliford R, Edwards A, Girling A, Hofer T, Di Tanna G, Petty J, et al. Inter-rater reliability of case-note audit: a systematic review. *J Health Serv Res Policy* 2007;12:173-80.

212. Naessens JM, O'Byrne TJ, Johnson MG, Vansuch MB, McGlone CM, Huddleston JM. Measuring hospital adverse events: assessing inter-rater reliability and trigger performance of the Global Trigger Tool. *Int J Qual Health Care* 2010;22(4):266-74.

213. Weingart SN, Davis RB, Palmer RH, Cahalane M, Hamel MB, Mukamal K, et al. Discrepancies between explicit and implicit review: physician and nurse assessments of complications and quality. *Health Serv Res* 2002;37(2):483-98.

214. Davis P, Lay-Yee R, Briant R, al e. Adverse events in New Zealand public hospitals I: occurance and impact. *N Z Med J* 2002;115(1167):U268.

215. Rubenstein LV, Kahn KL, Reinisch EJ, Sherwood MJ, Rogers WH, Kamberg C, et al. Changes in quality of care for five diseases measured by implicit review, 1981 to 1986. *JAMA* 1990;264(15):1974-9.

216. Caplan R, Posner K, Cheney F. Effect of outcome on physicians' judgements of appropriateness of care. *JAMA* 1991;265:1957-60.

217. Hale A. Why safety performance indicators? *Safety Science* 2008;10:1016-18.

218. Luck J, Peabody JW, Dresselhaus TR, Lee M, Glassman P. How well does chart abstraction measure quality? A prospective comparison of standardized patients with the medical record. *Am J Med* 2000;108(8):642-9.

219. National Patient Safety Agency. The fifth report of the Patient Safety Observatory. Safer care for the severely ill patient: learning from serious incidents,. London: National Patient Safety Agency, 2005.

220. Iezzoni LI. Assessing quality using administrative data. *Ann Intern Med* 1997;127(8 Pt 2):666-74.

221. Raftery J, Roderick P, Stevens A. Potential uses of routine data in health technology assessment. *Health Technology Assessment* 2005;9(20).

222. McCarthy EP, Iezzoni LI, Davis RB, Palmer RH, Cahalane M, Hamel MB, et al. Does clinical evidence support ICD-9-CM diagnosis coding of complications? *Med Care* 2000;38(8):868-76.

223. Lawthers AG, McCarthy EP, Davis RB, Peterson LE, Palmer RH, Iezzoni LI. Identification of in-hospital complications from claims data. Is it valid? *Med Care* 2000;38(8):785-95.

224. McKee M, Coles J, James P. 'Failure to rescue' as a measure of quality of hospital care: the limitations of secondary diagnosis coding in English hospital data. *J Public Health Med* 1999;21(4):453-8.

225. Griffiths P, Jones S, Bottle A. Is "failure to rescue" derived from administrative data in England a nurse sensitive patient safety indicator for surgical care? Observational study. *Int J Nurs Stud* 2013;50(2):292-300.

226. Pronovost PJ, Miller MR, Wachter RM. Tracking progress in patient safety: an elusive target. *JAMA* 2006;296(6):696-9.

227. 1000 Lives Plus. Providing assurance, driving improvement. Learning from mortality and harm reviews in NHS Wales. Improving Healthcare White Paper Series Number 10. Cardiff: 1000 Lives Plus, 2013.

228. Friese CR, Earle CC, Silber JH, Aiken LH. Hospital characteristics, clinical severity, and outcomes for surgical oncology patients. *Surgery* 2010;147(5):602-9.

229. Bates DW, O'Neil AC, Petersen LA, Lee TH, Brennan TA. Evaluation of screening criteria for adverse events in medical patients. *Med Care* 1995;33(5):452-62.

230. Olsen S, Neale G, Schwab K, Psaila B, Patel T, Chapman EJ, et al. Hospital staff should use more than one method to detect adverse events and potential adverse events: incident reporting, pharmacist surveillance and local real-time record review may all have a place. *Qual Saf Health Care* 2007;16(1):40-4.

231. Davis RE, Sevdalis N, Neale G, Massey R, Vincent CA. Hospital patients' reports of medical errors and undesirable events in their health care. *J Eval Clin Pract* 2012.

232. Brennan TA, Localio AR, Leape LL, Laird NM, Peterson L, Hiatt HH, et al. Identification of adverse events occurring during hospitalization. A cross-sectional study of litigation, quality assurance, and medical records at two teaching hospitals. *Ann Intern Med* 1990;112(3):221-6.

233. Shahian DM, Wolf RE, Iezzoni LI, Kirle L, Normand SL. Variability in the measurement of hospital-wide mortality rates. *N Engl J Med* 2010;363(26):2530-9.

234. Girling AJ, Hofer TP, Wu J, Chilton PJ, Nicholl JP, Mohammed MA, et al. Case-mix adjusted hospital mortality is a poor proxy for preventable mortality: a modelling study. *BMJ Qual Saf* 2012;21(12):1052-6.

235. Woloshynowych M, Neale G, Vincent C. Case record review of adverse events: a new approach. *Qual Saf Health Care* 2003;12:411-5.

236. Quach S, Hennessy DA, Faris P, Fong A, Quan H, Doig C. A comparison between the APACHE II and Charlson Index Score for predicting hospital mortality in critically ill patients. *BMC Health Serv Res* 2009;9:129.

237. Lemeshow S, Le Gall JR. Modeling the severity of illness of ICU patients. A systems update. *JAMA* 1994;272(13):1049-55.

238. Coventry PA, Grande GE, Richards DA, Todd CJ. Prediction of appropriate timing of palliative care for older adults with non-malignant life-threatening disease: a systematic review. *Age Ageing* 2005;34(3):218-27.

239. Lau F, Downing M, Lesperance M, Karlson N, Kuziemsky C, Yang J. Using the Palliative Performance Scale to provide meaningful survival estimates. *J Pain Symptom Manage* 2009;38(1):134-44.

240. Mor V, Laliberte L, Morris JN, Wiemann M. The Karnofsky Performance Status Scale. An examination of its reliability and validity in a research setting. *Cancer* 1984;53(9):2002-7.

241. Benning A, Ghaleb M, Suokas A, Dixon-Woods M, Dawson J, Barber N, et al. Large scale organisational intervention to improve patient safety in four UK hospitals: mixed method evaluation. *BMJ* 2011;342:d195.

242. Neale G, Chapman EJ, Hoare J, Olsen S. Recognising adverse events and critical incidents in medical practice in a district general hospital. *Clin Med* 2006;6(2):157-62.

243. Vincent C, Taylor-Adams S, Stanhope N. Framework for analysing risk and safety in clinical medicine. *BMJ* 1998;316(7138):1154-7.

244. Davis P, Lay-Yee R, Schug S, Briant R, Scott A, Johnson S, et al. Adverse events regional feasibility study: methodological results. *NZ Med J* 2001;114:200-2.

245. Vincent C, Young M, Phillips A. Why do people sue doctors? A study of patients and relatives taking legal action. *Lancet*;343(1609-13).

246. Bland M. *An introduction to medical statistics. Second edition*. Oxford: Oxford Medical Publications, 1999.

247. Spath P, editor. *Error reduction in healthcare: a systems approach to improving patient safety*. Washington: AHA Press, 1999.

248. Adams S, Vincent C. Systems analysis of clinical incidents: the London protocol. Available at: http://www1.imperial.ac.uk/resources/C85B6574-7E28-4BE6-BE61-E94C3F6243CE/londonprotocol_e.pdf. Accessed 20th January 2013.

249. Carroll JS, Rudolph JW, Hatakenaka S. Lessons learned from non-medical industries: root cause analysis as culture change at a chemical plant. *Qual Saf Health Care* 2002;11(3):266-9.

250. National Patient Safety Agency: National Reporting and Learning System. Root Cause Analysis Tools: Guide to report writing following Root Cause Analysis of Patient Safety Incidents. London: National Patient Safety Agency, 2008.

251. Pitches DW, Mohammed MA, Lilford RJ. What is the empirical evidence that hospitals with higher-risk adjusted mortality rates provide poorer quality care? A systematic review of the literature. *BMC Health Serv Res* 2007;7:91.

252. Scott IA, Brand CA, Phelps GE, Barker AL, Cameron PA. Using hospital standardised mortality ratios to assess quality of care--proceed with extreme caution. *Med J Aust* 2011;194(12):645-8.

253. Office of National Statistics. Deaths involving MRSA: England and Wales 2007 to 2011. Statistical Bulletin. London: Office of National Statistics, 2012.

254. Hutchinson A, Young TA, Cooper KL, McIntosh A, Karnon JD, Scobie S, et al. Trends in healthcare incident reporting and relationship to safety and quality data in acute hospitals: results from the National Reporting and Learning System. *Qual Saf Health Care* 2009;18(1):5-10.

255. 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(3):163-9.

256. Neale G, Woloshynowych M, Vincent C. Exploring the causes of adverse events in NHS hospital practice. *J R Soc Med* 2001;94:322-30.

257. Neale G, Woloshynowych M. Retrospective case record review: a blunt instrument that needs sharpening. *Qual Saf Health Care* 2003;12:2-3.

258. Hogan H, Healey F, Neale G, Thomson R, Vincent C, Black N. Preventable deaths due to problems in care in English acute hospitals: a retrospective case record review study. *BMJ Qual Saf* 2012**;**21(9):737-45.

259. Nolan J, Gallegher L, Lloyd-Scott L, Rowan K. National Cardiac Arrest Audit (NCAA). *JICS* 2009;10:313-15.

260. Seward E, Greig E, Preston S, Harris RA, Borrill Z, Wardle TD, et al. A confidential study of deaths after emergency medical admission: issues relating to quality of care. *Clin Med* 2003;3(5):425-34.

261. National Confidential Enquiry into Patient Outcome and Death. An acute problem? London: NCEPOD, 2005.

262. National Confidential Enquiry into Patient Outcome and Death. Extremes of age. London: NCEPOD, 1999.

263. Department of Health. Report of an independent expert working group on the prevention of venous thromboembolism in hospitalised patients. A report to Sir Liam Donaldson, Chief Medical Officer. London: Department of Health, 2007.

264. Robb G, Seddon M. A multi-faceted approach to the physiologically unstable patient. *Qual Saf Health Care* 2010;19(5):e47.

265. Raduma-Tomas MA, Flin R, Yule S, Williams D. Doctors' handovers in hospitals: a literature review. *BMJ Qual Saf* 2011;20(2):128-33.

266. Godhill D, McNarry A. Intensive care outreach services. *Current Anaesthesia and Critical Care* 2002;13(6):346-61.

267. NHS Institute for Innovation and Improvement. The Productive Series. Coventry: NHS Institute for Innovation and Improvement.

268. Pronovost P, Needham D, Berenholtz S, Sinopoli D, Chu H, Cosgrove S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med* 2006;355(26):2725-32.

269. Dixon-Woods M, Bosk CL, Aveling EL, Goeschel CA, Pronovost PJ. Explaining Michigan: developing an ex post theory of a quality improvement program. *Milbank Q* 2011;89(2):167-205.

270. Freemantle N, Richardson M, Wood J, Ray D, Khosla S, Shahian D, et al. Weekend hospitalization and additional risk of death: an analysis of inpatient data. *J R Soc Med* 2012;105(2):74-84.

271. Keogh B. Review into the quality of care and treatment provided by 14 hospital trusts in England: overview report by Professor Sir Bruce Keogh KBE. London: Department of Health, 2013.

272. Health and Social Care Information Centre. Hospital Episode Statistics: Admitted patient care 2011-12. Summary Report, 1st November 2012. Leeds: Health and Social Care Information Centre 2013.

273. Cornwall J, Levenson R, Sonola L, Poteliakhoff E. Continuity of care for older hospital patients: A call for action. London: King's Fund, 2012.

274. Lafont C, Gerard S, Voisin T, Pahor M, Vellas B. Reducing "iatrogenic disability" in the hospitalized frail elderly. *J Nutr Health Aging* 2011;15(8):645-60.

275. Abraham A. Care and Compassion? Report of the Health Service Ombudsman on ten investigations into NHS care of older people. Fourth report of the Health Service Commissioner for England; session 2010-2011. HC778. London: The Stationary Office, 2011.

276. Royal College of Physicians. Acute Care Tool Kit 3. Acute medical care for frail older people. London: Royal College of Physicians, 2012.

277. Royal College of Physicians. Acute Care Toolkit 2: High quality acute care. London: Royal College of Physicians, 2012.

278. Beswick AD, Rees K, Dieppe P, Ayis S, Gooberman-Hill R, Horwood J, et al. Complex interventions to improve physical function and maintain independent living in elderly people: a systematic review and meta-analysis. *Lancet* 2008;371(9614):725-35.

279. McCulloch P, Mishra A, Handa A, Dale T, Hirst G, Catchpole K. The effects of aviation-style non-technical skills training on technical performance and outcome in the operating theatre. *Qual Saf Health Care* 2009;18(2):109-15.

280. Robinson M, Chorley M, Borland S, Martin D. Labour's day of shame over the NHS: Thousands of unnecessary deaths, 14 hospital trusts condemned, and now devastating report demolished Labour's claim that the NHS was safe in its hands. *Daily Mail Online* 2013 16 July 2013.

281. Black N. Assessing the quality of hospitals. *BMJ* 2010;340:c2066.

282. Francis R. Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry. London: The Stationary Office, 2013.

283. Department of Health. Patients first and foremost. The initial government response to the report of the Mid Staffordshire NHS Trust public enquiry. London: Depatment of Health, 2013.

284. Department of Health. Hard truths: The journey to putting patients first. London: Department of Health, 2013.

285. Department of Health. The NHS Outcomes Framework 2013-14. London: Department of Health, 2012.

286. Nuffield Trust. Spending on health and social care to 2015-16. Parlimentary Briefing. London: Nuffield Trust, 2013.

287. National Audit Office. Progress on making NHS efficiency savings. London: The Stationery Office, 2012.

288. Vincent C, Burnett S, Carthey J. The measurement and monitoring of safety. London: The Health Foundation, 2013.

289. Berwick DM. Errors today and errors tomorrow. *N Engl J Med* 2003;348(25):2570-2.

290. Leistikow IP, Kalkman CJ, de Bruijn H. Why patient safety is such a tough nut to crack. *BMJ* 2011;342:d3447.