How effective are incident reporting systems for improving patient safety? A systematic literature review*

Charitini Stavropoulou,[1] Carole Doherty,[2] and Paul Tosey[2]

1 School of Health Sciences, City University London, UK

2 Faculty of Business, Economics and Law, University of Surrey, UK

**This is a preprint of an Article accepted for publication* in *The Milbank Quarterly © [2015] The Milbank Memorial Fund*

Abstract

Context: Incident reporting systems (IRSs) are used to gather information on patient safety incidents. However, and despite the financial burden they imply, little is known about their effectiveness. This paper reviews systematically the effectiveness of IRSs as a method of improving patient safety through organizational learning.

Method: This systematic literature review identified two groups of studies: a) studies comparing the effectiveness of IRSs relative to other methods of error reporting and b) studies examining the effectiveness of IRSs on settings, structures and outcomes in respect of improvements to patient safety. We used thematic analysis to compare the effectiveness of IRSs with other methods and to synthesize what was effective, where and why. Then, to assess the evidence concerning the ability of IRSs to facilitate organizational learning, we analyzed studies using the concepts of single loop and double loop learning.

Findings: In total, 43 studies were identified. Eight studies compared IRSs with other methods, while 35 explored the effectiveness of IRSs on settings, structures and outcomes. We did not find

strong evidence that IRSs perform better than other methods. We found some evidence of single loop learning, that is, changes to *clinical* settings or processes as a consequence of learning from IRSs, but little evidence either of improvements to outcomes or of changes to latent managerial factors involved in error production. In addition, there was insubstantial evidence of IRSs enabling double loop learning that is, cultural change or change of mindset.

Conclusions: The results indicate IRSs could be more effective if there were explicit criteria for what counts as an incident; they are owned and led by clinical teams rather than centralized hospital departments; and embedded within organizations as part of wider safety programs.

Key words

Patient safety, Incident reporting systems, organizational learning, single loop learning, double loop learning

Introduction

To improve patient safety experts have argued that major cultural change firmly rooted in continual improvement is required.[1] Necessary changes include constant evidence-based learning; managerial appreciation of the pressures that resource constraints can bring for front-line employees; avoidance of blame; and eschewing mechanistic performance objectives.[1] Incident reporting systems (IRSs) are designed to be used to obtain information about patient safety, with this knowledge translated into individual and organizational learning.[2–4] Organizational learning is described as 'a process of individual and shared thought and action in an organizational context',[5(p470)] from which cultural change ensues. This systematic review examines evidence concerning the effectiveness of IRSs as one mechanism to promote organizational learning to improve patient safety. We define effectiveness in both relative and absolute terms. In relative terms, we examine the quantity and type of incidents reported using IRSs by comparison with other forms of incident reporting, such as medical chart review. In absolute terms, we use Donabedian's[6] framework to explore the impact of IRSs on settings (structure), processes and safety outcomes.

IRSs have been in use in the healthcare field for many years, but it was following the publication of 'To Err is Human'[7] that systems were implemented more widely. For example all public hospitals in Australia were required to have an Advanced Incident Monitoring System (AIMS) in place by January 2005, in the UK the National Reporting and Learning System (NRLS) was set up in 2003[8] and in Ireland the STARSweb IRS was launched in 2004.[9] To put this in context, the number of patient safety incidents reported to the NRLS in England between October 2011 and March 2012 was 612,414. Six percent of incidents resulted in moderate harm and 1 percent (n= 5,235) resulted in severe harm or death.[10]

However, there are questions about the effectiveness and cost of IRSs.[11] Renshaw et al.[12(p383)] estimated that 'the cost of the system was equivalent to 1,184 UK National Health Service (NHS) employees spending all their time each month completing incident forms'; these being time consuming to complete.[13] Waring[14] argues that the rich detailed information in clinicians' stories is reassigned via IRSs into abstract, quantitative variables of the managerial system, reducing the effectiveness of IRSs for learning. Wachter[15] argues that incident reports do not provide information about the true frequency of organizational errors, are too expensive and bureaucratic.

Other problems associated with IRSs include: the number of incidents reported reflects employees' willingness to report rather than being an indicator of the safety of the system[16]; there is no shared understanding between clinicians (doctors, nurses and other healthcare professionals) about what constitutes an adverse event or near miss; lack of clarity about who, within the clinical team, is responsible for reporting respective incidents[17]; and some clinicians may be fearful of recriminations.[18] Generally, patients do not have independent access to IRSs and their experiences of harm may go unrecognized by clinicians.[19,20] This raises questions about their utility as mechanisms to promote organizational learning to improve patient safety.

Health expenditure in most countries has been declining since the beginning of the global financial crisis in 2008.[21] It is therefore important to consider whether investing in IRSs is money well spent, for both the public and private health sectors.[12,22] This paper reports on a parallel review of (a) studies comparing the effectiveness of IRSs relative to other methods of error reporting and (b) studies designed to measure the effectiveness of IRSs in absolute terms. For the latter, Donabedian's[6] settings, processes and outcomes framework is used to review systematically empirical evidence on how effective incident reporting systems are for patient

safety. Pursuing and measuring both systems and outcomes improvements may identify success factors, thereby contributing to their enhanced sustainability.[4] Then, to assess the evidence concerning the ability of IRSs as mechanisms to promote organizational learning, these studies are analysed using Argyris and Schön's[23] concepts of single loop and double loop learning.

We begin by examining the background to, rationale for, and practical application of, IRSs. Following this, we discuss perspectives on organizational learning and select a theory suitable for the present study. Then we describe our review method before presenting our findings, firstly comparing IRSs with other systems and then assessing their effectiveness.

Incident Reporting Systems (IRSs)

The theory underpinning IRSs is that for organizations to improve their safety performance managers should be aware of events in their organization and employees feel confident to report errors and near misses without fear of recrimination.[3,24] Managers and employees can obtain data about the frequency and severity of incidents, benchmark their performance against other similar organizations, and identify systems' deficiencies to improve performance and provide insights into human factors in areas such as management, training and fatigue. Experts have argued that organizations can learn from these data, using this learning to alter structures and processes to reduce both the *actual* harm and the *potential* for harm.[3,25]

IRSs are credited with helping to improve substantially the safety of airline travel and it was therefore assumed that there would be valuable lessons for healthcare.[7,26,27] There are two aspects to an IRS: firstly, the reporting of 'adverse events' or 'patient safety incidents' -any unintended or unexpected incident(s) that led to harm for one or more persons[28(p2)]; secondly, the reporting of 'near misses'-any event(s) that did not cause harm but had the potential to do so.

At the micro level, that is the level of organization where agents interact and rules are adopted, maintained, changed or resisted in their local context,[29] Reason[30] argues that IRSs provide a systematic method to enhance ongoing learning from experience for the primary purpose of improving patient safety. Voluntary confidential reporting is thought to enhance understanding of the frequency of types of adverse events, near misses and their patterns and trends, hence acting as a warning system. This information should then be utilized at the micro-meso level of organization, that is individual actors plus the system of rules.[29] At this level changes to common rules should occur to enable system redesign to reduce the possibility of adverse events (re)occurring. NASA claims that aviation safety reporting assists the identification of training needs; provides evidence that interventions have been effective; and engenders a more open culture in which incidents or service failures can be reported and discussed.[24] At the macro organizational level, that is, a higher order of organization which arises from the existence of interacting populations of meso rules,[29] IRSs are considered to be an accurate early warning system for the identification of problems related to emerging technologies and global economic trends.[16,31]

Several authors contend that adverse events occur when active failures, that is the errors, omissions or unsafe acts by individuals, interact with latent conditions (underlying structures and processes) within an organization, to cause harm. Near miss events occur up to 300 times more often than adverse events.[32,33] Evidence suggests that within health care organizations pressure on front-line employees to increase efficiency has created a safety culture where deviance is normalized[34(pi69)] as employees attempt to cope with competing demands by fixing or working around problems at the local level, their actions hiding latent conditions, which increase susceptibility to error, and at the same time instilling them into the system.[30] In addition, public

inquiries into UK NHS failings have reported that the dominance of doctors in the occupational hierarchy, in combination with a culture of fear can prevent other groups from speaking up about safety.[35,36] Turner maintains that readjustment of such cultural norms will lead to reduction of errors.[37] Theories emphasize how IRSs are a trigger for culture change, promoting knowledge sharing by aggregating data collected at a local level to reveal and disseminate more widely those patterns of cause and effect (latent conditions) which increase susceptibility to the same types of errors occurring in differing contexts.[30,38]

An IRS should be a secure information resource accessible and responsive to users.[24] The safety literature contends that to promote its widespread acceptance and use: all stakeholders should be committed and actively involved in its development; there should be consensus among stakeholders over its design; the system should be objective, not under the control of one or more stakeholders; and it should be designed to facilitate the collection of narratives about incidents in the respective reporter's own words.[31,39] Evidence suggests that critical to the success of any IRS is the quality of the feedback given to reporters to enable learning, encourage reporting and provide reporters with evidence that the information they are providing is being utilized appropriately.[40,41]

Organizational learning (Theory)

As noted above, IRSs are regarded as a mechanism to promote organizational learning to improve patient safety; and there is particular interest in organizational learning that leads to cultural change. We now consider the available theory or theories of organizational learning.

Since Cyert and March[42] coined the term `organizational learning', scholarship has burgeoned, reflected in reviews of the field such as Easterby-Smith et al.[43]; Easterby-Smith and Lyles[44] and Shipton.[45] While it is broadly acknowledged that an organization's ability to learn

and adapt to changing circumstances is critical to its performance and long-term success,[46] competing theoretical perspectives on organizational learning exist. Rashman et al.[5(p471)] for example, cite Chiva and Alegre's[47] identification of two broad perspectives, `cognitive possession' and `social-process'.

With regard to the latter, authors such as Rashman et al.[5] and Waring & Bishop[48] argue that social, situated theories of organizational learning are highly relevant to public service and healthcare contexts. This type of theory regards organizational learning as complex and emergent, occurring through and embedded in social practices.[49] Healthcare organizations in particular are characterized by professional communities that span organizational boundaries,[5] involving multiple stakeholders in a complex inter-professional setting.

A social perspective on organizational learning also highlights the political dimension of knowledge and the way this can influence and impede it. Hence in healthcare, an IRS may be perceived as a managerial control mechanism, existing for the purpose of governance or (self-)surveillance, or as bound up with organizational and inter-professional politics and agendas. Powerful professional interests can be projected onto initiatives such as IRSs, thus seizing them as new territory on which existing battles can be fought. In healthcare organizations, knowledge forms the basis of professional power and jurisdictional control; what counts as knowledge is contested terrain[50,51] hence a source of conflict between the various clinical professions and between clinicians and managers. Therefore, it may be wrong to assume that clinicians are willing to share information about errors widely.[14] Doctors, particularly surgeons, are often reluctant to report incidents seeing IRSs as a managerial encroachment on their professional status and individual autonomy.[41,52–55] Research suggests they are more inclined to participate when the IRS is situated and managed within the medical department.[56] Within the UK NHS,

evidence indicates an underlying hostility and distrust between doctors and managers with doctors prioritizing professional learning over organizational learning, their non-cooperation undermining the implementation of the NRLS.[55] In addition, Waring[55] contends that doctors are reluctant to report incidents both for fear of litigation and because they consider errors part of the inherent uncertainty of medical practice. Therefore, rather than facilitate organizational learning, IRSs may decontextualize knowledge and act as a structure for organizational power by engendering conflict and competition for control over what counts as an error and hence what type of knowledge is legitimate.[2,14,57,58] Although accompanied by a rhetoric of learning, IRSs may instead be the product of normative and coercive isomorphic pressure,[59] a method of maintaining and/or restoring an organization's legitimacy.[16]

While recognising the merits of a social perspective on organizational learning for the way in which learning from IRS is likely to occur in healthcare settings, our specific need in this paper is for a theory of organizational learning that enables us to assess the evidence presented in the studies examined. Hence we have chosen the seminal work of Argyris and Schön[23,60] on single and double loop learning. Argyris and Schön's theory represents a primarily (if not exclusively) cognitive perspective towards organizational learning, according to Chiva and Alegre,[47] being concerned with the process by which learning leads either to the correction of errors within existing goals, policies and values, or to changes in those goals, policies and values.

The principal reason for choosing Argyris and Schön's theory, in preference to social theories of organizational learning, is that their distinction between single and double loop learning enables us to interrogate evidence, provided in the papers we have reviewed to classify the type of organizational learning produced by IRSs. In particular, it enables us to differentiate between operational improvements and possible examples of cultural change. This is important

given the emphasis in the literature on the role of IRSs in changing patient safety culture. There are, nevertheless, potential limitations to using Argyris and Schön's theory, to which we return in the Discussion.

To describe Argyris and Schön's theory in more detail, it proposes two principal forms of organizational learning. `Single loop learning' refers to the correction of operational errors, without significant change in the overall safety culture and `double loop learning', the questioning and alteration of what Argyis and Schön called `governing variables'. Thus: 'When the error detected and corrected permits the organization to carry on its present policies or achieve its present objectives, then that error-and-correction process is single-loop learning. Double-loop learning occurs when error is detected and corrected in ways that involve the modification of an organization's underlying norms, policies and objectives.' [23(pp2-3)] To achieve cultural change, IRSs would appear to need to produce double, not single loop learning, equating to a shift in safety culture and `mindset' and resulting in a significantly different approach to the treatment of errors in healthcare organizations.

Argyris and Schön's theory[23] also identifies barriers to double loop learning in practice, which (they argue) make it more likely that organizations will undertake single loop learning. In particular, double loop learning is impeded by defensive behaviour that guards people against embarrassment and `exposure to blame'.[23(p40)] In relation to IRSs, defensive behaviour could lead not only to the non-reporting of errors, but also to the non-reporting itself being covered up. Hence Edmondson, based on Argyris' observation that `people tend to act in ways that inhibit learning when they face the potential for threat or embarrassment[61]', [62(p352)] argues that to achieve double loop learning in practice requires, a climate of sufficient psychological safety[62] to mitigate the propensity for defensive behaviour.

Argyris and Schön's theory therefore also enables review of evidence of potential barriers to double loop learning (the desired cultural change) in the studies examined. For example, fear of blame or reprisals, and the fact that `health care workers of all kinds are exposed to an inordinate amount of intimidating behavior',[11(p464)] would therefore appear incompatible with the requirement for sufficient psychological safety. Similarly, trying to enforce incident reporting through coercion (such as the threat of legal action) also seems likely to reinforce defensive behaviour.

In summary, in order to examine the relationship between IRSs and organizational learning a theory is necessary. While social theories of organizational learning acknowledge the complexity of healthcare contexts, the specific purpose of this paper led us to employ Argyris and Schön's theory of single and double loop learning to interrogate evidence of the type of organizational learning indicated by the studies in our review. As we have discussed, there are a number of reasons why IRSs may be problematic. Nonetheless, there has been no systematic review integrating the studies exploring the effectiveness of IRSs in the healthcare context.[12,33] The aim of this paper is to analyze and synthesize empirical evidence relating to the effectiveness of IRSs as a method of improving patient safety via organizational learning.

Method

Search strategy

Our search strategy was designed to find empirical studies about the effectiveness of IRSs as a method to improve patient safety. The search period was from January 1999, the year "To Err is Human"[7] was published, to March 2014. As indicated in Figure 1, we searched key healthcare journals, organization-based websites related to patient safety and online search engines. The search terms applied in all cases were: 'adverse event* reporting'; 'clinical incident *reporting';

‘incident reporting* safety’; ‘reporting medical errors’; ‘Reporting and Learning System(s)’; ‘Advanced Incident Monitoring System’; ‘Patient Safety Reporting System(s)’, ‘National Learning and Reporting Systems’; errors and organi?ational learning’; ‘Datix and organi?ational learning’; 'clinical incident analysis'; 'root cause analysis'; 'failure mode and effects analysis'; and ‘safer surgery checklist’.

We hand searched 11 key healthcare journals including Milbank Quarterly, Social Science and Medicine, Quality and Safety in Health Care, International Journal of Health Care Quality Assurance, Health, New England Journal of Medicine, Journal of the American Medical Association, British Medical Journal, Medical Journal of Australia, the Canadian Medical Association Journal and the New Zealand Medical Journal.

We also included in our search organization-based websites related to patient safety including Agency for Healthcare Research and Quality, The Joint Commission on Accreditation of Healthcare Organizations/ International Centre for Patient Safety, The National Patient Safety Agency (UK), The National Patients Safety Foundation (USA), The Health Foundation (UK), The Australian Patient Safety Foundation, Canadian Patient Safety Institute, the Scottish Patient Safety Programme, Health Quality and Safety Commission New Zealand, NASA and the WHO. Finally, we searched systematically for articles in PUBMED, ISI Web of Knowledge and Google Scholar.

Abstracts were obtained based on judgments about the content of each article using the title and key words. Two of the authors reviewed the abstracts independently, cross-referencing judgements on the papers. For abstracts to be included, they had to provide empirical data either on comparisons with other methods of incident reporting or in relation to changes to settings, processes or outcomes as a consequence of knowledge gained initially through information

derived from an IRS. We excluded systematic literature reviews. Where there was disagreement between reviewers, abstracts were included. Having agreed on the abstracts for inclusion, duplicates were removed and full papers retrieved. Following this we read, reread, and discussed the papers, again excluding those that did not meet the aims of our study. Finally, we hand searched the references of each full paper retrieved for titles and key words that included our search terms to identify further papers than may have been omitted by the search to date.

[Insert Figure 1 here]

Studies were limited to those published in English with no restrictions on the basis of country of origin or the context in which studies were undertaken. We included only empirical papers that sought to examine how effective IRSs are for patient safety, either by comparing them to other systems or by looking at improvements in structure, settings or outcomes according to Donabedian's[6] framework described below. We excluded opinion papers, systematic literature reviews and studies that analyzed the effectiveness of IRSs as a method to capture the number and type of near-miss and patient harm events. Barriers preventing clinicians from reporting incidents were beyond the scope of this paper.

Many of the studies on the final list were descriptive involving retrospective analysis of quality improvement work within single departments or national coordinating organizations. As explorations of a service quality intervention, they do not necessarily follow orthodox qualitative and quantitative research designs. Therefore, following Pawson et al.'s[63] argument that the value of such studies is demonstrated in synthesis, we took a pragmatic decision to include papers based on their relevance, that is if they addressed our research question using the data extraction process (see appendixes A and B), rather than by assessing the quality of the selected articles using a standard checklist.

Data Extraction

We identified two groups of studies: a) studies comparing the effectiveness of IRSs relative to other methods of error reporting such as medical chart review; b) studies that aimed to measure effectiveness in absolute terms. Parallel data analysis was carried out to address these different but related aspects of the effectiveness of IRSs as a method of improving safety.[64]

Appendix A summarizes studies comparing the effectiveness of IRSs relative to other methods of error reporting. Data for these studies were extracted by comparing and contrasting the various methods of reporting and their outcomes within and across the respective studies.

Appendix B summarizes studies that measured effectiveness in absolute terms. In respect of this second group of studies, we acknowledge that the measurement of effectiveness can be both complex and challenging.[65] Thus to ensure transparency data from these studies were extracted using Donabedian's[6] settings (structure), processes and outcomes framework. Hence we define 'effectiveness' in absolute terms as the following outcome types:

1) Changes made to the *setting* in which the process of care takes place, which refers to the structures that support the delivery of care.
2) Changes made to the *process* of care, which is to how care is delivered.
3) Effects of changes to settings and or process for the *outcomes* of care, in this case for the specific area of patient safety.

Donabedian[6] acknowledged that each approach has its own limitations. Outcomes are often difficult to measure and may be influenced by factors other than clinical care. Processes of care on the other hand are not as stable as outcome indicators. Furthermore, it is difficult to make causal links between settings, processes and outcomes: 'outcomes, by and large, remain the ultimate validators of the effectiveness and quality of medical care.'[6(p693)]

Data Synthesis

Having extracted the data from both data sets, we used an interpretative and integrative approach to evidence synthesis. This involved combining a summary of the data showing which types of changes to practice were made, with an interpretation of the data grounded in assumptions about how IRSs should work.[66]

In the first group of studies each paper was read initially to identify the comparative methods employed and their relative advantages/limitations. Then we compared and contrasted the studies to identify similarities, patterns and contradictions -a recursive process, which involved reading and rereading individual articles and moving back and forth between articles. For the second group of studies, having firstly extracted and tabulated (Table B) how adverse incidents were conceptualized; the types of changes made to practice; and whether these involved settings, processes or outcomes; we searched each article for evidence of improvements to patient safety as a result of the changes implemented. Following this, we used thematic analysis, considered as a suitable method of organizing and summarizing the findings from both qualitative and quantitative research,[64,67] to identify systematically across the studies the main themes in respect of what was effective (or ineffective), where (context) and why.

Findings

Descriptive analysis of studies

In total 43 studies were included in our analysis. The majority were conducted in the US (16), followed by the UK (14), Australia (4), Canada (3), France (1), the Netherlands (1), Denmark (1), India (1), Switzerland (1) and Japan (1).

The context for the studies varied. Most (29) took place at the micro-meso level. Of those, 15 were in general hospitals and 9 were in specialized units (three oncology departments, one pediatric unit, one obstetric unit, one hospital based transfusion service, one eye hospital, one psychiatric division of a teaching hospital and one tertiary cancer center). Of the remaining five studies, two were conducted in an intensive care unit (ICU), of which one was general and one neonatal-pediatric; two studies involved nursing homes; and the final study took place in a medium secure unit.

The other 14 of the 43 studies were at the macro level. Of these, nine investigated incidents reported in large scale reporting programs including the UK's NRLS and IRSs in NHS Scotland, the AIMS and the Food and Drugs Administration (FDA) in the US; one explored the US Veterans Health Administration; three referred to hemovigilance reporting programs at the macro level, including the Serious Hazard of Transfusion (SHOT); and one examined pharmacies.

Multiple Definitions

The analysis highlights that the studies used a wide variety of terms to describe adverse events. These terms included: clinical incidents, adverse reactions, adverse outcomes, adverse events, potential adverse events, adverse incidents, adverse drug reactions, errors, medical errors, drug errors, events, near-misses, medication errors, reviewable sentinel events. One study[68] used various terms including: clinical incident, clinical error, critical incident, adverse event or adverse incident. Weissman et al.[69] analyzed data from four hospitals all of which used different terminology for adverse events.

Of the 43 studies, 26 were considered to have provided clear definitions of what was considered as an adverse event Nine failed to provide any definition[40,70–77]; five used

classifications rather than definitions to categorize incidents[78–82]; and three acknowledged the difficulties of definition and raised the need for more conceptual clarity.[57,65,83]

Examples of the approaches taken by those studies that did provide definitions include Percarpio and Watts[84] who, using the Joint Commission's definitions, distinguished between "an adverse outcome that is primarily related to the natural course of the patient's illness or underlying condition" and a "reviewable sentinel event," which is "a death or major permanent loss of function that is associated with the treatment (including 'recognized complications') or lack of treatment of that condition, or otherwise not clearly and primarily related to the natural course of the patient's illness or underlying condition".[84(p35)] Sari et al.[85] and Wong et al.[68] used very broad definitions, which described adverse events as unplanned events with the potential to cause harm or undesired outcomes to patients. Some studies were more precise as to the type of outcome an adverse event can cause, the timing at which the adverse event can take place and who experiences it. Marang-van de Mheen et al.[86] specified that an adverse event can happen during or following medical care and can be noted during the treatment or after discharge or transfer to another department. The outcomes of an incident almost always included disability or death, but also prolonged hospitalization.[87–89] Cooke et al.[90] defined an adverse event as any impairment in quality, efficiency or effectiveness of the patient care system. Only one study discussed damage or loss of equipment or property and one study discussed incidents of violence, aggression and self-harm.[91] All definitions talked about harm to the patient, with only three extending their definitions to include a staff member[88–90] and one to include a visitor.[88]

The definitions of medication errors were more exact, although again these varied between studies. Jayaram et al.[92] and Zhan et al.[22] used similar definitions, Zhan et al.'s being the more precise: "any preventable event that may cause or lead to inappropriate medication use or

patient harm while the medication is in the control of a healthcare professional, patient or consumer".[22(p37)] Boyle et al.[93] specified the type of error, including incorrect drug quality, dose or patient.

Adverse events in blood transfusion were generally well defined. In the UK, in addition to detailing categories of adverse events reportable to, and monitored by, the SHOT scheme, Stainsby et al.[94,95] listed non-reportable events, such as reactions to plasma products. This scheme is professionally led and affiliated to the Royal College of Pathologists.[94] In the US, Askeland et al.[96] gave a very detailed description of the categories of adverse events that can occur during the 'blood product history'. Similarly, Callum et al.[78] described causal codes, to classify latent and active failures and patient related factors in a Canadian hospital and, Rebibo et al.[97] defined 'hemovigilance' describing how the national system for surveillance and alert in France operated at each organizational level.

In conclusion, the variability in terminology and definitions suggests that assessing the effectiveness of IRSs may be hampered by problems of conceptual clarity and comparability of studies. We address these implications further in the Discussion. In the following sections we compare IRSs with other systems before going on to explore the effectiveness of IRSs in absolute terms.

Studies comparing IRSs with other systems

Of the 43 studies, eight compared IRSs with other reporting methods, while 35 examined the effectiveness of IRSs themselves.

To begin with the studies that compared IRSs with other systems (presented in Appendix A), four of these eight compared IRSs with retrospective medical chart review.[83,85–87] In a study by Beckman et al., conducted in an ICU in Australia,[87] senior intensive care clinicians

encouraged staff to write incident reports, using the established IRS, by discussing incident monitoring at ward rounds and similar clinical sessions. The IRS identified a larger number of preventable incidents, provided richer contextual information about them and required significantly fewer resources than the retrospective medical chart review. There were qualitative differences in the type of adverse events highlighted by the two forms of reporting. Equipment problems and adverse events related to the retrieval team were only reported by the IRS. The authors speculated that staff believed that the patient's medical record was not the correct place for reporting such problems. The IRS identified near misses, the medical chart review did not. Unplanned readmissions were deemed to be due to adverse events in only three cases in the medical chart review, whereas the IRS detected six. Medical chart review identified incidents such as iatrogenic infections and unrelieved pain, which were not identified by the IRS. Additionally, medical chart review found evidence of patient's breathing problems not found in the IRS, possibly because they did not lead to an obvious adverse event such as increased length of stay in the ICU. Beckman et al. argued that both the IRS and medical chart review were able to identify problems of patient safety in intensive care responsive to actions to improve the quality of care, but they did not provide evidence of changes to process or outcomes.

Sari et al.[85] compared an IRS with retrospective medical chart review in an English NHS hospital. They found that the medical records had documented cases of unplanned transfers to ICU, unplanned return of patients to the operating theatre, inappropriate self-discharge and unplanned readmission. Not one of these cases was reported in the IRS, indicating under reporting.

Similarly, Stanhope et al.[83] examined the reliability of IRSs in two obstetric units in London, concluding that although IRSs can provide useful information, they may seriously underestimate the overall numbers of incidents.

Marang-van de Mheen et al.[86] compared the incident reports of clinically occurring adverse events gathered by surgeons and discussed at their weekly specialty meeting, with retrospective medical chart review, in a sample of high risk surgical patients in a Dutch hospital. They found adverse events were missed by both the IRS and medical chart review again suggesting under reporting. Medical chart review identified significantly more adverse events overall than routine reporting, supporting the findings of Sari et al.[85] However, the IRS identified serious adverse events that were missed by medical chart review. Adverse events occurring after discharge or ward/hospital transfer were not identified by the IRS. Marang-van de Mheen et al.[86] argued that when incident reporting was under the control of the clinicians and supported by discussion at regular peer-led meetings it had distinct advantages in comparison with macro-level quality improvement initiatives such as National Confidential Enquiry into Patient Outcome and Death reports. The authors maintained that local ownership of the data gave clinicians the opportunity to study adverse events within their specialty; responsibility for implementing recommendations; and longitudinal data to study trends and monitor the effectiveness of changes to practice. The studies by both Beckman et al.[87] and Marang-van de Mheen et al.[86] similarly highlight the importance of ownership of the IRS at the micro level for individual and departmental learning.

Of the other four studies that compared IRSs with other systems,[76,93,98,99] Olsen et al.[98] compared three different methods of detecting drug related adverse events in an English NHS hospital: the IRS; active surveillance of prescription charts by pharmacists; and medical chart

review. Similar to Beckmann et al.,[87] Stanhope et al.[83] and Marang-van de Mheen et al.[86] they found that the IRS provided a less acceptable indication of clinical adverse events relative to the two other methods, concluding that the IRS was effective only when supplemented with other data collection. Flynn et al.[99] also compared three methods for detecting medication errors: an IRS; medical chart review; and direct observation. Direct observation involved researchers observing nurses administering 50 prescriptions during the morning medication administration round. Observers, including nurses and pharmacy technicians, were paid to collect the data. The study concluded that direct observation was the most efficient and accurate of the three methods. However, similar to the other studies cited it gave no indication of the relative resources involved.

The third study, by Wagner et al.,[76] tested the effectiveness of a computerized falls IRS providing a standardized structure and consistency for which items to include in the report, comparing this with a semi-structured open-ended description type of report often used in US nursing homes . Findings suggested that there was more complete documentation of the post-fall evaluation process in the medical records in nursing homes using the computerized IRS than in nursing homes using non-standardized descriptive type of reporting. Similarly, Boyle et al.[93] assessed manual versus computerized IRSs in pharmacies in Canada; pharmacists reported that both computerized and manual incident reporting were cost-effective and easy to complete. However, pharmacists using computerized reporting systems assessed their utility higher than those working with manual systems.

To summarise, the eight studies that compared IRSs with other reporting methods show no firm evidence that an IRS performs better than any other method of reporting.

Studies Examining the Effectiveness of IRSs

We turn now to the remaining 35 (of the total of 43) studies that examined the impact of IRSs themselves on settings, processes and outcomes (summarized in Appendix B). The micro-meso changes reported in these studies were of three types: a) changes to policies, guidelines and documentation , b) provision of staff training, and c) implementation of technology, discussed in turn below. Following these, we summarise macro-level impacts that were reported in nine of the 35 studies. Finally, we present our analysis through the lens of organizational learning theory. .

Changes to Policies, Guidelines and Documentation

Frey et al[100] reported on changes to drug administration in a Swiss neonatal ICU, including the introduction of a standardized prescription form, compulsory double-checking for a list of specified drugs and new labelling of infusion syringes. However, no evaluation of the effectiveness of the changes for patient safety was reported. Anderson et al.[40] reported that many policy changes had been introduced in both an acute and in a mental health hospital in London. Again no evaluation of the impact on safety was provided. Few frontline clinicians participated in this study because they did not have knowledge about incident reporting and were often not consulted about the feasibility and potential benefits of recommended solutions. This suggests that the IRS in both these hospitals had limited effectiveness at the micro level.

Wong et al.[68] described 15 changes to practice directly resulting from data specific to vitreoretinal patient safety incident reports at the Moorfields Eye Hospital, England, concluding that these changes had improved patient safety. Grant et al.[80] examined patterns of adverse events in an Australian hospital using data from an electronic record-keeping system. Two problematic areas were identified: sedation for colonoscopy and inhalational anesthesia with desflurane. Using this information anesthetists developed specific departmental guidelines for

these procedures. Subsequently, adverse events during these two procedures were significantly reduced. Ross et al.[72] reported a reduction of medication errors from 9.8 to 6 per year when, highlighted by the IRS, dispensing checking by two people was initiated. An IRS was implemented in a surgical unit at The Johns Hopkins Hospital in the US in 2007 with 'good catch' awards being made to staff that reported and helped to prevent safety hazards.[101] At the time of publication in 2012 the authors noted that quality improvements associated with 25 of the 29 "Good Catch" awards had been sustained. The changes described included the removal of high-concentration heparin vials and daily equipment checks. The authors did not measure directly the impact of the IRS on safety culture, noting that the project coincided with several other quality improvement initiatives, hence they were unable to attribute changes in safety culture to any one initiative.

Wolff et al.[88] reported a reduction in the number of falls resulting in fractures following implementation of falls risk assessments, after the IRS identified falls as the most common adverse events. This was a cross-sectional study therefore sustainability was not measured. Hospital acquired hip fractures still result in poor outcomes such as increased mortality and doubling of the mean length of patient stay and mean cost of admission[102,103] suggesting that IRSs have made little impact on patient safety in respect of falls.

Checklists and time-outs for delivering radiation therapy were implemented in a Chicago hospital's Department of Oncology in response to errors related to wrong site or wrong patient.[104] Daily pre-treatment timeouts had to be accomplished by at least two therapists in the treatment room before delivering treatment to the patient, followed by post treatment planning timeouts completed by physicians. The checklists included reviews of treatment parameters before each treatment step. The authors reported that the use of these relatively simple measures

significantly reduced error rates related to wrong treatment site, wrong patient and wrong dose in patients receiving radiation therapy.[104]

In a medium secure hospital in Wales, by analyzing data from an IRS, Sullivan and Ghroum[91] identified the peak periods for adverse events involving violence, aggression and self-harm. An improvement plan was implemented, this included flexible patterns of staffing and the introduction of therapeutic treatment groups. As a consequence, the authors reported a significant reduction in reported adverse events over a two-year period. The context for this study is relatively unusual in that staff are often the recipients of violence and such adverse events are highly visible, hence employees may be more motivated to learn from them.

Provision of staff training

In a number of studies data from the IRSs identified the need for staff training. In some cases training was introduced to raise awareness of risks and establish a culture of safety,[92] in others to improve clinicians' skills. For example training for nurses to improve their ability to administer drugs[72]; safe prescribing teaching sessions for residents[100]; training to improve clinicians' recognition of mental health issues in young people[81]; education on preventing incompatible blood transfusions[78]; and training for staff on how to improve communication of adverse events to their supervisors and for supervisors on how to give feedback from adverse events to support and encourage learning.[90]

Callum et al.[78] showed that educational sessions on preventing ABO-incompatible transfusions were ineffective, the rate of adverse events remaining unchanged. Similarly, Cooke et al.[90] found no evidence that training improved processes of care or outcomes. Indeed, most respondents believed that the incidents they reported were not investigated. The findings by

Cooke et al.[90] and Anderson et al.[40] suggest disconnection between the micro and the meso levels of organization.

The impact of training on improving the actual process of care and ultimately improving outcomes was not reported in many of the studies, despite this being one of the quality improvement methods used.[68,73,79,81,92,94,95,100,105] Indeed, the only study reporting evidence of a direct impact from training was that by Ross et al.[72] in a UK pediatric hospital, who showed that training provided to all nurses administering intravenous (IV) drugs resulted in a reduction of errors. However, it should be noted that this occurred when nurses were beginning to take over IV drug administration from doctors and the authors noted that 'nurses are increasingly responsible for giving all medications, precisely because they have better error trapping systems in place'.[72(p495)]

Implementation of Technology

Implementation of technology was the third commonly documented change to practice in the studies we reviewed. Askeland et al.[96] reported on the introduction of bar code technology throughout the blood transfusion process in a US hospital to assist in the prevention of transfusion errors. They found that the bar code system was considered three times safer than the old manual system. Callum et al.[78] described the implementation of an IRS for transfusion medicine in a Canadian teaching hospital. Information from the system was forwarded to the Canadian Blood Services who established implementation and expiration date labeling as priorities. Callum et al.[78] argued that this would reduce the errors associated with labeling of the expiration date, but no actual evidence was provided. In addition, they implemented a trial mandating labeling at the bedside via a system using wristband barcodes, portable handheld data terminals and printers to allow easy bedside labeling. The authors reported an improvement in

blood group determination and antibody screens in the emergency room as a result. A new requisition form was also introduced on which the area to sign was delineated by a thick black box and written above the box in big letters 'please read and sign'. However, they argued that this change did not provide sufficient reinforcement suggesting the need to evaluate an electronic signature as a 'forcing function'[78(p1209)] to eliminate this type of error.

Ford et al[106] reported that in the John Hopkins Department of Radiation Oncology a change was implemented such that the medical physicists "hid" the treatment fields which were not in use for respective patients, thereby eliminating human error. Following implementation, not one out-of-sequence treatment was reported.

Finally, a significant reduction in reported prescribing errors was found by Jayaram et al.[92] following the introduction of an electronic system allowing pharmacists immediately to page any doctor who entered an incorrect order so that it could be remedied.

In conclusion, three types of micro-meso level changes ensuing from IRSs were reported in these 35 studies. All four instances of the implementation of technology were reported as being successful. However, studies did not always evaluate the effectiveness of the changes reported for patient safety outcomes; for example only one out of 12 studies that reported provision of training did so. Few studies reported the outcomes of IRSs, therefore, evidence of the effectiveness of the changes ensuing from IRSs remains partial.

Macro level changes

Nine of the 35 studies reported on changes to practice at the macro level. Roughead et al.[107] analyzed the case of the antibiotic flucloxacilin in Australia. Data from the Drug Reaction Advisory Committee had raised awareness amongst health professionals of the adverse hepatic reaction associated with the use of flucloxacillin, resulting in a significant decrease in its use.

Wysowski and Swartz[108] analyzed all reports of suspected adverse drug reactions submitted to the FDA from 1969 to 2002. During this period, numerous drug reactions were identified and added to the product labeling as warnings, precautions, contraindications, and adverse reactions. Further, 75 drug products were removed from the market due to safety concerns and 11 had special requirements for prescription or restricted distribution programs.

Similarly two guidelines, one on the management of a suspected transfusion transmitted bacterial contamination and one on the process of transfusion in France, were published in 2003.[97] The authors reported that incompatible ABO transfusions were reduced between 2002 and 2003 and misdiagnosis of adverse blood transfusion events were better identified and investigated.

Zhan et al.[22] analyzed voluntary reports of errors related to the use of warfarin in a large number of hospitals in the US from 2002 to 2004 and mention a number of changes in patient care including increased monitoring and alterations to protocols. They did not state if such changes reduced errors. Grissinger et al.[74] analyzed errors involving heparin from data aggregated from three large IRSs. The three programs used different terms to categorize the areas where errors occurred, complicating the aggregation of this information at the macro-level. This cross-sectional study identified significant harm caused by heparin but it did not explore whether organizations learned from the IRSs and if this resulted in reduced levels of harm. The authors found common patterns of events across all three IRSs, arguing that in the case of common events such as medication errors, additional learning about the origination and causes of errors can be obtained only if incident reports provide rich qualitative data on the event and the context in which it occurred, rather than aggregating quantitative data.

Spigelman and Swan[73] surveyed 12 organizational users of the AIMS. Respondents reported numerous settings and process changes including equipment standardization, new standards for medication prescribing and administration, and staffing level improvements. The authors noted a poor level of reporting by medical staff; that improvements in outcomes, as a result of changes implemented were difficult to ascertain; and if the AIMS was to show outcome improvements to patient safety the level of resources required were not to be underestimated.

In the UK, Hutchinson et al[109] argued that the NPSA provided hospitals with feedback, which enabled them to benchmark their performance against other similar hospitals. However, improvements to processes and outcomes at the meso level, arising from aggregation of data at the macro level of the NPSA were not reported.

Conlon et al.[70] analyzed the IRS introduced from 2001 in 36 Trinity hospitals and affiliates in the US. Numerous changes to practice as a consequence of learning from IRS data are described. The authors conceded that it was difficult to attribute improvements solely to the IRS as the organization employed various improvement efforts at any given time. However, it had achieved a 26 percent decrease in severity-adjusted mortality rates since January 2005 and a reduction in liability costs following the implementation of the IRS. Overall, there is some evidence of effectiveness for improving patient safety at the macro level.

Organizational Learning (Analysis)

We then applied Argyris and Schön's definitions of single and double loop learning to the second group of (35) studies to assess the extent to which there is evidence that IRSs prompted any of the two types of learning. This was an interpretive process that entailed debate about how to apply Argyris and Schön's theory rigourously and consistently. In essence, we focused on whether evidence was of technical and operational improvements (single loop learning), or of

changes in governing variables (double loop learning). The detailed results are shown in Appendix C.

First, we observe that the evidence presented by 33 of the 35 studies could be classified as single loop learning, such as direct improvements to procedures. Examples include a new bar code system leading to correction of errors and improvement in patient safety[96]; a variety of changes including new labelling[100,105]; and the implementation of new blood transfusion guidelines[97]. Furthermore, there were reasons why the remaining (2) studies did not contain such evidence: one study analyzed the causes of errors but did not report actions taken[74]; and the other PSA[81] was concerned with making recommendations towards improving patient safety.

Turning to double loop learning, based on our review we consider there to be little conclusive or convincing evidence within the studies analysed that shows IRSs leading to changes in governing variables. As noted earlier, the absence of such evidence does not necessarily mean that IRSs are ineffective in this respect. There are several alternative explanations for this lack of evidence. First, with some studies it could be inferred that an effective safety culture already exists[105,106]; if so, double loop learning would effectively be redundant. Second, with the exception of Aagard et al.,[57] Cooke et al.,[90] NHS QI[89] and Nicolini et al.,[77] the studies reviewed made little explicit use of organizational learning theory and lacked theoretically-informed conceptualisations of cultural change. In the absence of a theoretical framework, such studies may inevitably struggle to capture convincing evidence of cultural shifts in patient safety. Third, where studies have confined themselves to investigating outcomes that ensue directly and immediately from an IRS, they may have failed to capture the more indirect and diffuse learning that social theories of organizational learning suggest could be present.

Even given these reasons, it is an important finding that the studies reviewed are more

successful at producing evidence of single than of double loop learning.

Ten of the studies contain claims that *could* correspond to double loop learning. The most detailed description of organizational learning that appears compatible with double loop learning is that by Conlon et al.,[70] who state that `A systemwide council of PEERs Coordinators meets regularly to share lessons learned and best practices related to patient safety. This information is routinely shared with management. The PEERs system nurtures a blame-free environment where reporting is encouraged',[70(p1)] and `The PEERs system has become part of the culture within Trinity Health… This leads to a common understanding and helps to foster a consistent culture within Trinity Health'.[70(p12)]

In most other instances, the studies infer that an improved safety culture has been achieved; for example, `Conceptual changes included changes in risk perceptions and awareness of the importance of good practice'[40(p148)]; `the belief that some changes are contributing to an `enhanced "safety culture"'[78(p1209)]; `a focused, hospital-wide effort to improve the system of medication preparation, processing, and delivery'[80(p217)]; indicators of a positive safety culture[109]; creating a safety culture through a multi-disciplinary effort involving combination of interventions[92]; `Changing the error reporting form to make it less punitive'[72(p492);] `Developing an awareness of error and a safety culture with less emphasis on the "blame" approach'[73(Table 2,p658)]; and that `successive SHOT reports have encouraged open reporting of adverse events and near-misses in a supportive, learning culture'[94(p.281)].

Not one of these studies, however, contains sufficient information about the action taken towards organizational learning, or sufficient evidence about the consequences of such action, to conclude that double loop learning resulted from an IRS.

What studies do indicate, however, are potential facilitators of organizational learning and/or potential barriers, in the absence of such facilitators (see Table 1 and Appendix C).[40,77]

First, it was noted above that, according to Argyris and Schön's theory, psychological safety is likely to be important for double loop learning. There is regular and repeated reference across more than half (18) of the 35 studies to the need to make reporting less punitive, to the benefits of anonymous, confidential reporting, and to the absence of a `blame culture' or fear of reprisals. Where studies have used medical definitions of error, this may have contributed to the research agenda focusing on the micro level, thereby implicitly blaming individuals.

Second, the emphasis on learning needs to be genuine, rather than rhetorical or espoused. Four studies[68,71,77,89] raise awareness of the need for learning to be the function or output of an IRS. This is contrasted with IRSs being driven by an `audit culture' where its agenda may be (perceived to be) the reassertion of management control; and with the possibility that an IRS exists (or is perceived to) for the purpose of surveillance.

Third, although rarely adopting a social perspective on organizational learning, many studies drew attention to the complex, emergent nature of it. The review did not find one paper that examined explicitly the effectiveness of IRSs for identifying latent error promoting organizational (managerial) factors such as decisions about resource allocation. Yet, it is the accumulation of dysfunctional organisational processes which eventually result in adverse events.[110] An important point made by a number of studies is that IRSs are most effective when part of wider quality improvement programs.[70,75,104,107] Being embedded within, or linked to, organization-wide interventions may offer one way to overcome the difficulty of achieving organizational learning in a complex multi-professional setting. Several studies refer to the need

for an IRS to be cross-departmental, multi-professional or inter-organizational.[70,71,90,92,94,97,101,105] Others highlight the way that multiple interventions are more likely to be effective than single interventions.[72,78,92,107] Thus Callum et al.[78] comment on the ineffectiveness of small-group educational sessions if used in isolation; and Ross et al.[72] highlight the need for an intervention to be complemented by other changes. Finally, some studies emphasised the benefits of an IRS being locally designed, and/or enabling the participation of staff who are directly concerned with patient care in that setting.[70,77,90,101]

Discussion

We conducted a parallel review of studies comparing IRSs with other forms of reporting and of studies designed to measure the effectiveness of IRSs in absolute terms, with the aim of exploring whether IRSs improve patient safety through organizational learning.

The analysis of the former group of studies showed no strong evidence that IRSs perform better than other methods. Indeed, medical chart review may have greater effectiveness in identifying clinical incidents than IRSs. What is more, there was very little focus on resource utilization with only two studies looking at this issue.[87,93] Therefore, there is no clear evidence that IRSs are more cost -effective than other systems.

The analysis of the second group of studies looked for evidence of changes implemented as a consequence of information gained by IRSs on settings, processes and outcomes, using Donabedian's[6] framework. At the macro level of organization we found evidence that IRSs could trigger single loop learning primarily in the context of drug prescribing by action forcing changes such as withdrawal of certain medicines from the market. There was also some limited evidence

of changes to processes and outcomes at the micro-meso level triggered by dissemination of IRSs data on adverse events arising from blood transfusion and use of flucloxacillin.

At the micro-meso level of organization there were few studies that reported on outcomes and those that did acknowledged the difficulty of demonstrating a causal relationship between IRSs and safety improvements, as IRSs were often part of a wider program of safety improvement.[70,101,104,107] Further, our synthesis supports Waring's[14] argument that centralized systems, at the micro-meso level, such as those used within UK hospitals, might not yield the depth of learning anticipated by policy-makers. Consistent with this, our review indicates that meso level changes may have little impact at the micro level. While, at the intra organizational micro-meso level, where there is ownership of incidents and clinical commitment to safety improvement, changes to settings and processes can be implemented successfully using learning from IRSs. The imposition of changes generated at the organizational level violates norms of collegiality and self-regulation and creates distrust of managerial motives.[14] Our synthesis suggests that IRSs are at their most effective when used and owned by clinical teams or communities of practice[111] within specific departments rather than at the wider organization level. Such communities have been shown to be nurtured by opportunities for interaction and communication[111] and are likely arenas for the development of reciprocal ties, shared commitment to group goals, trust and psychological safety required for organizational learning.[111]

Notably, the absence of standard, agreed universal definitions for adverse events or near misses and lack of clear definitions and measurement of outcomes makes it difficult to compare, identify and correct errors, or to evaluate the impact of doing so, reliably. Without clear definition of what counts as an adverse event, assessing the effectiveness of IRSs is problematic.

Our analysis showed that when definitions were clear, such as in studies of blood transfusion and macro level drug reporting, IRSs were more likely to improve safety. In contrast, although anticoagulation is an area of high risk, IRSs relating to anticoagulant therapy did not have agreed definitions of harm hence aggregation of information from various data bases was problematic. Another factor impeding organizational learning was the absence of a feedback loop; staff not always receiving feedback about incidents reported.[13,40]

Our review identified both potential facilitators of and barriers to double loop learning and indicates that to achieve it, an IRS needs to satisfy certain conditions. Reported incidents should be regarded as errors resulting from wider, potentially complex settings and processes; rather than narrowly focused on clinical practice or on `solvable' errors. To deal with such complexity, an IRS needs to work across functional, organizational and professional boundaries; be contextually located and participative, rather than imposed and managed hierarchically. IRSs should be tailored to local conditions to create a sense of ownership and involvement in efforts towards organizational learning. Resulting action is likely to require multiple, complementary interventions. Studies indicate that interventions used in isolation (for example training) are unlikely to be effective. Employees need to have confidence that `learning' is the authentic purpose and raison d'etre for an IRS; as distinct from the perception that an IRS exists for procedural purposes, or as a managerial instrument for the purpose of surveillance. Hence, a more effective method might be the development of IRSs at the micro-meso department level, provided they retained the main principles.[112] This finding concurs with the principle from organizational learning theory that the processes through which double loop learning occurs are multifaceted, emergent and embedded in social practices.

Study Limitations

Our review has relied mainly on formal research in academic journals; therefore, and although we searched a range of relevant organization data bases, we may have missed some evidence of effectiveness of IRSs within organizations that has not been subjected to empirical investigation and reporting.

The choice of Argyris and Schön's theory means that we have adopted a cognitive rather than social perspective on organizational learning. We have acknowledged that social theories of organizational learning may account for the way organizational learning is likely to emerge through complex processes that involve multiple actors and multiple agencies, therefore this is an area of potential for future research. Nevertheless, we would maintain that Argyris and Schön's theory is fit for purpose given the aims of our paper.

Conclusions

Overall, the studies reviewed did show some evidence that IRSs can lead to single loop learning, that is, corrections to errors in procedures and improvements in techniques. However, we found little evidence that IRSs ultimately improve patient safety outcomes or that single loop learning changes were sustained, although this may be a consequence of measurement difficulties[65,109] and the need for agreed definitions for both adverse events and the types of incident that should be reported. An important point made by a number of studies is that at the micro-meso level of organization, IRSs are most effective when combined with other improvement efforts as part of wider quality improvement programs, supporting an argument that 'reporting systems should complement, not replace practices used by hospitals to review and analyze their health safety incidents'.[65(p3)] Our review found little evidence of IRSs leading to double loop learning, that is, cultural change or change of mindset.

In sum, one way of improving both the efficiency and effectiveness of IRSs might be by embedding them as part of wider safety programs and devolving their control and management from centralized hospital departments to clinical teams. Results of our study suggest that healthcare organizations should consider carefully the opportunity costs involved in IRSs and whether they provide value for money. Further work on the cost-effectiveness of IRSs would shed more light on this issue. In addition, further longitudinal research is required to explore: the impact of IRSs on patient safety outcomes; and how/if IRSs detect, and organizations learn from, the wider latent managerial factors involved in patient safety and harm. Finally, future studies designed to investigate the capacity of IRSs should be better theorized in respect of organizational learning.

References

1. Berwick D. *A Promise to Learn – a Commitment to Act: Improving the Safety of Patients in England.*; 2013.

2. Currie G, Waring J, Finn R. The Limits of Knowledge Management for Uk Public Services Modernization: The Case of Patient Safety and Service Quality. *Public Adm*. 2008;86(2):363-385.

3. Hudson P. Applying the lessons of high risk industries to health care. *Qual Saf Health Care*. 2003;12(suppl 1):i7-i12.

4. McCarthy D, Blumenthal D. Stories from the sharp end: case studies in safety improvement. *Milbank Q*. 2006;84(1):165-200.

5. Rashman L, Withers E, Hartley J. Organizational learning and knowledge in public service organizations: A systematic review of the literature. *Int J Manag Rev*. 2009;11(4):463-494.

6. Donabedian A. Evaluating the Quality of Medical Care. *Milbank Q*. 2005;83(4):691-729.

7. Kohn LT, Corrigan JM, Donaldson MS, others. *To Err Is Human: Building a Safer Health System*. Vol 627. National Academies Press; 1999.

8. NPSA. Seven Steps to Patient Safety. 2004. http://www.nrls.npsa.nhs.uk/resources/collections/seven-steps-to-patient-safety/. Accessed October 16, 2014.

9. Fitzgerald E, Cawley D, Rowan NJ. Irish staff nurses perceptions of clinical incident reporting. *Int J Nurs Midwifery*. 2011;3(2):14-21.

10. NPSA. Organisation Patient Safety Incident Reports. 2014. http://www.nrls.npsa.nhs.uk/patient-safety-data/organisation-patient-safety-incident-reports/. Accessed October 16, 2014.

11. Chassin MR, Loeb JM. High-Reliability Health Care: Getting There from Here. *Milbank Q*. 2013;91(3):459-490.

12. Renshaw M, Vaughan C, Ottewill M, Ireland A, Carmody J. Clinical incident reporting: wrong time, wrong place. *Int J Health Care Qual Assur*. 2008;21(4):380-384.

13. Travaglia JF, Westbrook MT, Braithwaite J. Implementation of a patient safety incident management system as viewed by doctors, nurses and allied health professionals. *Health*. 2009;13(3):277-296.

14. Waring JJ. Constructing and re-constructing narratives of patient safety. *Soc Sci Med*. 2009;69(12):1722-1731.

15. Wachter B. Hospital Incident Reporting Systems: Time to Slay the Beast. *Wachters World*. http://community.the-hospitalist.org/2009/09/20/hospital-incident-reporting-systems-time-to-slay-the-monster/. Accessed October 17, 2014.

16. Vincent C. Incident reporting and patient safety. *BMJ*. 2007;334(7584):51-51.

17. Dixon-Woods M. Why is patient safety so hard? A selective review of ethnographic studies. *J Health Serv Res Policy*. 2010;15(suppl 1):11-16.

18. Mahajan RP. Critical incident reporting and learning. *Br J Anaesth*. 2010;105(1):69-75.

19. Doherty C, Stavropoulou C. Patients' willingness and ability to participate actively in the reduction of clinical errors: A systematic literature review. *Soc Sci Med*. 2012;75(2):257-263.

20. Doherty C, Saunders MNK. Elective surgical patients' narratives of hospitalization: the co-construction of safety. *Soc Sci Med*. 2013;98:29-36.

21. OECD. *Health at a Glance 2013*. Paris; 2013. http://www.oecd.org/els/health-systems/Health-at-a-Glance-2013.pdf.

22. Zhan C, Smith SR, Keyes MA, Hicks RW, Cousins DD, Clancy CM. How useful are voluntary medication error reports? The case of warfarin-related medication errors. *Jt Comm J Qual Patient Saf Jt Comm Resour*. 2008;34(1):36-45.

23. Argyris C, Schön DA. *Organizational Learning: A Theory of Action Perspective*. Addison-Wesley Pub. Co.; 1978.

24. ASRS. *ASRS Program Briefing.*; 2014. http://asrs.arc.nasa.gov/docs/ASRS_ProgramBriefing2013.pdf.

25. OGP. *Health and Safety Incident Reporting System Users' Guide, 2010 Data.* International Association of Oil & Gas Producers; 2010. http://www.ogp.org.uk/pubs/433.pdf.

26. Leape LL. EDITORIAL: Why should we report adverse incidents? *J Eval Clin Pract*. 1999;5(1):1-4.

27. Lewis GH, Vaithianathan R, Hockey PM, Hirst G, Bagian JP. Counterheroism, common knowledge, and ergonomics: concepts from aviation that could improve patient safety. *Milbank Q*. 2011;89(1):4-38.

28. HSCIC. *NHS Outcomes Framework 2014/15: Domain 5 – Treating and Caring for People in a Safe Environment and Protecting Them from Avoidable Harm.*; 2014. https://indicators.ic.nhs.uk/download/Outcomes%20Framework/Specification/NHSOF_Domain_5_S_V2.pdf.

29. Dopfer K, Foster J, Potts J. Micro-meso-macro. *J Evol Econ*. 2004;14(3):263-279.

30. Reason JT. *The Human Contribution*. Ashgate Burlington, VT; 2008. http://www.ashgatepublishing.com/pdf/leaflets/The_Human_Contribution_2009.pdf. Accessed February 1, 2015.

31. ASRS. *ASRS: The Case for Confidential Incident Reporting Systems*.; 2006. http://asrs.arc.nasa.gov/docs/rs/60_Case_for_Confidential_Incident_Reporting.pdf.

32. Barach P, Small SD. Reporting and preventing medical mishaps: lessons from non-medical near miss reporting systems. *Bmj*. 2000;320(7237):759-763.

33. Wu AW, Pronovost P, Morlock L. ICU incident reporting systems. *J Crit Care*. 2002;17(2):86-94.

34. Amalberti R, Vincent C, Auroy Y, de Saint Maurice G. Violations and migrations in health care: a framework for understanding and management. *Qual Saf Health Care*. 2006;15(suppl 1):i66-i71.

35. Kennedy I. The report of the public inquiry into children's heart surgery at the Bristol Royal Infirmary 1984-1995: learning from Bristol. 2001. http://webarchive.nationalarchives.gov.uk/+/www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_4005620.

36. Francis R. *Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry*. London: The Stationery Office.; 2013.

37. Turner BA. The organizational and interorganizational development of disasters. *Adm Sci Q*. 1976:378-397.

38. Reason JT. *Managing the Risks of Organizational Accidents*. Vol 6. Ashgate Aldershot; 1997.

39. Seckler-Walker J, Taylor-Adams S. Clinical incident reporting. In: *Clinical Risk Management: Enhancing Patient Safety*. Vol London: BMJ Books; 2001.

40. Anderson JE, Kodate N, Walters R, Dodds A. Can incident reporting improve safety? Healthcare practitioners' views of the effectiveness of incident reporting. *Int J Qual Health Care*. 2013:mzs081.

41. Waring JJ. Beyond blame: cultural barriers to medical incident reporting. *Soc Sci Med*. 2005;60(9):1927-1935. doi:10.1016/j.socscimed.2004.08.055.

42. Cyert RM, March JG, others. A behavioral theory of the firm. *Englewood Cliffs NJ*. 1963;2.

43. Easterby-Smith M, Araujo L, Burgoyne J. *Organizational Learning and the Learning Organization: Developments in Theory and Practice*. SAGE; 1999.

44. Easterby-Smith M, Lyles MA. Organizational learning and knowledge management. 2003. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.129.408. Accessed March 9, 2015.

45. Shipton H. Cohesion or confusion? Towards a typology for organizational learning research1. *Int J Manag Rev*. 2006;8(4):233-252.

46. Argote L, Miron-Spektor E. Organizational learning: From experience to knowledge. *Organ Sci*. 2011;22(5):1123-1137.

47. Chiva R, Alegre J. Organizational learning and organizational knowledge towards the integration of two approaches. *Manag Learn*. 2005;36(1):49-68.

48. Waring JJ, Bishop S. Lean healthcare: rhetoric, ritual and resistance. *Soc Sci Med*. 2010;71(7):1332-1340.

49. Gherardi S, Nicolini D, Odella F. Toward a Social Understanding of How People Learn in Organizations The Notion of Situated Curriculum. *Manag Learn*. 1998;29(3):273-297.

50. Abbott A. *The System of Professions: An Essay on the Division of Expert Labor*. University of Chicago Press; 2014.

51. Freidson E. *Professionalism, the Third Logic: On the Practice of Knowledge*. University of Chicago Press; 2001.

52. Braithwaite J, Westbrook M, Travaglia J. Attitudes toward the large-scale implementation of an incident reporting system. *Int J Qual Health Care*. 2008;20(3):184-191.

53. Evans SM. Attitudes and barriers to incident reporting: a collaborative hospital study. *Qual Saf Health Care*. 2006;15(1):39-43. doi:10.1136/qshc.2004.012559.

54. Lawton R, Parker D. Barriers to incident reporting in a healthcare system. *Qual Saf Health Care*. 2002;11(1):15-18.

55. Waring J. Adaptive regulation or governmentality: patient safety and the changing regulation of medicine. *Sociol Health Illn*. 2007;29(2):163-179.

56. Waring JJ. A qualitative study of the intra-hospital variations in incident reporting. *Int J Qual Health Care J Int Soc Qual Health Care ISQua*. 2004;16(5):347-352.

57. Aagaard L, Soendergaard B, Andersen E, Kampmann JP, Hansen EH. Creating knowledge about adverse drug reactions: a critical analysis of the Danish reporting system from 1968 to 2005. *Soc Sci Med*. 2007;65(6):1296-1309.

58. Yeung K, Dixon-Woods M. Design-based regulation and patient safety: a regulatory studies perspective. *Soc Sci Med*. 2010;71(3):502-509.

59. DiMaggio PJ, Powell WW. The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *Am Sociol Rev*. 1983:147-160.

60. Argyris C. Single-loop and double-loop models in research on decision making. *Adm Sci Q*. 1976:363-375.

61. Argyris C. *Reasoning, Learning, and Action: Individual and Organizational*. Jossey-Bass San Francisco, CA; 1982.

62. Edmondson A. Psychological safety and learning behavior in work teams. *Adm Sci Q*. 1999;44(2):350-383.

63. Pawson R, Greenhalgh T, Harvey G, Walshe K. Realist synthesis: an introduction. *Manch ESRC Res Methods Programme Univ Manch*. 2004. http://betterevaluation.org/sites/default/files/RMPmethods2.pdf. Accessed February 1, 2015.

64. Pope C, Mays N, Popay J. *Synthesising Qualitative and Quantitative Health Evidence: A Guide to Methods: A Guide to Methods*. McGraw-Hill International; 2007.

65. Rosenthal J, Booth M. *Maximizing the Use of State Adverse Event Data to Improve Patient Safety*. National Academy for State Health Policy; 2005. https://www.premierinc.com/quality-safety/tools-services/safety/safety-share/03-06-downloads/01-patient-safety-gnl61.pdf. Accessed October 16, 2014.

66. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesising qualitative and quantitative evidence: a review of possible methods. *J Health Serv Res Policy*. 2005;10(1):45-53B.

67. Dixon-Woods M, Agarwal S, Jones D, Young B, Sutton A. Synthesising qualitative and quantitative evidence: a review of possible methods. *J Health Serv Res Policy*. 2005;10(1):45-53B.

68. Wong SC, Kelly SP, Sullivan PM. Patient safety in vitreoretinal surgery: quality improvements following a patient safety reporting system. *Br J Ophthalmol*. 2013;97(3):302-307.

69. Weissman JS, Annas CL, Epstein AM, et al. Error reporting and disclosure systems: views from hospital leaders. *Jama*. 2005;293(11):1359-1366.

70. Conlon P, Havlisch R, Porter C. Using an Anonymous Web-Based Incident Reporting Tool to Embed the Principles of a High-Reliability Organization. In: *Advances in Patient Safety: New Directions and Alternative Approaches (Vol. 1: Assessment)*. Vol Agency for Healthcare Research and Quality; 2008. http://www.ncbi.nlm.nih.gov/books/NBK43630. Accessed October 16, 2014.

71. Pierson S, Hansen R, Greene S, et al. Preventing medication errors in long-term care: results and evaluation of a large scale web-based error reporting system. *Qual Saf Health Care*. 2007;16(4):297-302.

72. Ross LM, Wallace J, Paton JY. Medication errors in a paediatric teaching hospital in the UK: five years operational experience. *Arch Dis Child*. 2000;83(6):492-497.

73. Spigelman AD, Swan J. Review of the Australian incident monitoring system. *ANZ J Surg*. 2005;75(8):657-661.

74. Grissinger MC, Hicks RW, Keroack MA, Marella WM, Vaida AJ. Harmful medication errors involving unfractionated and low-molecular-weight heparin in three patient safety reporting programs. *Jt Comm J Qual Patient Saf*. 2010;36(5):195-202.

75. Savage SW, Schneider PJ, Pedersen CA. Utility of an online medication-error-reporting system. *Am J Health Syst Pharm*. 2005;62(21):2265-2270.

76. Wagner LM, Capezuti E, Clark PC, Parmelee PA, Ouslander JG. Use of a falls incident reporting system to improve care process documentation in nursing homes. *Qual Saf Health Care*. 2008;17(2):104-108.

77. Nicolini D, Waring J, Mengis J. Policy and practice in the use of root cause analysis to investigate clinical adverse events: mind the gap. *Soc Sci Med*. 2011;73(2):217-225.

78. Callum JL, Kaplan HS, Merkley LL, et al. Reporting of near-miss events for transfusion medicine: improving transfusion safety. *Transfusion (Paris)*. 2001;41(10):1204-1211.

79. Elhence P, Veena S, Sharma RK, Chaudhary RK. Root cause analysis of transfusion error: identifying causes to implement changes. *Transfusion (Paris)*. 2010;50(12pt2):2772-2777.

80. Grant C, Ludbrook G, Hampson EA, Semenov R, Willis R. Adverse physiological events under anaesthesia and sedation: a pilot audit of electronic patient records. *Anaesth Intensive Care*. 2008;36(2):222-229.

81. NPSA. *Review of Patient Safety for Children and Young People*. London; 2009. http://www.nrls.npsa.nhs.uk/resources/?entryid45=59864. Accessed October 17, 2014.

82. Takeda H, Matsumura Y, Nakajima K, et al. Health care quality management by means of an incident report system and an electronic patient record system. *Int J Med Inf*. 2003;69(2):285-293.

83. Stanhope N, Crowley-Murphy M, Vincent C, O'Connor AM, Taylor-Adams SE. An evaluation of adverse incident reporting. *J Eval Clin Pract*. 1999;5(1):5-12.

84. Percarpio KB, Watts BV. A cross-sectional study on the relationship between utilization of root cause analysis and patient safety at 139 Department of Veterans Affairs medical centers. *Jt Comm J Qual Patient Saf Jt Comm Resour*. 2013;39(1):32-37.

85. Sari AB-A, 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.

86. Marang-van de Mheen PJ, van Hanegem N, Kievit J. Effectiveness of routine reporting to identify minor and serious adverse outcomes in surgical patients. *Qual Saf Health Care*. 2005;14(5):378-382.

87. Beckmann U, Bohringer C, Carless R, et al. Evaluation of two methods for quality improvement in intensive care: Facilitated incident monitoring and retrospective medical chart review. *Crit Care Med*. 2003;31(4):1006-1011.

88. Wolff AM, Bourke J, Campbell IA, Leembruggen DW. Detecting and reducing hospital adverse events: outcomes of the Wimmera clinical risk management program. *Med J Aust*. 2001;174(12):621-625.

89. NHS Quality Improvement Scotland. *Safe Today - Safer Tomorrow: Patient Safety - Review of Incident and Near-Miss Reporting*. NHS Quality Improvement Scotland; 2006.

90. Cooke DL, Dunscombe PB, Lee RC. Using a survey of incident reporting and learning practices to improve organisational learning at a cancer care centre. *Qual Saf Health Care*. 2007;16(5):342-348.

91. Sullivan M, Ghroum P. Incident reporting to improve clinical practice in a medium-secure setting. *Ment Health Pract*. 2013;16(7):16-20.

92. Jayaram G, Doyle D, Steinwachs D, Samuels J. Identifying and reducing medication errors in psychiatry: creating a culture of safety through the use of an adverse event reporting mechanism. *J Psychiatr Pract*. 2011;17(2):81-88.

93. Boyle TA, Scobie AC, MacKinnon NJ, Mahaffey T. Quality-related event learning in community pharmacies: manual versus computerized reporting processes. *J Am Pharm Assoc JAPhA*. 2011;52(4):498-506.

94. Stainsby D, Jones H, Asher D, et al. Serious hazards of transfusion: a decade of hemovigilance in the UK. *Transfus Med Rev*. 2006;20(4):273-282.

95. Stainsby D, Williamson L, Jones H, Cohen H. 6 Years of shot reporting—its influence on UK blood safety. *Transfus Apher Sci*. 2004;31(2):123-131.

96. Askeland RW, McGrane S, Levitt JS, et al. Improving transfusion safety: implementation of a comprehensive computerized bar code–based tracking system for detecting and preventing errors. *Transfusion (Paris)*. 2008;48(7):1308-1317.

97. Rebibo D, Hauser L, Slimani A, Hervé P, Andreu G. The French Haemovigilance System: organization and results for 2003. *Transfus Apher Sci*. 2004;31(2):145-153.

98. Olsen S, Neale G, Schwab K, 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):40-44.

99. 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):436-446.

100. Frey B, Buettiker V, Hug MI, et al. Does critical incident reporting contribute to medication error prevention? *Eur J Pediatr*. 2002;161(11):594-599.

101. Herzer KR, Mirrer M, Xie Y, et al. Patient safety reporting systems: sustained quality improvement using a multidisciplinary team and "good catch" awards. *Jt Comm J Qual Patient SafetyJoint Comm Resour*. 2012;38(8):339.

102. Murray GR, Cameron ID, Cumming RG. The consequences of falls in acute and subacute hospitals in Australia that cause proximal femoral fractures. *J Am Geriatr Soc*. 2007;55(4):577-582.

103. Zapatero A, Barba R, Canora J, et al. Hip fracture in hospitalized medical patients. *BMC Musculoskelet Disord*. 2013;14:15.

104. Kalapurakal JA, Zafirovski A, Smith J, et al. A comprehensive quality assurance program for personnel and procedures in radiation oncology: Value of voluntary error reporting and checklists. *Int J Radiat Oncol Biol Phys*. 2013;86(2):241-248.

105. Kivlahan C, Sangster W, Nelson K, Buddenbaum J, Lobenstein K. Developing a comprehensive electronic adverse event reporting system in an academic health center. *Jt Comm J Qual Patient Saf*. 2002;28(11):583-594.

106. Ford EC, Smith K, Harris K, Terezakis S. Prevention of a wrong-location misadministration through the use of an intradepartmental incident learning system. *Med Phys*. 2012;39(11):6968-6971.

107. Roughead EE, Gilbert AL, Primrose JG. Improving drug use: a case study of events which led to changes in use of flucloxacillin in Australia. *Soc Sci Med*. 1999;48(6):845-853.

108. Wysowski DK, Swartz L. Adverse drug event surveillance and drug withdrawals in the United States, 1969-2002: the importance of reporting suspected reactions. *Arch Intern Med*. 2005;165(12):1363-1369.

109. Hutchinson A, Young TA, Cooper KL, 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.

110. Vincent C, Taylor-Adams S, Chapman EJ, et al. How to investigate and analyse clinical incidents: clinical risk unit and association of litigation and risk management protocol. *Bmj*. 2000;320(7237):777-781.

111. Wenger E. Communities of Practice and Social Learning Systems. *Organization*. 2000;7(2):225-246.

112. 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.

Figure 1: Search Strategy

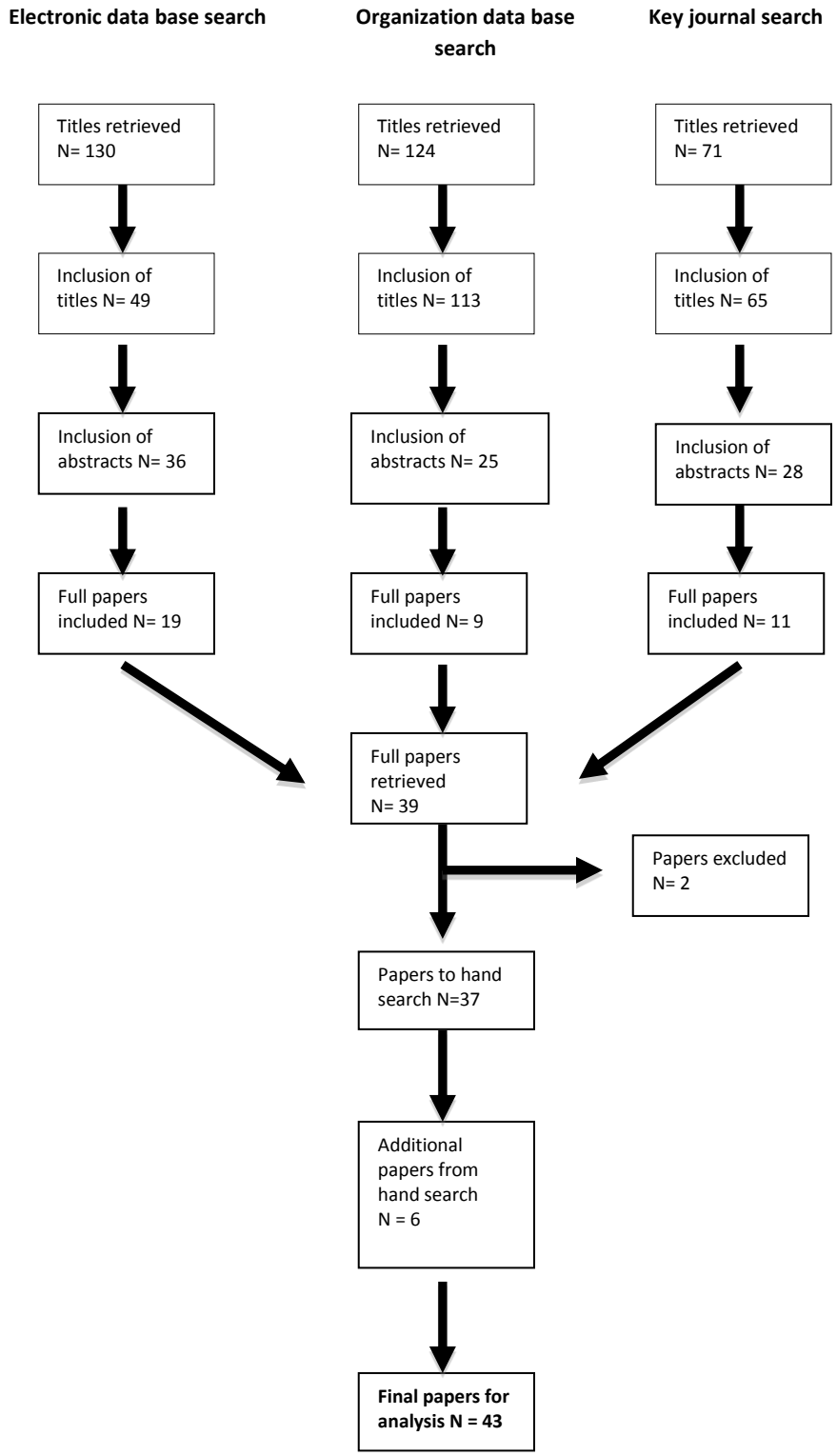

Table 1: Summary of potential facilitators of double loop learning

Facilitator	Characteristics	Studies
Psychological safety	Non-punitive; making reporting less punitive; anonymous, confidential; absence of `blame culture'; removing fear of reprisals.	Anderson et al., 2013; Conlon et al., 2013; Cooke et al., 2007; Elhence et al., 2010; Frey et al., 2002; Jayaram et al., 2011; Kalapurakal et al., 2013; Kivlahan et al., 2002; NHS Quality Improvement Scotland, 2006; Nicolini et al., 2011; Pierson et al., 2007; Ross et al., 2000; Savage et al., 2005; Spiegelman and Swan, 2005; Stainsby et al., 2006; Stainsby et al., 2004; Takeda et al., 2003; Weissman et al., 2005; Wong et al., 2013
Focus on learning	Learning as the function/output (vs `audit culture' etc.); actual, genuine focus on learning (vs rhetorical/espoused); allows for discrepancies, emotion etc.	NHS Quality Improvement Scotland, 2006; Nicolini et al., 2011; Pierson et al., 2007; Wong et al., 2013
Reflects complexity		
Cross-departmental/ organizational/professional	Multi-agency; inter-organizational; multi- or cross- disciplinary; cross-functional, breaks down silos and barrier between departments.	Conlon et al., 2013; Cooke et al., 2007; Herzer et al., 2012; Jayaram et al., 2011; Kivlahan et al., 2002; Pierson et al., 2007; Rebibo et al., 2004; Stainsby et al., 2006
Multiple interventions	Holistic/systemic approach; complementary, system-wide interventions vs single interventions in isolation (e.g. training).	Callum et al., 2001; Cooke et al., 2007; Ross et al., 2000; Roughead et al., 1999; Sullivan and Ghroum, 2013
Local and participative	Built within the context (vs imposed); locally-designed vs centrally or externally designed. Participants are involved in problem-solving; vs hierarchical, in the hands of specialists.	Conlon et al., 2013; Cooke et al., 2007; Herzer et al., 2012; NHS Quality Improvement Scotland, 2006; Nicolini et al., 2011