Review Article

How Effective Are Incident-Reporting Systems for Improving Patient Safety? A Systematic Literature Review

CHARITINI STAVROPOULOU,* CAROLE DOHERTY,† and PAUL TOSEY†

City University London; †University of Surrey

Policy Points:

- Incident-reporting systems (IRSs) are a method of error reporting to enable organizational learning. Despite their significant cost, however, little is known about their effectiveness for improving patient safety.
- Our systematic literature review found no strong evidence that IRSs perform better than other forms of reporting. In addition, although we show that IRSs can improve clinical settings and processes, we found little evidence that they ultimately improve outcomes or enable cultural changes.
- IRSs could work more effectively if the reportable incidents used are defined more clearly and the IRSs have clinical ownership and integration with wider safety programs.

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

Methods: Our systematic literature review identified 2 groups of studies: (1) those comparing the effectiveness of IRSs with other methods of error reporting and (2) those examining the effectiveness of IRSs on settings, structures, and outcomes in regard to improving 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, we identified 43 studies, 8 that compared IRSs with other methods and 35 that explored the effectiveness of IRSs on settings, structures, and outcomes. We did not find strong evidence that IRSs performed better than other methods. We did find some evidence of single-loop learning, that is, changes to *clinical* settings or processes as a consequence of learning from IRSs, but little evidence of either improvements in outcomes or changes in the latent managerial factors involved in error production. In addition, there was insubstantial evidence of IRSs enabling double-loop learning, that is, a cultural change or a change in mind-set.

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

Keywords: patient safety, incident-reporting systems, organizational learning, single-loop and double-loop learning.

T O IMPROVE PATIENT SAFETY, EXPERTS HAVE ARGUED THAT major cultural changes, firmly rooted in continual improvement, are necessary.[1] These changes include constant evidence-based learning; managerial appreciation of the pressures that resource constraints can impose on frontline employees; avoidance of blame; and disregard of mechanistic performance objectives.[1] Incident-reporting systems (IRSs) are designed to obtain information about patient safety which can then be 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. In this systematic review, we examine evidence concerning the effectiveness of IRSs as one way of promoting organizational learning in order 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 and compare them with other forms of incident reporting, such as medical chart reviews. In absolute terms, we use Donabedian's framework[6] to explore the impact of IRSs on settings (structure), processes, and safety outcomes.

IRSs have been used in the health care field for many years, but it was only after the publication of *To Err Is Human*[7] that these 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 United Kingdom, 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% ($n = 5,235$) resulted in severe harm or death.[10]

There are questions about the effectiveness and cost of IRSs, however.[11] Renshaw and colleagues[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," which were time-consuming to complete.[13] Waring[14] argues that the detailed information in clinicians' stories is reassigned via IRSs into abstract, quantitative variables of the managerial system, thereby reducing the effectiveness of IRSs for learning. Wachter[15] contends that incident reports do not provide information about the true frequency of organizational errors and are too expensive and bureaucratic.

Other problems associated with IRSs include the number of incidents reflecting employees' willingness to report rather than indicating the system's safety[16]; the lack of shared understanding among clinicians (doctors, nurses, and other health care professionals) about what constitutes an adverse event or near miss; the lack of clarity about who in the clinical team is responsible for reporting such incidents[17]; and some clinicians' fear of recriminations.[18] Generally, patients do not have independent access to IRSs, and clinicians may not recognize their experiences of harm.[19,20] These concerns raise questions about the utility of IRSs to promote organizational learning to improve patient safety.

In most countries, health expenditures have been declining since the beginning of the global financial crisis in 2008.[21] It therefore is important to determine whether investing in IRSs is money well spent, for both the public and private health sectors.[12,22] This article offers a parallel review of (1) studies comparing the effectiveness of IRSs with other methods of error reporting and (2) studies measuring the effectiveness of IRSs in absolute terms. For the latter, we used Donabedian's[6] settings, processes, and outcomes framework to

systematically review the empirical evidence on the effectiveness of incident-reporting systems for patient safety. Comparing 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, we analyzed these studies using Argyris and Schön's concepts of single-loop and double-loop learning.[23]

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

Incident-Reporting Systems (IRSs)

The theory underpinning IRSs is that in order for organizations to improve their safety performance, managers must be aware of events in their organization and employees must feel confident about reporting 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 in order 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 substantially improve the safety of airline travel, thus leading to the assumption that IRSs would also offer valuable lessons for health care.[7,26,27] An IRS has 2 aspects: first, it reports "adverse events" or "patient safety incidents," that is, any unintended or unexpected incident(s) that led to harm for one or more persons[28(p2)]; and second, it reports "near misses," any event(s) that did not cause harm but could have done so.

At the micro level, that is, the level of an organization on which agents interact and rules are adopted, maintained, changed, or resisted in their local context,[29] Reason[30] argues that IRSs provide a systematic method of enhancing ongoing learning from experience for the primary purpose of improving patient safety. Voluntary confidential reporting is thought

to deepen understanding of the frequency of types of adverse events, near misses, and their patterns and trends and hence acts as a warning system. This information should then be used at the organization's micro and meso levels, that is, individual actors plus the system of rules.[29] At the meso level, common rules should be changed so the system can be redesigned to reduce the possibility of adverse events (re)occurring. Indeed, NASA claims that aviation safety reporting helps identify 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 level, that is, a higher order of organization arising from the existence of interacting populations of meso rules,[29] IRSs are considered to be an accurate early warning system for problems related to emerging technologies and global economic trends.[16,31]

Several authors contend that adverse events occur when active failures—that is, individuals' errors, omissions, or unsafe acts—interact with an organization's latent conditions (underlying structures and processes) to cause harm. Near-miss events are up to 300 times more common than adverse events.[32,33] Evidence suggests that within health care organizations, pressure on frontline employees to be more efficient has created a safety culture in which deviance is normalized[34(p69)] as employees attempt to cope with competing demands by fixing or working around problems at the local level. Accordingly, their actions often hide latent conditions, which increase susceptibility to error and, at the same time, introduce them into the system.[30] In addition, public inquiries into the failings of the United Kingdom's National Health Service (NHS) have found 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 the incidence of errors can be reduced by readjusting such cultural norms.[37] Theories emphasize that 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) that increase susceptibility to the same types of errors occurring in different 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 to and actively involved in its development; stakeholders should reach

consensus on 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 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 give reporters evidence that the information they are providing is being used appropriately.[40,41]

Organizational Learning (Theory)

Since Cyert and March[42] coined the term "organizational learning," scholarship has burgeoned, as reflected in reviews of the field by Easterby-Smith, Araujo, and Burgoyne,[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, Withers, and Hartley,[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, Withers, and Hartley[5] and Waring and Bishop[48] argue that social, situated theories of organizational learning are highly relevant to public service and health care. This type of theory regards organizational learning as complex and emergent, occurring through and embedded in social practices.[49] Health care organizations in particular are characterized by professional communities that span organizational boundaries[5] and involve multiple stakeholders in a complex interprofessional setting.

A social perspective on organizational learning also highlights how the political dimension of knowledge can influence and impede it. Hence in health care, 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 interprofessional politics and agendas. Powerful professional interests can be projected onto initiatives like IRSs, thus seizing them as new territory on which existing battles can be fought. In health care organizations, knowledge forms the basis of professional power and jurisdictional control; what counts as knowledge is contested terrain[50,51] and thus is a source of conflict among the various clinical professions and between clinicians and managers. It may be wrong, therefore, to assume that clinicians are willing to share widely

any information about errors.[14] Doctors, particularly surgeons, are often reluctant to report incidents, and they see IRSs as a managerial encroachment on their professional status and individual autonomy.[41,52-55] But research suggests that they are more inclined to participate when the IRS is situated and managed within the medical department.[56] Within the NHS, evidence indicates an underlying hostility and distrust between doctors and managers, with doctors prioritizing professional learning over organizational learning, and their lack of cooperation undermining the NRLS's implementation.[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 IRSs are accompanied by a rhetoric of learning, they 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 we recognize the merits of a social perspective on organizational learning for the way in which learning from IRSs is likely to occur in health care settings, our specific need here is for a theory of organizational learning that enables us to assess the evidence presented in the studies we examined. Accordingly, 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 of organizational learning, which according to Chiva and Alegre[47] is 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.

Our principal reason for choosing Argyris and Schön's theory, instead of a social theory 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 in order to classify the type of organizational learning produced by IRSs. In particular, their theory 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 section.

Argyris and Schön's theory proposes 2 principal forms of organizational learning. "Single-loop learning" refers to the correction of operational errors without significantly changing the overall safety culture, and "double-loop learning" refers to the questioning and alteration of what Argyis and Schön call "governing variables."

> 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, equivalent to a shift in safety culture and "mind-set" and resulting in a significantly different approach to the treatment of errors in health care 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 behavior that guards people against embarrassment and "exposure to blame."[23(p40)] Defensive behavior in relation to IRSs could lead to not only the nonreporting of errors but also the nonreporting itself being covered up. Hence Edmondson, based on Argyris's 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 double-loop learning in practice requires a climate of sufficient psychological safety[62] to reduce the likelihood of defensive behavior.

Argyris and Schön's theory therefore also enabled us to review the evidence of potential barriers to double-loop learning (the desired cultural change) in the studies we 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 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 behavior.

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 health care contexts, the specific purpose of this article led us to use Argyris and Schön's theory of single- and double-loop learning to interrogate evidence of the type of organizational learning indicated by the studies we reviewed. As we have discussed, IRSs may be problematic for a number of reasons. Nonetheless, there has been no systematic review integrating the studies exploring the effectiveness of IRSs in the health care context.[12,33] Our aim, therefore, is to analyze and synthesize empirical evidence relating to the effectiveness of IRSs as a method of improving patient safety via organizational learning.
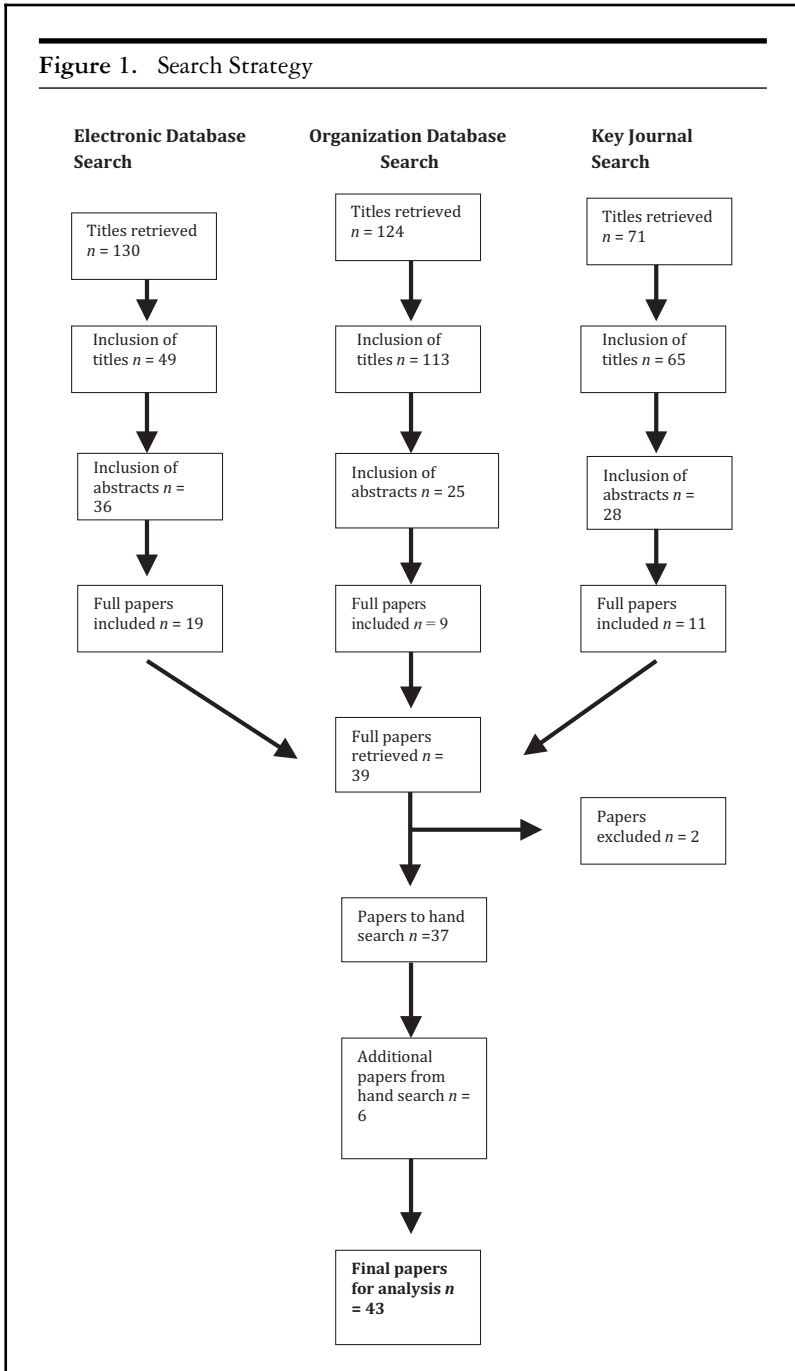
Methods

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 Figure 1 shows, we searched key health care 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 health care journals: *The Milbank Quarterly*, *Social Science and Medicine*, *Quality and Safety in Health Care*, the *International Journal of Health Care Quality Assurance*, *Health*, the *New England Journal of Medicine*, the *Journal of the American Medical Association*, the *British Medical Journal*, the *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: the Agency for Healthcare Research and Quality, The Joint Commission on Accreditation of Healthcare Organizations/International Centre for Patient Safety, the National Patient

Figure 1.   Search Strategy

Electronic Database Search

Titles retrieved
n = 130

↓

Inclusion of
titles *n* = 49

↓

Inclusion of
abstracts *n* =
36

↓

Full papers
included *n* = 19

Organization Database Search

Titles retrieved
n = 124

↓

Inclusion of
titles *n* = 113

↓

Inclusion of
abstracts *n* = 25

↓

Full papers
included *n* = 9

↓

Full papers
retrieved *n* =
39

→ Papers
excluded *n* = 2

↓

Papers to hand
search *n* =37

↓

Additional
papers from
hand search *n* =
6

↓

**Final papers
for analysis *n*
= 43**

Key Journal Search

Titles retrieved
n = 71

↓

Inclusion of
titles *n* = 65

↓

Inclusion of
abstracts *n* =
28

↓

Full papers
included *n* = 11

Safety Agency (UK), the National Patients Safety Foundation (USA), the Health Foundation (UK), the Australian Patient Safety Foundation, the Canadian Patient Safety Institute, the Scottish Patient Safety Programme, the Health Quality and Safety Commission New Zealand, NASA, and the World Health Organization. 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 keywords. Two of us reviewed the abstracts independently, cross-referencing our judgments on the papers. For abstracts to be included, they had to contain empirical data either on comparisons with other methods of incident reporting or in relation to changes in settings, processes, or outcomes as a consequence of knowledge gained initially through information derived from an IRS. We excluded systematic literature reviews. When the reviewers disagreed, we included the abstracts. Once we agreed on the abstracts for inclusion, we removed duplicates and retrieved full papers. Then 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 keywords that included our search terms to identify further papers that we may have omitted in our search to date.

Studies were limited to those published in English with no restrictions on the country of origin or the context in which studies were undertaken. We included only empirical papers that examined the effectiveness of IRSs for patient safety, by either comparing them with other systems or looking at improvements in structure, settings, or outcomes according to Donabedian's[6] framework (described later). We excluded opinion papers, systematic literature reviews, and studies analyzing the effectiveness of IRSs in capturing the number and type of near misses and patient harm events. Barriers preventing clinicians from reporting incidents were beyond the scope of our article.

Many of the studies on the final list described retrospective analyses of quality improvement work in single departments or national coordinating organizations. As explorations of a service quality intervention, they did not necessarily follow orthodox qualitative and quantitative research designs. Therefore, following Pawson and colleagues'[63] argument that the value of such studies is demonstrated in synthesis, we made a pragmatic decision to include papers based on their relevance, that is, if they addressed our research question using the data extraction process

(see online Appendices A and B), rather than by assessing the quality of the selected articles using a standard checklist.

Data Extraction

We identified 2 groups of studies: (1) those comparing the effectiveness of IRSs with other methods of error reporting, such as medical chart review, and (2) those measuring effectiveness in absolute terms. We carried out parallel data analysis to address these different but related aspects of IRSs' effectiveness in improving safety.[64]

Online Appendix A summarizes those studies comparing the effectiveness of IRSs with other methods of error reporting. We extracted the data for these studies by comparing and contrasting the various methods of reporting and their outcomes within and across the studies.

Online Appendix B summarizes the studies measuring effectiveness in absolute terms. We acknowledge that measuring effectiveness can be both complex and challenging.[65] Thus, to ensure transparency, we extracted the data from these studies using Donabedian's settings (structure), processes, and outcomes framework,[6] thereby defining "effectiveness" in absolute terms as the following outcome types:

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

Donabedian 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, however, are not as stable as outcome indicators. Furthermore, it is difficult to make causal links among 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 synthesizing the evidence. We did this by combining a summary of the data showing which types of changes

in practice were made with an interpretation of the data grounded in assumptions about how IRSs should work.[66]

We initially read each paper in the first group of studies to identify the comparative methods employed and their relative advantages and/or limitations. Then we compared and contrasted the studies to identify similarities, patterns, and contradictions, a recursive process that involved reading and rereading individual articles and moving back and forth between articles. For the second group of studies, we first extracted and tabulated (online Appendix B) how adverse incidents were conceptualized, the types of changes made in the practice, and whether these involved settings, processes, or outcomes. We then searched each article for evidence of improvements in patient safety as a result of the changes implemented. Following this, we used thematic analysis, which is considered a suitable method of organizing and summarizing the findings from both qualitative and quantitative research,[64,66] to identify systematically across the studies the main themes regarding what was effective (or ineffective), where (context), and why.

Findings

Descriptive Analysis of Studies

In total we included 43 studies in our analysis, the majority of which were conducted in the United States (16), followed by the United Kingdom (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 and meso levels. Of those, 15 were in general hospitals, and 9 were in specialized units (3 oncology departments, 1 pediatric unit, 1 obstetric unit, 1 hospital-based transfusion service, 1 eye hospital, 1 psychiatric division of a teaching hospital, and 1 tertiary cancer center). Two of the remaining 5 studies were conducted in an intensive care unit (ICU), one of which was general and the other neonatal-pediatric; 2 studies were done in 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, 9 investigated incidents reported in large-scale reporting programs: the United Kingdom's NRLS; Australia's AIMS; the IRSs in NHS Scotland, the US Food and Drug Administration (FDA), and the US

Veterans Health Administration; 3 hemovigilance reporting programs, including the Serious Hazard of Transfusion (SHOT); and 1 program on pharmacies.

Multiple Definitions

Our analysis shows that the studies used a wide variety of terms to describe adverse events, such as "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," and "reviewable sentinel events." One study[67] used "clinical incident," "clinical error," "critical incident," "adverse event," and "adverse incident." Weissman and colleagues[68] analyzed data from 4 hospitals, all of which used different terminology for adverse events.

Of the 43 studies, 26 provided clear definitions of what was considered an adverse event; 9 failed to provide any definition[40,69-76]; 5 used classifications rather than definitions to categorize incidents[77-81]; and 3 acknowledged the difficulties of definition and recommended more conceptual clarity.[57,65,82]

One example of an approach taken by those studies that did provide definitions is that by Percarpio and Watts[83] 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."[83(p35)] Sari and colleagues[84] and Wong, Kelly, and Sullivan[67] used very broad definitions to describe adverse events, like unplanned events that could cause harm to or undesired outcomes for patients. Some studies specified the type of outcome caused by an adverse event, its timing, and who experienced it. Marang-van de Mheen, van Hanegem, and Kievit[85] specified that an adverse event could take place either during or following medical care and could 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.[86-88]

Cooke, Dunscombe, and Lee[89] defined an adverse event as any impairment in the patient care system's quality, efficiency, or effectiveness. Only 1 study discussed damage or loss of equipment or property, and only 1 discussed incidents of violence, aggression, and self-harm.[90] All the definitions cited harm to the patient, with only 3 extending their definitions to include a staff member[87-89] and 1 to include a visitor.[87]

The definitions of medication errors were more exact but again varied among the studies. Although Jayaram and colleagues[91] and Zhan and colleagues[22] used similar definitions, Zhan and colleagues' were 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 health care professional, patient or consumer."[22(p37)] Boyle and colleagues[92] specified the type of error, including incorrect drug quality, dose, or patient.

Adverse events in blood transfusion were generally well defined. In the United Kingdom, in addition to detailing the categories of adverse events reportable to, and monitored by, the SHOT scheme, Stainsby and colleagues[93,94] listed nonreportable events, such as reactions to plasma products. This scheme is professionally led by and affiliated with the Royal College of Pathologists.[93] In the United States, Askeland and colleagues[95] gave a very detailed description of the categories of adverse events that can occur during the "blood product history." Similarly, Callum and colleagues[77] described causal codes for latent and active failures and patient-related factors in a Canadian hospital, and Rebibo and colleagues[96] defined "hemovigilance" as 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 the studies' conceptual clarity and comparability, whose implications we address in the discussion section.

Studies Comparing IRSs With Other Systems

Of the 8 studies that compared IRSs with other systems (presented in online Appendix A), 4 compared IRSs with retrospective medical chart reviews.[82,84-86] Beckman and colleagues conducted a study in an ICU in Australia[86] in which 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 did the retrospective medical chart review. The qualitative differences in the types of adverse events were highlighted in the 2 forms of reporting. Equipment problems and adverse events related to the retrieval team were reported only by the IRS. The authors speculated that staff believed that the patients' medical records were not the correct place for reporting such problems. The IRS identified near misses, but the medical chart review did not. Unplanned readmissions were deemed to be due to adverse events in only 3 cases in the medical chart review, whereas the IRS detected 6. The medical chart review cited incidents like iatrogenic infections and unrelieved pain, which were not identified by the IRS. The medical chart review also found evidence of patients' breathing problems not found in the IRS, possibly because they did not lead to an obvious adverse event like a longer stay in the ICU. Beckman and colleagues argued that both the IRS and the medical chart review were able to identify problems of patient safety in intensive care that were responsive to actions to improve the quality of care, but they did not provide evidence of changes in process or outcomes.

Sari and colleagues[84] compared an IRS with a retrospective medical chart review in an English NHS hospital. They found that the medical records had documented cases of unplanned transfers to ICU, unplanned returns of patients to the operating theater, inappropriate self-discharges, and unplanned readmissions. Not one of these cases was reported in the IRS, indicating underreporting.

Similarly, Stanhope and colleagues[82] examined the reliability of IRSs in 2 obstetric units in London, concluding that although IRSs can provide useful information, they may seriously underestimate the overall number of incidents.

Marang-van de Mheen, van Hanegem, and Kievit[85] compared the incident reports of clinically occurring adverse events gathered by surgeons and discussed at their weekly specialty meeting with a retrospective medical chart review in a sample of high-risk surgical patients in a Dutch hospital. They found that both the IRS and the medical chart review missed adverse events, again suggesting underreporting. The medical chart review identified significantly more adverse events overall than did routine reporting, thus supporting Sari and colleagues' findings.[84] But the IRS identified serious adverse events that were missed by the medical

chart review. The IRS did not identify adverse events occurring after discharge or ward/hospital transfer. Marang-van de Mheen, van Hanegem, and Kievit[85] argued that when incident reporting was controlled by the clinicians and supported by discussion at regular, peer-led meetings, it had distinct advantages compared with macro-level quality improvement initiatives such as the National Confidential Enquiry into Patient Outcome and Death reports. The authors maintained that local ownership of the data gave clinicians an opportunity to study adverse events in their specialty, responsibility for implementing recommendations, and longitudinal data to study trends and monitor the effectiveness of changes in practice. The studies by both Beckman and colleagues[86] and Marang-van de Mheen, van Hanegem, and Kievit similarly highlight the importance of ownership of the IRS at the micro level for individual and departmental learning.

Of the other 4 studies comparing IRSs with other systems,[75,92,97,98] Olsen and colleagues[97] looked at 3 different methods of detecting drug-related adverse events in an English NHS hospital: the IRS, the active surveillance of prescription charts by pharmacists, and a medical chart review. Similar to Beckmann and colleagues,[86] Stanhope and colleagues[82] and Marang-van de Mheen, van Hanegem, and Kievit[85] found that the IRS provided a less acceptable indication of clinical adverse events than did the 2 other methods, thereby concluding that the IRS was effective only when supplemented with other data collection. Flynn and colleagues[98] also compared 3 methods for detecting medication errors: an IRS, a medical chart review, and direct observation. Direct observation involved the researchers observing nurses administering 50 prescriptions during the morning medication administration round. The observers were nurses and pharmacy technicians, and they were paid to collect the data. The study concluded that direct observation was the most efficient and accurate of the 3 methods. But like the other studies cited, it gave no indication of the relative resources involved.

The third study, by Wagner and colleagues,[75] tested the effectiveness of a computerized "falls IRS" providing a standardized structure and consistency for those items to include in the report, comparing this with a semistructured, open-ended description type of report often used in US nursing homes. Their findings suggested that the post-fall evaluation process was documented more completely in the medical records in those nursing homes using the computerized IRS than in nursing homes using a nonstandardized descriptive type of reporting. Similarly, Boyle

and colleagues[92] assessed manual versus computerized IRSs in pharmacies in Canada. The pharmacists reported that both computerized and manual incident reporting were cost-effective and easy to complete. Those pharmacists using computerized reporting systems, however, rated their utility higher than did those working with manual systems.

To summarize, the 8 studies that compared IRSs with other reporting methods showed 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 the IRSs themselves on settings, processes, and outcomes (summarized in online Appendix B). The micro and meso changes reported in these studies were of 3 types: (1) changes to policies, guidelines, and documentation; (2) provision of staff training; and (3) implementation of technology. Then we summarize the macro-level impacts reported in 9 of the 35 studies, and finally, we present our analysis through the lens of organizational learning theory.

Changes to Policies, Guidelines, and Documentation

Frey and colleagues[99] reported 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 labeling of infusion syringes. But they did not evaluate the effectiveness of these changes for patient safety. Anderson and colleagues[40] discovered that many policy changes had been introduced in both an acute care and a mental health hospital in London. Again, no evaluation of the impact on safety was provided. Only a few frontline clinicians participated in this study because they were not familiar with incident reporting and often were 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, Kelly, and Sullivan[67] 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 and colleagues[79] examined patterns of adverse events in an Australian hospital using data from an electronic record-keeping system. They found 2 problematic areas: sedation for colonoscopy and inhalational anesthesia with desflurane. Using this information, the anesthetists developed specific departmental guidelines for these procedures. Subsequently, the number of adverse events during these 2 procedures was significantly reduced. Ross, Wallace, and Paton[71] reported a reduction of medication errors from 9.8 to 6 per year when, after being highlighted by the IRS, 2 people began checking the dispensing of medications. In 2007, an IRS was implemented in a surgical unit at the Johns Hopkins Hospital in Baltimore, with "Good Catch" awards given to staff that reported and helped prevent safety hazards.[100] At the time of publication in 2012, the authors noted that the 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 directly measure the impact of the IRS on safety culture, noting that because their project coincided with several other quality improvement initiatives, they were unable to attribute the changes in safety culture to any one initiative.

Wolff and colleagues[87] reported a reduction in the number of falls resulting in fractures following the implementation of falls risk assessments, after the IRS identified falls as the most common adverse event. This was a cross-sectional study, so 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,[101,102] suggesting that IRSs have made little impact on patient safety in regard to 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.[103] As a result, at least 2 therapists in the treatment room had to take daily pretreatment time-outs before delivering treatment to the patient, followed by posttreatment-planning time-outs 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.[103]

In a medium-secure hospital in Wales, Sullivan and Ghroum[90] analyzed data from an IRS to find the peak periods for adverse events involving violence, aggression, and self-harm. An improvement plan was implemented that 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 2-year period. The context for this study was relatively unusual in that because staff are often the recipients of violence and such adverse events are highly visible, they may be more motivated to learn from IRSs.

Provision of Staff Training

A number of studies found data from the IRSs that indicated a need for staff training. In some cases, training was introduced to raise awareness of risks and establish a culture of safety,[91] and in others, it was to improve clinicians' skills. Examples are training to improve nurses' ability to administer drugs,[71] safe-prescribing teaching sessions for residents,[99] training to improve clinicians' recognition of mental health issues in young people,[80] education on preventing incompatible blood transfusions,[77] 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.[89]

Callum and colleagues[77] showed that educational sessions on preventing ABO-incompatible transfusions were ineffective, as the rate of adverse events remained unchanged. Similarly, Cooke, Dunscombe, and Lee[89] 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, Dunscombe, and Lee[89] and Anderson and colleagues[40] suggest a disconnect between the micro and the meso levels of organization.

Many of the studies did not report the impact of training on improving the actual process of care and ultimately improving outcomes, even though this was one of the quality improvement methods used.[67,72,78,80,91,93,94,99,104] Indeed, the only study reporting evidence of a direct impact from training was that by Ross and colleagues[71] in a UK pediatric hospital, which showed that training provided to all nurses administering intravenous (IV) drugs reduced errors. We should note, though, that this occurred when nurses were beginning to take over IV drug administration from doctors, and the authors explained that

"nurses are increasingly responsible for giving all medications, precisely because they have better error trapping systems in place."[71(p495)]

Implementation of Technology

The implementation of technology was the third most commonly documented change to practice in the studies we reviewed. Askeland and colleagues[95] reported on the introduction of bar code technology throughout the blood transfusion process in a US hospital in order to help prevent transfusion errors. They found that the bar code system was considered 3 times safer than the old manual system. Callum and colleagues[77] 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, which established implementation and expiration date labeling as priorities. Although Callum and colleagues[77] argued that this would reduce the errors associated with labeling of the expiration date, they offered no actual evidence. In addition, the hospital implemented a trial that mandated labeling at the bedside via a system using wristband bar codes and portable handheld data terminals and printers to allow easy bedside labeling. The authors reported an improvement as a result in blood group determination and antibody screens in the emergency room. A new requisition form was introduced as well, on which the area to sign was delineated by a thick black box and written above the box in big letters was "Please read and sign." They pointed out, however, that this change did not provide sufficient reinforcement, suggesting the need to evaluate the electronic signature as a "forcing function"[77(p1209)] to eliminate this type of error.

Ford and colleagues[105] reported that the Johns Hopkins Hospital's Department of Radiation Oncology implemented a change in which the medical physicists "hid" the treatment fields not currently being used for patients, thereby eliminating human error. Indeed, after this change was made, not one out-of-sequence treatment was reported.

Finally, a significant reduction in reported prescribing errors was found by Jayaram and colleagues[91] following the introduction of an electronic system allowing pharmacists to page immediately any doctor who entered an incorrect order so that it could be corrected.

In conclusion, these 35 studies found 3 types of micro- and meso-level changes prompted by IRSs. All 4 instances of the implementation

of technology were reported as being successful, although the studies did not always evaluate the effectiveness of the changes reported for patient safety outcomes. For example, only 1 out of 12 studies that reported provision of training did so. Because only a few studies reported the outcomes of IRSs, 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, Gilbert, and Primrose[106] analyzed the case of the antibiotic flucloxacillin in Australia. After data from the Drug Reaction Advisory Committee raised health professionals' awareness of the adverse hepatic reaction associated with the use of flucloxacillin, its use was significantly decreased.

Wysowski and Swartz[107] 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. Furthermore, 75 drug products were removed from the market owing to safety concerns, and 11 had special requirements for prescription or restricted distribution programs.

Similarly, two guidelines, one regarding the management of a suspected transfusion-transmitted bacterial contamination and the other regarding the process of transfusion in France, were published in 2003.[96] The authors reported that the number of incompatible ABO transfusions was reduced between 2002 and 2003 and that the misdiagnosis of adverse blood transfusion events was more efficiently identified and investigated.

Zhan and colleagues[22] analyzed voluntary reports of errors related to the use of warfarin in a large number of hospitals in the United States from 2002 to 2004, and they mentioned a number of changes in patient care, including increased monitoring and alterations to protocols. They did not state whether such changes reduced errors. Grissinger and colleagues[73] analyzed errors involving heparin gleaned from data aggregated from 3 large IRSs. The 3 programs used different terms to categorize the areas where errors occurred, complicating the aggregation of this information at the macro level. Although this cross-sectional study found significant harm caused by heparin, it did not explore

whether organizations learned from the IRSs and whether this reduced levels of harm. The authors found common patterns of events in all 3 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[72] surveyed 12 organizational users of the AIMS. The respondents reported numerous settings and process changes, including equipment standardization, new standards for medication prescribing and administration, and improvements in staffing level. The authors noted that the medical staff had a poor level of reporting; that improvements in outcomes resulting from changes implemented were difficult to ascertain; and that if the AIMS was to show outcome improvements in patient safety, the level of resources required should not be underestimated.

In the United Kingdom, Hutchinson and colleagues[108] contended that the NPSA gave hospitals feedback that enabled them to benchmark their performance against other similar hospitals. Nonetheless, improvements in processes and outcomes at the meso level, arising from the aggregation of data at the NPSA's macro level, were not reported.

Conlon, Havlisch, and Porter[69] analyzed the IRS introduced in 2001 in 36 Trinity hospitals and affiliates in the United States, and they described numerous changes in practice as a consequence of learning from the IRS data. The authors conceded that it was difficult to attribute improvements solely to the IRS, as the organization employed various improvement efforts. It had achieved a 26% 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 of evidence that IRSs prompted any of the 2 types of learning. This was an interpretive process that entailed debate about how to apply Argyris and Schön's theory rigorously 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 online Appendix C.

First, we observed that the evidence presented by 33 of the 35 studies could be classified as single-loop learning, such as direct improvements to procedures. Examples are a new bar code system leading to the correction of errors and improvements in patient safety[95]; new labeling[99,104]; and the implementation of new blood transfusion guidelines.[96] 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,[73] and the other PSA[80] was concerned with making recommendations for improving patient safety.

Turning to double-loop learning, based on our review we discovered little conclusive or convincing evidence in the studies we analyzed 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, it could be inferred from some studies that an effective safety culture already exists[104,105]; if so, double-loop learning would effectively be redundant. Second, with the exception of Aagard and colleagues,[57] Cooke, Dunscombe, and Lee,[89] NHS QI,[88] and Nicolini, Waring, and Mengis,[76] the studies we reviewed made little explicit use of organizational learning theory and lacked theoretically informed conceptualizations of cultural change. In the absence of a theoretical framework, such studies inevitably struggle to capture convincing evidence of cultural shifts in patient safety. Third, those studies confined to investigating outcomes that ensue directly and immediately from an IRS 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-loop than of double-loop learning.

Ten of the studies contain claims that *could* refer to double-loop learning. The most detailed description of organizational learning that appears compatible with double-loop learning is that by Conlon, Havlisch, and Porter,[69] 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,"[69(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."[69(p12)]

In most other instances, the studies imply that the safety culture has been improved. 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'"[77(p1209)]; "a focused, hospital-wide effort to improve the system of medication preparation, processing, and delivery"[79(p217)]; indicators of a positive safety culture[108]; creating a safety culture through a multidisciplinary effort involving a combination of interventions[91]; "changing the error reporting form to make it less punitive"[71(p492)]; "developing an awareness of error and a safety culture with less emphasis on the 'blame' approach"[72(Table 2,p658)]; and "successive SHOT reports have encouraged open reporting of adverse events and near-misses in a supportive, learning culture."[93(p281)]

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

What the studies do indicate, however, are potential facilitators of organizational learning and/or potential barriers in the absence of such facilitators (see Table 1 and online Appendix C).[40,76]

First, we noted earlier that according to Argyris and Schön's theory, psychological safety is likely to be important to double-loop learning. Eighteen of the 35 studies regularly and repeatedly refer to the need to make reporting less punitive, recommending anonymous, confidential reporting and the absence of a "blame culture" or fear of reprisals. Those studies using medical definitions of error 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 raise awareness of the need for learning to be the function or output of an IRS.[67,70,76,88] This is contrasted with IRSs being driven by an "audit culture" whose agenda may be (perceived to be) the reassertion of management control and with the

Table 1. Summary of Potential Facilitators of Double-Loop Learning

Facilitator	Characteristics	Studies
Psychological safety	Nonpunitive; less punitive reporting; anonymous, confidential; absence of "blame culture" and fear of reprisals.	Anderson et al. 2013[40]; Conlon, Havlisch, and Porter 2013[69]; Cooke, Dunscombe, and Lee 2007[89]; Elhence et al. 2010[78]; Frey et al. 2002[99]; Jayaram et al. 2011[91]; Kalapurakal et al. 2013[103]; Kivlahan et al. 2002[104]; NHS Quality Improvement Scotland 2006[88]; Nicolini, Waring, and Mengis 2011[76]; Pierson et al. 2007[70]; Ross, Wallace, and Paton 2000[71]; Savage, Schneider, and Pedersen 2005[74]; Spigelman and Swan 2005[72]; Stainsby et al. 2006[93]; Stainsby et al. 2004[94]; Takeda et al. 2003[81]; Weissman et al. 2005[68]; Wong, Kelly, and Sullivan 2013[67]
Focus on learning	Learning as the function/output (vs "audit culture," etc.); actual, genuine focus on learning (vs rhetorical/espoused); discrepancies, emotion, etc., allowed.	NHS Quality Improvement Scotland 2006[88]; Nicolini, Waring, and Mengis 2011[76]; Pierson et al. 2007[70]; Wong, Kelly, and Sullivan 2013[67]

Table 1. *Continued*

Facilitator	Characteristics	Studies
Reflects complexity		
Cross-departmental/ organizational/ professional	Multiagency; interorganizational; multi- or cross-disciplinary; cross-functional; no silos and barriers between departments.	Conlon, Havlisch, and Porter 2013[69]; Cooke, Dunscombe, and Lee 2007[89]; Herzer et al. 2012[100]; Jayaram et al. 2011[91]; Kivlahan et al. 2002[104]; Pierson et al. 2007[70], Rebibo et al. 2004[96], Stainsby et al. 2006[93]
Multiple interventions	Holistic/systemic approach; complementary, systemwide interventions vs single interventions in isolation (eg, training).	Callum et al. 2001[77]; Cooke, Dunscombe, and Lee 2007[89], Ross, Wallace, and Paton 2000[71], Roughead, Gilbert, and Primrose 1999[106], Sullivan and Ghroum 2013[90]
Local and participative	Built within the context (vs imposed); locally designed vs centrally or externally designed; participants involved in problem-solving instead of being hierarchical, in the hands of specialists.	Conlon, Havlisch, and Porter 2013[69]; Cooke, Dunscombe, and Lee 2007[89]; Herzer et al. 2012[100]; NHS Quality Improvement Scotland 2006[88]; Nicolini, Waring, and Mengis 2011[76]

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 its complex, emergent nature. Our review did not find one paper that explicitly examined 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 organizational processes that eventually result in adverse events.[109] An important point made by a number of studies is that IRSs are most effective when they are part of wider quality improvement programs.[69,74,103,106] Being embedded in, or linked to, organizationwide interventions may be one way to overcome the difficulty of achieving organizational learning in a complex, multiprofessional setting. Several studies refer to the need for an IRS to be cross-departmental, multiprofessional, or interorganizational.[69,70,89,91,93,96,100,104] Others emphasize that multiple interventions are more likely to be effective than single interventions.[71,77,91,106] Thus Callum and colleagues[77] comment on the ineffectiveness of small-group educational sessions if used in isolation; and Ross, Wallace, and Paton[71] underscore the need for an intervention to be complemented by other changes. Finally, some studies stressed 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.[71,76,89,100]

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, in order to explore 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 reviews may have greater effectiveness in identifying clinical incidents than IRSs do. Moreover, there was very little focus on resource utilization, with only 2 studies looking at this issue.[86,92] Therefore, there is no clear evidence that IRSs are more cost-effective than other systems.

Our 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 framework.[6] At the macro level of organization, we found evidence that IRSs could trigger single-loop learning primarily in the context of prescribing drugs, by means of action forcing such changes as the withdrawal of certain medicines from the market. There was also some limited evidence of changes to processes and outcomes at the micro and meso levels triggered by the dissemination of IRS data on adverse events arising from blood transfusions and the use of flucloxacillin.

At the micro and meso levels of organization, few studies 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.[71,100,103,106] Furthermore, our synthesis supports Waring's[14] argument that centralized systems at the micro and meso levels, such as those used in UK hospitals, might not yield the depth of learning anticipated by policymakers. Consistent with this, our review indicates that meso-level changes may have little impact at the micro level. At the intraorganizational micro and meso levels, where there is ownership of incidents and clinical commitment to safety improvement, settings and processes can be changed 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 most effective when used and owned by clinical teams or communities of practice[110] in specific departments rather than at the wider organization level. Such communities have been shown to be nurtured by opportunities for interaction and communication[110] and are likely arenas for the development of reciprocal ties, shared commitment to group goals, trust, and the psychological safety required for organizational learning.[110]

Notably, the absence of standard, agreed universal definitions of adverse events or near misses and the lack of clear definitions and measurement of outcomes make it difficult to compare, identify, and correct errors or to evaluate reliably the impact of doing so. Without a 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 transfusions and macro-level drug reporting, IRSs were more likely to improve safety. In contrast, although anticoagulation is an area of high risk, because IRSs relating to

anticoagulant therapy did not have agreed-on definitions of harm, the aggregation of information from various databases was problematic. Another factor impeding organizational learning was the absence of a feedback loop; staff did not always receive feedback about incidents reported.[13,40]

Our review identified both potential facilitators of and barriers to double-loop learning and indicates that in order 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 "solvable" errors. To deal with such complexity, an IRS needs to work across functional, organizational, and professional boundaries and to 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 toward organizational learning. The resulting action is likely to require multiple, complementary interventions. Studies indicate that interventions used in isolation (eg, training) are unlikely to be effective. Employees need to be confident that "learning" is the authentic purpose and raison d'être 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- and mesodepartment levels, provided they retain the main principles.[111] 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.

Limitations of Our Study

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

Our choice of Argyris and Schön's theory means that we adopted a cognitive rather than a social perspective on organizational learning. We have acknowledged that social theories of organizational

learning may account for the way that organizational learning is likely to emerge through complex processes that involve multiple actors and multiple agencies. Accordingly, this is a promising area for future research. Nevertheless, we believe that Argyris and Schön's theory is suitable for our purpose given the aims of our article.

Conclusions

Overall, the studies we reviewed did show some evidence that IRSs can lead to single-loop learning, that is, corrections to errors in procedures and improvements in techniques. We found little evidence, however, that IRSs ultimately improve patient safety outcomes or that single-loop learning changes were sustained, although this may be a consequence of measurement difficulties[65,108] and the need for agreed-on 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 organizations' micro and meso levels, IRSs are most effective when combined with other improvement efforts as part of wider quality improvement programs. This supports the 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 a change of mind-set.

In sum, one way of improving both the efficiency and effectiveness of IRSs might be to embed them as part of wider safety programs and devolve their control and management from centralized hospital departments to clinical teams. The results of our study suggest that health care organizations should consider carefully the opportunity costs of 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, more longitudinal research is required to explore the impact of IRSs on patient safety outcomes and how and/or 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 regard to organizational learning.

References

1. Berwick D. *A Promise to Learn—A Commitment to Act: Improving the Safety of Patients in England*. London: National Advisory Group on 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 Manage 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. *To Err Is Human: Building a Safer Health System*. Washington, DC: National Academy Press; 2000.
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 Quality Assurance*. 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.

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: OECD; 2013. http://www.oecd.org/els/health-systems/Health-at-a-Glance-2013.pdf. Accessed June 10, 2014.

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*. Reading, MA: Addison-Wesley; 1978.

24. ASRS. ASRS program briefing. 2014. http://asrs.arc.nasa.gov/docs/ASRS_ProgramBriefing2013.pdf. Accessed June 9, 2014.

25. International Association of Oil & Gas Producers. Health and safety incident reporting system users' guide, 2010 data. 2010. http://www.ogp.org.uk/pubs/433.pdf. Accessed September 10, 2014.

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. Accessed September 10, 2014.

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

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

31.   ASRS. ASRS: the case for confidential incident report-
      ing systems. 2006. http://asrs.arc.nasa.gov/docs/rs/60_Case_
      for_Confidential_Incident_Reporting.pdf. Accessed September
      10, 2014.

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 sys-
      tems. *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. Accessed June
      10, 2014.

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 develop-
      ment of disasters. *Adm Sci Q*. 1976:378-397.

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

39.   Seckler-Walker J, Taylor-Adams S. Clinical incident reporting.
      In: Vincent C, ed., *Clinical Risk Management: Enhancing Patient
      Safety*. 2nd ed. London: BMJ Books; 2001.

40.   Anderson JE, Kodate N, Walters R, Dodds A. Can incident re-
      porting improve safety? Healthcare practitioners' views of the
      effectiveness of incident reporting. *Int J Quality Health Care*.
      2013;25(2):141-150.

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

42. Cyert RM, March JG. *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice-Hall; 1963.

43. Easterby-Smith M, Araujo L, Burgoyne J. *Organizational Learning and the Learning Organization: Developments in Theory and Practice*. Thousand Oaks, CA: 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 research. *Int J Manage 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. *Manage Learning*. 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. *Manage Learning*. 1998;29(3):273-297.

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

51. Freidson E. *Professionalism, the Third Logic: On the Practice of Knowledge*. Chicago, IL: 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 Quality 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 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*. San Francisco, CA: Jossey-Bass; 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*. Berkshire, United Kingdom: McGraw-Hill International; 2007.

65. Rosenthal J, Booth M. *Maximizing the Use of State Adverse Event Data to Improve Patient Safety*. Portland, ME: 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. 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.

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

69. Conlon P, Havlisch R, Porter C. Using an anonymous web-based incident reporting tool to embed the principles of a high-reliability organization. In: Henriksen K, Battles J, Keyes MA, et al., eds., *Advances in Patient Safety: New Directions and Alternative Approaches. Vol. 1: Assessment*. Rockville,

MD: Agency for Healthcare Research and Quality; 2008. http://www.ncbi.nlm.nih.gov/books/NBK43630. Accessed October 16, 2014.

70. 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.

71. 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.

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

73. 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.

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

75. 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.

76. 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.

77. 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.

78. Elhence P, Veena S, Sharma RK, Chaudhary RK. Root cause analysis of transfusion error: identifying causes to implement changes. *Transfusion (Paris)*. 2010;50(12, pt. 2):2772-2777.

79. 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.

80. 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.

81. 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.

82. 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.

83. 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.

84. 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.

85. 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.

86. 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.

87. 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.

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

89. 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.

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

91. 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.

92. 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.

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

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

95. 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.

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

97. 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.

98. 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.

99. 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.

100. 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 Safety Jt Comm Resour*. 2012;38(8):339.

101. 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.

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

103. 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.

104. 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.

105. 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.

106. 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.

107. 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.

108. 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.

109. 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.

110. Wenger E. Communities of practice and social learning systems. *Organization*. 2000;7(2):225-246.

111. 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.

Supplementary Material

Additional supporting information may be found in the online version of this article at http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1468-0009:

Online Appendix A. Data Extraction Table: Comparison of Systems Papers
Online Appendix B. Data Extraction Table: Evidence of Changes in Setting, Process, and Outcomes
Online Appendix C. Interpretation of Studies as Single- or Double-Loop Learning