Enhancing Patient Safety through Organizational Learning: Are Patient Safety Indicators a Step in the Right Direction?

Peter E. Rivard, Amy K. Rosen, and John S. Carroll

Objective. To assess the potential contribution of the Agency for Healthcare Research and Quality Patient Safety Indicators (PSIs) to organizational learning for patient safety improvement.

Principal Findings. Patient safety improvement requires organizational learning at the system level, which entails changes in organizational routines that cut across divisions, professions, and levels of hierarchy. This learning depends on data that are varied along a number of dimensions, including structure-process-outcome and from granular to high-level; and it depends on integration of those varied data. PSIs are inexpensive, easy to use, less subject to bias than some other sources of patient safety data, and they provide reliable estimates of rates of preventable adverse events.

Conclusions. From an organizational learning perspective, PSIs have both limitations and potential contributions as sources of patient safety data. While they are not detailed or timely enough when used alone, their simplicity and reliability make them valuable as a higher-level safety performance measure. They offer one means for coordination and integration of patient safety data and activity within and across organizations.

Key Words. Patient safety, patient safety indicators, high-reliability organizations, organizational learning, administrative data

Heightened attention is being paid to the quality and safety of medical care. This attention brings with it an increased demand for data on the quality and safety performance of the individuals and organizations that provide health care (Kohn, Corrigan, and Donaldson 2000; Hurtado, Swift, and Corrigan 2001; Pronovost et al. 2004). While health care is behind other industries in its ability to measure safety (Gaba 2003), evidence-based measures are being developed and improved (Romano et al. 2003).

The Patient Safety Indicators (PSIs), developed by the Agency for Healthcare Research and Quality (AHRQ) and the UCSF–Stanford

Evidence-Based Practice Center (EPC) (AHRQ 2005), are in increasing use by hospitals, health systems, and others who monitor hospital patient safety performance (Miller et al. 2001; Romano et al. 2003; HealthGrades 2004; Rosen et al. 2005). The PSIs were developed as a tool for tracking and improving patient safety (AHRQ 2005). The PSIs use administrative data (hospital discharge records), including ICD-9 diagnosis and procedure codes, to identify potential in-hospital patient safety events (McDonald et al. 2002). The PSIs are expressed as rates: the numerator is the number of occurrences of the outcome of interest and the denominator is the total population at risk (AHRQ 2005). PSIs track possible surgical complications and other nosocomial events, screening for "potential problems that patients experience resulting from exposure to the health care system, and that are likely amenable to prevention by changes at the level of the system" (McDonald et al. 2002).

Harm to patients is often the result of system-level failures as well as individual error (Kohn, Corrigan, and Donaldson 2000). Improving patient safety therefore requires learning by groups and organizations as well as by individuals. We argue in this paper that, in order for adoption of safety performance measures and indicators such as the PSIs to lead to safety improvement, these measures must contribute to patient safety learning at the organizational level, and not merely to the evaluation of providers. We integrate findings from research on organizational learning, quality and safety measurement, and the AHRQ PSIs, to propose a framework for integrating multiple sources of patient safety data in support of organizational learning for patient safety improvement. We address the specific question: How can the AHRQ PSIs facilitate hospitals' organizational learning?

First, we describe aspects of a normative model of organizational learning for safety improvement in health care. Second, we compare sources of patient safety data and describe the role of data and measurement in organizational learning for patient safety improvement. Third, we assess the AHRQ PSIs in this context and propose a framework for integrating a variety of data and measures to support patient safety improvement at the organizational level. Following our description of this framework, we conclude with a discussion of some implications of our findings for research and practice.

Address correspondence to Peter E. Rivard, M.H.S.A., Research Associate, Center for Health Quality, Outcomes and Economic Research, Bedford VAMC, 200 Springs Road (152), Bedford, MA 01730. Amy K. Rosen, Ph.D., Senior Research Scientist, is with the Center for Health Quality, Outcomes and Economic Research, Bedford, MA. John S. Carroll, Ph.D., is with the MIT Sloan School of Management, Cambridge, MA.

ORGANIZATIONAL LEARNING FOR SAFETY IMPROVEMENT

Risk and safety are properties of whole systems as well as properties of individual actors and their actions (Weick 1993a; Edmondson 1996; Reason 1997, 2000; Carroll, Rudolph, and Hatakenaka 2002). Patient care is delivered by systems as well as by the individuals in them; therefore, patient safety improvement must entail group and organizational learning as well as individual learning (Edmondson 1996; Gaba 2000a). For example, when the wrong "Mary Jones" undergoes an invasive procedure, and any of several individuals might have averted the error, what is the learning that will most effectively and efficiently prevent this scenario from recurring? Assuring that a few individuals are trained and responsible for verifying patient identity will not address larger issues, e.g., a system where key roles are chronically under excessive time pressure, or an organizational culture where patient identification is always somebody else's job and error is always somebody else's fault. Organizational learning is manifested by new or modified organizational routines (Levitt and March 1988). Organizational routines can be defined as repetitive, recognizable patterns of interdependent actions involving multiple actors (Feldman and Pentland 2003). Therefore, one key to the distinction between individual and organizational learning is that organizational learning becomes apparent and measurable in the enactment of organizational routines.

Safety Improvement and High-Reliability Organizations (HROs)

Organizational learning for safety improvement entails changes in organizational routines that affect risk and safety (Weick 1987). Not all organizational learning is system-level learning: for example, Tucker and Edmondson (2003) distinguish first-order from second-order learning in organizations. First-order learning occurs at the front lines, where individuals and groups change routines to solve problems. Those solutions may maintain safety, but they may also work around, and mask, underlying system problems (Spear and Schmidhofer 2005). To the extent that patient care routines involve multiple actors, professions, and divisions within an organization, changing routines to reliably improve patient safety is often a matter of second-order learning, which occurs across and between those groups and divisions (Tucker and Edmondson 2003).

When it comes to safety improvement, the complex task of system-level organizational learning faces additional challenges related both to the data

themselves and to the cognitive and cultural constraints on communication and interpretation of the data. Data-related challenges include limited availability of organization-specific data on risk and safety, chronological sequencing of events (e.g., in inpatient hospital care, knowing whether a condition was present on admission), and in health care, issues related to coding, reliability, and validity. Safety data differ from other types of feedback on organizational performance, such as volume or quality: most types of adverse events are infrequent and the ideal rate is low or zero (Reason 1997; Gaba 2000b). Infrequent adverse outcomes, e.g., some surgical complication rates, make it impossible to measure "true" rates because it takes multiple years of data or the aggregation of many providers in order for the data to be actionable (Mant 2001).

In addition to data challenges, there are cognitive and cultural constraints on collecting, communicating about, analyzing and interpreting safety data. People often prefer not to acknowledge or discuss adverse events to which they may have contributed; they wish to avoid embarrassment, blame, or sanction (Weick 1993b; Edmondson 1996). The health care professions' culture of perfection and the persistence of cultural norms that support "blaming and shaming" tends to enhance those impediments (Leape 1994; Sharpe and Faden 1998; Kohn, Corrigan, and Donaldson 2000).

The work of health care organizations is often both complex and high-risk, as is the work of organizations categorized as HROs. HROs can be defined as: complex organizations engaged in high-hazard activity, that continually face unexpected developments, yet "manage to have fewer than their fair share of accidents" (Weick and Sutcliffe 2001). Among the population of HROs are certain nuclear power plants and air traffic control systems. HROs are noteworthy for their success in confronting the data challenges and cultural and cognitive challenges to safety improvement described above, and in achieving second-order, whole-system organizational learning. Although the literature is mixed about whether hospitals can be characterized as HROs (Roberts and Rousseau 1989; Reason 2000), the similarities are strong enough to justify drawing on HRO research in developing a normative model of organizational learning for safety improvement in health care (Gaba 2000b).

Safety researchers already have encouraged health care organizations to adopt some of the safety learning practices of HROs. Many hospitals are actively engaged in practices similar to those of HROs (van der Schaaf 2002; Gaba 2003). Examples include: implementation of methods developed in the aviation industry for improving communication; changing the environment so that event reporting is conducted more routinely (Thomas and Helmreich 2002; Woolever 2005); and reduction of the cultural impediments that

promote shaming and blaming by leaders (Pate and Stajer 2001). However, while HROs are known for their vicarious learning from the experiences of other organizations (Weick, Sutcliffe, and Obstfeld 1999), health care organizations appear more resistant to inter-organizational quality and safety learning (Lichtman et al. 2001; Mills, Weeks, and Surott-Kimberly 2003). Also, while the successes of HROs are compelling, adapting the HRO model of safety to health care may present some significant cultural hurdles: for example, health professionals' cultures of autonomy and their notions of what constitutes high performance may need to change (Amalberti et al. 2005).

The high safety performance of HROs, and their success in meeting the challenges to organizational learning about safety described above, is attributable to their distinctive structural and cultural characteristics. HROs approach safety learning with an organizational culture that frames safety and learning in particular ways and influences both the choice of data to collect and the meaning ascribed to the data. One such frame is the HRO's *preoccupation with failure:* HROs look carefully at actual and potential failures, rather than downplaying and minimizing them (Reason 2000; Weick and Sutcliffe 2001). HROs actively imagine possible future failures. Preoccupation with failure helps the HRO address the scarcity of data on safety failures by focusing attention on the data that are available.

A second HRO safety-learning frame is *reluctance to simplify* interpretations of safety issues. This is the HRO's recognition of a high degree of interdependency among the organization's dissimilar parts. Because of this interdependency, a problem in one location may reflect one or several problems "upstream" and/or may contribute to problems "downstream" (Weick and Sutcliffe 2001). Reluctance to simplify helps counter cognitive biases toward the most obvious and close-at-hand explanations for adverse events. It also addresses cultural resistance to safety learning by insisting on explanations that go beyond individual blame.

In an HRO, the systemic nature of risk and safety is confronted by these learning frames, leading to a holistic and systemic approach to safety learning. However, understanding how all the pieces of an HRO work doesn't necessarily lead to understanding how the whole system works; and reliable parts do not necessarily constitute a safe system (Perrow 1984; Bierly and Spender 1995). Similarly, in health care, understanding process-outcome links and assessing the preventability of a particular outcome is rarely straightforward. Therefore, in addition to multiple sources and interpretations of data, an HRO seeks data across several levels of analysis and it seeks to integrate those data (Weick and Roberts 1993).

DATA TO SUPPORT ORGANIZATIONAL LEARNING FOR PATIENT SAFETY IMPROVEMENT

The AHRQ PSIs

The PSIs were developed by AHRQ and the Stanford–UCSF EPC through literature review, physician panels, and extensive statistical testing, with the goal of making the most effective use of administrative data to detect possible adverse events (McDonald et al. 2002). Table 1 provides examples of PSIs, their definitions, and summaries of expert panel reviews of the indicators. The PSI definitions attempt to limit PSI events to those that are reasonably preventable. For example, the definition for the iatrogenic pneumothorax PSI excludes certain types of hospitalizations where the patient's condition or treatment places the patient at much higher risk for pneumothorax.

The AHRQ PSIs are part of a broad array of patient safety data tools. In this section, we describe that array in light of the data requirements for organizational learning and specify the place of the PSIs within that array. The following criteria, sources, and dimensions of patient safety data describe differences that are important to the utility of the data and to how the data may contribute to organizational learning (see Figure 1 for an illustration of the array of data and their uses).

Criteria for Evaluating Patient Safety Data and Their Sources

A choice among safety data alternatives is likely to involve tradeoffs among different criteria including validity, reliability, cost of acquisition and analysis, and how actionable the findings are. Validity itself is multidimensional. While face validity is important for data to be meaningful to users, construct validity indicates the extent to which the intended phenomena are being measured. Within construct validity, a more sensitive measure will identify more "true" occurrences, possibly at the risk of false positives, whereas a measure with high positive predictive value (PPV) may yield some false negatives, but the occurrences identified are highly likely to be "true." High PPV is an important criterion for patient safety data that are used for organizational learning because a low rate of false negatives can contribute to a climate of fairness, trust, and open discussion of safety.

Validation of the AHRQ PSIs is still in its early stages. Despite the care taken in developing the PSI definitions, the PSIs retain some likelihood of both false positives and false negatives (Zhan and Miller 2003b; Gallagher, Cen, and Hannan 2005; Rosen, Rivard et al. 2006). An additional limitation to PSIs' sensitivity is that they are intermediate outcome measures and do not capture

Table 1: Observed Hospital-Level PSI Rates*, October 1, 2000–September 30, 2001, 127 Veterans Health Administration Facilities for Six Selected Nonobstetric Indicators

#	Name	Patient Safety Indicator Definition	Hospital-Level Occurrences per 1,000 Discharges at Risk			AHRQ/EPC Expert Panel Evaluation of Indicators
			Lowest	Median	Highest	
3	Decubitus ulcer	Cases of decubitus ulcer per 1,000 discharges with length of stay≥5 days. *Excludes patients with:* paralysis, diseases/disorders of skin, subcutaneous tissue, or breast; excludes obstetrical patients; and patients admitted from a long-term care facility	0.00	13.00	70.25	Very favorable; best used for overall rate rather than case finding
4	Failure to rescue	Deaths per 1,000 patients having developed specified complications of care during hospitalization. *Excludes patients:* age 75+, neonates, and patients admitted from long-term care facility, or transferred to or from other acute care facility	0.00	151.52	333.33	Favorable; however, it combines different aspects of quality/safety
6	Iatrogenic pneumothorax	Cases of iatrogenic pneumothorax per 1,000 discharges. *Excludes:* patients with trauma, thoracic surgery, lung or pleural biopsy, cardiac surgery, and obstetrical patients	0.00	0.91	3.65	Important preventable complication; however, rates may overstate actual medical error
7	Infection due to medical care	Cases of secondary infection after infusion, injection, transfusion, vaccination; cases of reaction from internal device or graft, per 1,000 discharges. *Excludes:* immunosuppressed and cancer patients	0.00	1.87	6.70	Favorable, best used as an overall rate because of the existence of an underlying rate even with the highest quality care
12	Postoperative PE/DVT	Cases of deep vein thrombosis or pulmonary embolism per 1,000 surgical discharges. *Excludes:* obstetrical patients	0.00	9.63	34.64	Important preventable complication; this indicator highly rated
15	Accidental puncture or laceration	Cases of accidental cut or laceration during procedure per 1,000 surgical discharges. *Excludes:* obstetrical patients	0.00	1.42	12.36	Favorably rated, although rates may overstate actual medical error

*Observed rates are raw rates, not risk adjusted.

Figure 1:   Patient safety data sources and uses, along the course of care delivery
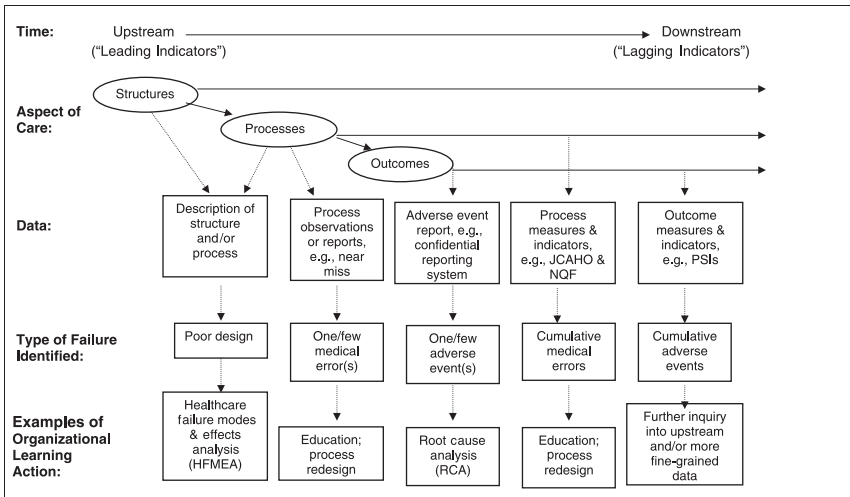
posthospitalization outcomes such as complications, readmission, or death. Nonetheless, studies have consistently found statistically significant associations between the PSIs and increased mortality, length of stay, and cost, although they were not able to isolate the extent to which these outcomes were attributable to actual adverse events (Miller et al. 2001; Zhan and Miller 2003a; Rosen et al. 2005). Results of a VA study measuring the sensitivity of the AHRQ PSIs using chart-based data from the VA National Surgical Quality Improvement Program as the "gold standard" suggest that certain PSIs have moderate-to-high levels of sensitivity and PPV (Rosen et al. 2005; Rosen, Zhao et al. 2006). Another study found that, because occurrence of a PSI event is determined by data from one hospitalization only, the "infections due to medical care" PSI misses cases where infections first appear in administrative data on short-term readmissions (Gallagher, Cen, and Hannan 2005).

Because the PSIs generate rates of events, they offer opportunities for comparisons across time periods, providers, and systems. Therefore, it is important to examine their reliability. The PSIs demonstrate good reliability, based on the consistent results found across studies (Rivard et al. 2005; Rosen et al. 2005) and across years (Rosen, Rivard et al. 2006). However, their reliability could prove to be a concern over time: to the extent that rewards and/or sanctions are associated with PSI rates, incentives are created to code

in ways that reflect desired rates on performance indicators. Some PSIs may be amenable to "downcoding" (i.e., adopting a higher severity threshold for coding events that are tracked by the PSIs). In addition, there may be challenges associated with implementing PSIs consistently across multiple facilities (e.g., in a health care system or region) and with making the best use of available data (Rivard et al. 2005).

Table 1 shows low, median, and high occurrence rates of six PSIs in a population of 127 Veterans Administration hospitals for Fiscal Year 2001 (10/1/00–9/30/01). These PSIs, selected for inclusion here primarily because of their normally higher occurrence rates compared with other PSIs, suggest that at least some of the PSIs can provide meaningful and timely event rate and trend data. The data in Table 1 are from a VA-funded study that examined the feasibility of using VA administrative data and the PSIs to identify potential patient safety events in the VA (Rosen et al. 2005).

As facilities and systems seek to expand their array of safety data, cost of acquisition and cost of analysis are important criteria to be considered. Because they are based on administrative data that facilities already collect for other purposes, PSIs are a low-cost safety data source. Finally, how actionable are the data—i.e., do the data shed light on the organization's achievement of important goals, and does the organization have the will and capacity to act on the findings? How actionable the PSIs are will vary by facility.

Dimensions of Patient Safety Data

Donabedian's (1980) quality paradigm—structure, process, and outcome—serves to illuminate some important distinctions among safety data. Donabedian's concept of structure is particularly relevant to organizational learning. Structure encompasses the more stable characteristics of the system of care delivery, including not only staffing, equipment, and facilities, but also how those elements are organized to deliver care. Structure includes formalized organizational routines, such as the process of passing patient information across caregiver work shifts. System improvement is change in structure; structure data are therefore frequently essential to system-level organizational learning. HRO research suggests that HROs are likely to focus "upstream" in the chain of organizational action in their search for data on risk and safety. To borrow a concept from economics: HROs look for "leading indicators" of risk that are predictive of potential adverse events, thus focusing the HRO on process and structure data (Roberts and Rousseau 1989).

Research on quality improvement suggests that process measures are more sensitive than outcome measures to actual differences in quality across providers and/or time (Mant and Hicks 1995; Mant 2001). Process measures of health care quality are also likely to be more actionable in that they are easier to interpret (Mant 2001), in part because accountability is clearer (Rubin, Pronovost, and Diette 2001; Pronovost et al. 2004). At the same time, "lagging indicators" such as outcome measures tend to have more face validity than process measures and they are more meaningful in public discussions of patient safety (Kasandjian 2003; Pronovost et al. 2004). Nonetheless, use of outcome measures carries the risk of outcome bias (e.g., if users of a safe process happen to experience adverse outcomes due simply to random variation, purely outcome-driven learning might lead to adoption of a less safe process) (Caplan, Posner, and Cheney 1991); and outcome measures prompt search for safety problems only where things have already gone wrong, whether or not the outcome was preventable (Henriksen and Kaplan 2003).

Patient safety data typically identify *medical errors, near misses, or adverse events*. Medical errors lead to adverse events or to near misses. Near miss data are valuable for timely organizational learning because they may act more as leading indicators than do adverse event data. However, in order to facilitate effective safety learning, near miss data must be interpreted from a preoccupation-with-failure perspective that seeks to improve understanding of the riskiness of structures and processes, and not from the perspective of "it's a good thing we caught that one in time" (Dillon and Tinsley 2005; Harris, Pace, and Fernald 2005).

HROs' reluctance to simplify, as well as their deference to expertise, suggests the need for *granular data* that support a comprehensive understanding of the processes of care that are to be improved. However, system learning requires integration, and as we discuss in more detail below, *more holistic* data such as the PSIs can serve that purpose. Another important distinction is between *measures and indicators*. The PSIs are indicators: they point to events that are potentially preventable. Indicators have a probabilistic relationship to the focal events, whereas measures are more definitive, but measures may be more costly to obtain.

A final dimension to consider is *sample size*. Rare or one-time occurrences, such as sentinel events, can provide richly detailed descriptions of care structures and processes that are needed for system improvement (March, Sproull, and Tamuz 1991). However, cumulative indicators or measures based on large administrative databases, such as the PSIs, have the potential for reliability and statistical significance. In addition, excessive focus on the worst

outcomes, such as death and severe disability, could lead to insufficient attention to prevention of minor disabilities, despite the fact that the latter may represent a larger portion of the excess resource use resulting from adverse events (Runciman, Edmonds, and Pradhan 2002).

Sources of Patient Safety Data

Potential sources of patient safety data include: direct observation; medical records; administrative data (which the PSIs use); reporting of errors, near misses, and adverse events; and malpractice claims. The following comparison of data sources draws on the work of Michel et al. (2004) and Thomas and Peterson (2003). *Direct observation* is costly but provides fine-grained, potentially valid, and useful data on processes of care. Medical records, albeit costly to review, are generally a detailed and valid source of adverse event data. *Administrative data* are low cost and generally complete. The data are routinely collected and, while inaccuracy and variability of coding practices always characterizes administrative data to some degree (Romano et al. 2002; Zhan and Miller 2003b), they generally have few data elements missing. Incentives affecting administrative data have related more to maximizing reimbursement than to recording (or choosing not to record) safety events historically. Administrative data are a "bigger net" than event reporting for catching possible safety events as they cover many more episodes of care.

    Event reporting can be a timely and low-cost means of acquiring patient safety data. It relies on an individual's decision to report, which requires observing and interpreting an occurrence as a safety issue, as well as overcoming any existing practical and cultural barriers to reporting. Event tracking that relies on medical records is likely to be more sensitive than tracking based on reporting (Classen et al. 1991). *Malpractice claims analysis* can offer rich data, but it is subject to hindsight bias and significant selection factors in what is reported. Hindsight bias is the tendency to conflate correlation with causality, e.g., to assume that the acts that preceded an adverse outcome were unsafe acts (Henriksen and Kaplan 2003).

    Given the relatively low incremental cost of generating PSI reports from existing administrative data, the PSIs offer a useful additional data source. The amount of information they contribute to the understanding of particular events, however, is limited. First, the PSIs may or may not point to actual preventable events. Second, by providing only a count and rate of possible preventable events, by themselves they offer relatively little toward understanding of the cause of the problem. PSIs offer multiple measures of safety in

that they cover a range of safety concerns: for example, some are predominantly medical, such as decubitus ulcer, while others focus primarily on surgical care, such as postoperative sepsis. Because PSIs are not driven by one individual's perspective on safety events, they are well suited to the need for multiple interpretations and for a balanced discussion of alternative interpretations.

PSI rates that accrue quarterly or annually are a relatively slow feedback loop. Particularly with regard to surgical events, recording events at the point of care would provide quicker feedback (Bent et al. 2002). In comparison with PSIs, collecting near-miss data moves the organization even further upstream (see Figure 1) in the detection of risk. However, while PSIs do not serve as "leading indicators" for active problems that the organization can—and does—track by other more rapid-cycle methods, PSIs can serve as timely indicators for latent problems that the organization is not otherwise tracking.

Integrating Patient Safety Data

The previous section described a broad array of patient safety data. Choices among the measures are a matter of tradeoffs among such factors as timeliness, validity, reliability, cost, and usefulness. Also important, however, are balance and integration among the forms of patient safety data used in an organization. This is important because of at least three sets of considerations. First, individuals and groups in different parts of the organization have differing accountabilities for safe structures, processes, and outcomes of care; different accountabilities call for different safety data. Senior leadership is likely to experience accountability for high-visibility aspects of safety such as sentinel events, mortality rates, and other measures easily understood by the public; middle managers, service chiefs, and department chairs may have a production-line quality perspective on safety; and individual clinicians are often concerned with the impact of particular safety issues on their own work and livelihood (Amalberti et al. 2005). Integration for organizational learning is challenging. Earlier we made the distinction between first- and second-order change (Tucker and Edmondson 2003); in fact, the organizational routines that affect patient safety exist at several levels of the organization. In addition to senior and middle management and front line staff, a distinct level is represented by the clinical microsystem, which has been proposed as an integrating construct that captures the structures of care at the midlevel between an individual organizational unit and the larger organization (Mohr and Batalden 2002). Measures relevant to one level, e.g., senior management and macrosystems, may not seem relevant to another level, e.g., those concerned with microsystem safety.

Second, the players in different parts of the system with different accountabilities also need to communicate and coordinate with each other in order for organizational learning and system improvement to happen. This needs to occur vertically, up and down the hierarchy, and horizontally, across functions and specialties. This integration can be supported by system-level data, such as outcomes of complex care processes, and data that are meaningful to members of different professions and microsystems. The goal of a system should be to maximize its desired whole-system outputs, not the outputs of its components (Schyve 2005); an outcome measure reflects the whole system of care processes that produced the outcome (Mant 2001).

Third, data need to be linked across structures, processes, and outcomes. Perhaps most importantly, because much of the uncertainty in health care is due to the ongoing challenge of linking processes of care with health-related outcomes, quality and safety improvement efforts need to determine process-outcome linkages (Kasandjian 2003). Structure–process links also are important, however: in one VA study, an intervention for improvement in surgical quality that combined structural change (coordination, in the form of protocols and pathways) with process monitoring and feedback resulted in the highest quality of care (Young et al. 1998).

The PSIs have a role to play in an integrated approach to patient safety data. They can serve the needs of senior managers for relatively straightforward outcome data with high face validity. They provide an opportunity to track facility-level indicators over time, and to benchmark across facilities and systems. They can be "rolled up" to produce system-wide rates within a system of many facilities, and they can be "drilled down" to focus further inquiry within a facility if the PSI rates are unfavorable. For example, if postoperative sepsis rates are higher than expected, further analysis can determine whether a concentration of cases is associated with one subspecialty or one operating room. They can be linked with process measures, e.g., comparing changes in sepsis rates with changes in preoperative antibiotic therapy. PSIs can be linked with structural measures, as well: for example, a change in decubitus ulcers could be associated with structural changes such as nursing staffing.

Safety performance measures such as PSIs have a role to play in organizational learning that is different from incident-specific data such as event reports. Huber's (1991) review of organizational learning research makes a distinction between experiential learning and "learning from searching and noticing." The former is learning from the results of the organization's own actions, including its experiments and failures; the latter is learning from scanning the environment, engaging in focused search for new opportunities

and solutions, and measuring performance. Learning from experiments and failures is more reactive and is subject to cognitive limitations such as hindsight bias; whereas "search" suggests a role for performance measures in a more proactive approach to performance improvement. While PSIs are fundamentally lagging indicators, they can serve a proactive "search" function by identifying safety issues that hitherto were unnoticed or ignored. Also, as performance indicators, PSIs can be a useful tool of organizational learning and change when used to normalize the conversation about failure and medical harm, to bring it out of the shadows of sanction and blame and into the light of inquiry for learning and improvement. This is consistent with the HRO preoccupation with failure: in an HRO, it is normal to discuss risk and failure. As indicators and not measures, and as trends that do not in and of themselves identify specific cases, PSI rates that are deemed unsatisfactory by their users are akin to small failures that do not carry the emotional impact of large failures. Small failures are good sources of learning because the organizational defenses around them are not so rigid (Sitkin 1992).

Using the PSIs for Organizational Learning

While there appears to be no published research to date on hospitals' use of PSIs for safety improvement, there is evidence that the PSIs are being used for hospital assessment and benchmarking. For example, the CMS Premier Hospital Quality Incentive Demonstration project is examining quality and safety in a nationwide organization of not-for-profit acute care hospitals and offering financial rewards to participating top performing hospitals. Two PSIs are among the evidence-based indicators being collected (CMS 2005). Reports such as AHRQ's National Healthcare Quality Report and National Healthcare Disparities Report, which incorporate the PSIs, provide the first accurate, albeit limited, assessment of national performance on quality and patient safety in the United States (AHRQ 2004a, b). States and hospital systems are making PSI data available to facilities electronically. One approach currently being tested uses web-based query tools for hospitals to quickly access their own PSI rates quarterly (Pickard et al. 2004). The tool allows hospitals to query their own PSI rates, compare themselves with other hospitals, and drill down to individual discharge data (all data are secure and de-identified in compliance with HIPAA) (Savitz, Sorensen, and Bernard 2004). This is similar to use of process indicators, such as the IHI Trigger Tools, which prompt further search for potential adverse drug events (Rozich, Haraden, and Resar 2003).

DISCUSSION AND CONCLUSIONS

The AHRQ PSIs are easily implemented indicators of possible preventable adverse events. While low-cost, they are at a relatively high level of abstraction in comparison with the granular detail that is potentially available from other sources such as medical records or incident reports. PSIs are generated by software that screens administrative data. They are less subject to bias than other retrospective safety data collection methods such as event reporting: while administrative data are not immune to manipulation, there are detailed coding guidelines and conventions and countervailing incentives in place to keep this in check. For these reasons, the PSIs also have the potential to be both valid and reliable. They are potentially valid not as counts of actual preventable events, but as moderately sensitive indicators of event rates.

Implications for Practice

From an organizational learning perspective, the PSIs have a potential position in an array of patient safety data that also includes process and structure data, more granular data than the PSIs, and medical-record-based data. As a whole-system measure, the PSIs are useful to senior managers and they may assist in the integration of outcome data from specific units or microsystems within an organization. Because PSIs are built on discharge records, they can easily be "drilled down" to support further inquiry into potential patient safety issues. As administrative data, the PSIs also are likely to prove useful to the extent that they can be linked to data on safety-related processes, such as process monitoring, or structures, such as variation in staffing, procedures, or safety culture, to create a more complete picture of the organization's safety performance (Zhan et al. 2005). If hospitals are to adopt the HROs' preoccupation with failure, they must develop the organizational structures and culture to bear the burden of that preoccupation, because it is not sustainable for people to bear that burden individually (Reason 2000). Ongoing reporting and discussion of PSIs can be part of that structure.

Implications for Research

Viewing the PSIs and their use from the perspective of organizational learning and the example of HROs suggests opportunities, both to continue development of the indicators themselves and to learn how they contribute to safety learning. Given the low incremental cost of producing PSI rate reports, and given the HRO's thirst for safety data, patient safety could be served by

developing and validating more indicators. An increase in the number of indicators would reduce the focus on particular diagnoses and diffuse pressure to downcode administrative data for safety reporting purposes. In addition, a sufficient number of indicators would create the opportunity to test correlations and identify factors from among indicators. This in turn would enable the indicators to depict patient safety on a number of distinct dimensions. The HRO emphasis on timely data suggests that the value of PSIs would be enhanced through complementary development and implementation of rapid turnaround on PSI reporting. The web-based query tool described above is one example. The PSIs can also contribute to safety knowledge and learning to the extent that, as outcome measures, they can be linked by researchers to measures of structures and processes of care.

Conclusions

The PSIs' strengths lie in their reliability and their ease of use; they are likely to be strongest not when used alone, but when part of an integrated system of measures. This is not easy in health care, with its predisposition to focus safety concerns on individual performance rather than on system performance (Schyve 2005). Evidence of this challenge can be found in a recent Stanford–VA safety culture survey, where roughly a third of hospital staff reported that they were not rewarded for promptly identifying errors, mistakes were hidden, and people were punished for their errors. Navy aviators, in comparison, had much lower rates of problematic responses to the same survey (Singer 2003). In sum, higher-level safety data such as the PSIs are only a beginning to the work that is needed to build safer health care systems; however, if used in nonpunitive and system-oriented ways, they have the potential to contribute to safety culture change as well as to the technical side of safety learning.

ACKNOWLEDGMENTS

This paper does not represent the policies or the positions of the Department of Veteran Affairs. No official endorsement by this organization is intended or should be inferred.

REFERENCES

AHRQ (Agency for Healthcare Research and Quality). 2004a. *National Healthcare Disparities Report Publication # 05-0014.* Rockville, MD: Department of Health and Human Services and the Agency for Healthcare Research and Quality.

———. 2004b. *National Healthcare Quality Report AHRQ Publication # 05-0013.* Rockville, MD: Department of Health and Human Services and the Agency for Healthcare Research and Quality.

———. 2005. "Guide to Patient Safety Indicators, Version 2.1, Revised January 3, 2005" [accessed on May 10, 2005]. Available at http://www.qualityindicators.ahrq.gov/psi_download.htm

Amalberti, R., Y. Auroy, D. Berwick, and P. Barach. 2005. "Five Barriers to Achieving Ultrasafe Health Care." *Annals of Internal Medicine* 142: 756–64.

Bent, P. D., S. N. Bolosin, B. J. Creati, A. J. Patrick, and M. E. Colson. 2002. "Professional Monitoring and Critical Incident Reporting Using Personal Digital Assistants." *Medical Journal of Australia* 177 (9): 496–99.

Bierly, P. E., and J.-C. Spender. 1995. "Culture and High Reliability Organizations: The Case of the Nuclear Submarine." *Journal of Management* 21 (4): 639–56.

Caplan, R. A., K. L. Posner, and F. W. Cheney. 1991. "Effect of Outcome on Physician Judgments of Appropriateness of Care." *Journal of the American Medical Association* 265 (15): 1957–60.

Carroll, J. S., J. W. Rudolph, and S. Hatakenaka. 2002. "Learning from Experience in High-Hazard Organizations." *Research in Organizational Behavior* 24: 87–137.

Classen, D. C., S. L. Pestotnik, R. S. Evans, and J. P. Burke. 1991. "Computerized Surveillance of Adverse Drug Events in Hospital Patients." *Journal of the American Medical Association* 266: 2847–51.

CMS. 2005. "Rewarding Superior Quality Care: The Premier Hospital Quality Incentive Demonstration Project 'Centers for Medicare and Medicaid Services' [accessed on November 24, 2005]. Available at http://www.cms.hhs.gov/quality/hospital/PremierFactSheet.pdf

Dillon, R. L., and C. H. Tinsley. 2005. "Whew That Was Close: How Near-Miss Events Bias Subsequent Decision-Making under Risk." Paper Presented at Academy of Management Annual Meeting, Honolulu, HI.

Donabedian, A. 1980. *The Definition of Quality and Approaches to Its Assessment.* Ann Arbor, MI: Health Administration Press.

Edmondson, A. E. 1996. "Learning from Mistakes Is Easier Said Than Done: Group and Organizational Influences on the Detection and Correction of Human Error." *Journal of Applied Behavioral Science* 32 (1): 5–28.

Feldman, M. S., and B. T. Pentland. 2003. "Reconceptualizing Organizational Routines as a Source of Flexibility and Change." *Administrative Science Quarterly* 48: 94–118.

Gaba, D. M. 2000a. "Anaesthesiology as a Model for Patient Safety in Health Care." *British Medical Journal* 320: 785–8.

———. 2000b. "Structural and Organizational Issues in Patient Safety: A Comparison of Health Care to Other High-Hazard Industries." *California Management Review* 43 (1): 83–102.

———. 2003. "Safety First: Ensuring Quality Care in the Intensely Productive Environment—The HRO Model." *Anesthesia Patient Safety Foundation Newsletter* 18 (1): 13–4.

Gallagher, B., L. Cen, and E. L. Hannan. 2005. "Readmissions for Selected Infections Due to Medical Care: Expanding the Definition of a Patient Safety Indicator." In *Advances in Patient Safety: From Research to Implementation*, vol. 2, edited by K. Henriksen, J. Battles, D. I. Lewin, and E. Marks, pp. 39–50. Rockville, MD: Agency for Healthcare Research and Quality.

Harris, D., W. D. Pace, and D. H. Fernald. 2005. "Mitigating Factors in Near Miss Events Reported to a Patient Safety Reporting System: A Qualitative Study from the ASIPS Collaborative." Paper Presented at Academy Health Annual Research Meeting, Boston.

HealthGrades. 2004. "Patient Safety in American Hospitals" [accessed on October 15, 2004]. Available at http://www.healthgrades.com/media/english/pdf/HG_Patient_Safety_Study_Final.pdf

Henriksen, K., and H. Kaplan. 2003. "Hindsight Bias, Outcome Knowledge and Adaptive Learning." *Quality and Safety in Health Care* 12 (suppl 2): 46–50.

Huber, G. P. 1991. "Organizational Learning: The Contributing Processes and the Literatures." *Organization Science* 2 (1): 88–115.

Hurtado, M. P., E. K. Swift, and J. M. Corrigan, (eds.) 2001. *Institute of Medicine, Committee on the National Quality Report on Health Care Delivery. Envisioning the National Health Care Quality Report*. Washington, DC: National Academy Press.

Kasandjian, V. A. 2003. *Accountability through Measurement: A Global Healthcare Imperative*. Milwaukee, WI: ASQ Quality Press.

Kohn, L. T., J. M. Corrigan, and M. S. Donaldson, (eds.) 2000. *To Err Is Human: Building a Safer Health System. Report of the Institute of Medicine's Committee on Quality of Health Care in America*. Washington, DC: National Academy Press.

Leape, L. L. 1994. "Error in Medicine." *Journal of the American Medical Association* 272 (23): 1851–7.

Levitt, B., and J. G. March. 1988. "Organizational Learning." *Annual Review of Sociology* 14: 319–40.

Lichtman, J. H., S. A. Roumanis, M. J. Radford, M. S. Riedinger, S. Weingarten, and H. M. Krumholz. 2001. "Can Practice Guidelines Be Transported Effectively to Different Settings? Results from a Multicenter Interventional Study." *Joint Commission Journal of Quality Improvement* 2 (1): 42–53.

Mant, J. 2001. "Process vs. Outcome Indicators in the Assessment of Quality of Health Care." *International Journal for Quality in Health Care* 13 (6): 475–80.

Mant, J., and N. Hicks. 1995. "Detecting Differences in Quality of Care: The Sensitivity of Measures of Process and Outcome in Treating Acute Myocardial Infarction." *British Medical Journal* 311: 793–6.

March, J. G., L. S. Sproull, and M. Tamuz. 1991. "Learning from Samples of One or Fewer." *Organization Science* 2 (1): 1–13.

McDonald, K., P. Romano, J. Geppert, and S. Davies. 2002. "Evidence Report for Measure of Patient Safety Based on Hospital Administrative Data—The Patient Safety Indicators (Draft Report)." UCSF–Stanford Evidence-Based Practice Center, Contract # 290-97-0013 for the Agency for Healthcare Research and Quality, Stanford, CA.

Michel, P., J.-L. Quenon, S. M. D. Sarasqueta, and L. Scemama. 2004. "Comparison of Three Methods for Estimating Rates of Adverse Events and Rates of Preventable Adverse Events in Acute Care Hospitals." *British Medical Journal* 328: 199–203.

Miller, M., A. Elixhauser, C. Zhan, and G. Meyer. 2001. "Patient Safety Indicators: Using Administrative Data to Identify Potential Patient Safety Concerns." *Health Services Research* 36 (6): 110–32.

Mills, P. D., W. B. Weeks, and B. C. Surott-Kimberly. 2003. "A Multihospital Safety Improvement Effort and the Dissemination of New Knowledge." *Joint Commission Journal on Quality and Safety* 29 (3): 124–33.

Mohr, J. J., and P. B. Batalden. 2002. "Improving Safety on the Front Lines: The Role of Clinical Microsystems." *Quality and Safety in Health Care* 11: 45–50.

Pate, B., and R. Stajer. 2001. "The Diagnosis and Treatment of Blame." *Journal for Healthcare Quality* 23 (1): 4–8.

Perrow, C. 1984. *Normal Accidents: Living with High Risk Systems*. New York: Basic Books.

Pickard, S., W. Xu, P. Hougland, Z. Liu, L. Savitz, and A. Sorensen. 2004. "Hospital Patient Safety Indicator Web Querying Tool: Facilitating Knowledge Transfer and Utilization." Utah Department of Health and RTI International. Presented at "Making the Health Care System Safer: Annual Patient Safety Research Conference, September 26–28, 2004.

Pronovost, P. J., T. Nolan, S. Zeger, M. Miller, and H. Rubin. 2004. "How Can Clinicians Measure Safety and Quality in Acute Care?" *Lancet* 363 (9414): 1061–73.

Reason, J. 1997. *Managing the Risks of Organizational Accidents*. Aldershot, U.K.: Ashgate.

———. 2000. "Human Error: Models and Management." *British Medical Journal* 320: 768–70.

Rivard, P. E., A. R. Elwy, S. Loveland, S. Zhao, D. Tsilimingras, A. Elixhauser, P. Romano, and A. Rosen. 2005. "Applying Patient Safety Indicators across Healthcare Systems: Achieving Data Comparability." In *Advances in Patient Safety: From Research to Implementation*, vol. 2, edited by K. Henriksen, J. Battles, D. I. Lewin, and E. Marks, pp. 7–24. Rockville, MD: Agency for Healthcare Research and Quality.

Roberts, K. H., and D. M. Rousseau. 1989. "Research in Nearly Failure-Free, High-Reliability Systems: Having the Bubble." *IEEE Transaction on Engineering Management* 36: 132–9.

Romano, P. S., B. K. Chan, M. E. Schembri, and J. A. Rainwater. 2002. "Can Administrative Data Be Used to Ascertain Clinically Significant Postoperative Complication Rates across Hospitals?" *Medical Care* 40 (10): 856–67.

Romano, P. S., J. J. Geppert, S. Davies, M. R. Miller, A. Elixhauser, and K. M. McDonald. 2003. "A National Profile of Patient Safety in U.S. Hospitals." *Health Affairs* 22 (2): 154–66.

Rosen, A. K., P. E. Rivard, S. Loveland, S. Zhao, C. L. Christiansen, D. Tsimilimingras, A. Elixhauser, P. S. Romano, W. Henderson, and S. Khuri. 2006. "Identification of Patient Safety Events from VA Administrative Data: Is It Valid?" VA Health Services Research and Development National Meeting, Arlington, VA.

Rosen, A. K., P. E. Rivard, S. Zhao, S. Loveland, D. Tsilimingras, C. Christiansen, A. Elixhauser, and P. S. Romano. 2005. "Evaluating the Patient Safety Indicators: How Well Do They Perform on Veterans Health Administration Data?" *Medical Care* 143 (9): 873–84.

Rosen, A. K., S. Zhao, P. E. Rivard, S. Loveland, M. E. Montez, A. Elixhauser, and P. S. Romano. 2006. "Tracking Rates of Patient Safety Indicators over Time: Lessons from the VA." *Medical Care*, in press.

Rozich, J. D., C. R. Haraden, and R. K. Resar. 2003. "Adverse Drug Event Trigger Tool: A Practical Methodology for Measuring Medication Related Harm." *Quality and Safety in Health Care* 12: 194–200.

Rubin, H. R., P. J. Pronovost, and G. B. Diette. 2001. "The Advantages and Disadvantages of Process-Based Measures of Health Care Quality." *International Journal for Quality in Health Care* 13 (6): 469–74.

Runciman, W. B., M. J. Edmonds, and M. Pradhan. 2002. "Setting Priorities for Patient Safety." *Quality and Safety in Health Care* 11: 224–9.

Savitz, L., A. Sorensen, and S. Bernard. 2004. *Using AHRQ Patient Safety Indicators to Drive Quality Improvement.* Report. Submitted to the Agency for Healthcare Research and Quality by RTI International, Research Triangle Park, NC.

Schyve, P. M. 2005. "Systems Thinking and Patient Safety." In *Advances in Patient Safety: From Research to Implementation*, vol. 2, edited by K. Henriksen, J. Battles, D. I. Lewin, and E. Marks, pp. 1–4. Rockville, MD: Agency for Healthcare Research and Quality.

Sharpe, V. A., and A. I. Faden. 1998. *Medical Harm: Historical, Conceptual, and Ethical Dimensions of Iatrogenic Illness.* Cambridge, U.K.: Cambridge University Press.

Singer, S. J. 2003. "Stanford-VA Palo Alto Survey Examines Health Care Workers' Attitudes Toward Safety." *Anesthesia Patient Safety Foundation Newsletter* 18 (1): 7–8.

Sitkin, S. B. 1992. "A Strategy of Learning through Failure: The Strategy of Small Losses." In *Research in Organizational Behavior*, vol. 14, edited by B. M. Staw and L. Cummings, pp. 231–66. Greenwich, CT: JAI Press.

Spear, S. J., and M. Schmidhofer. 2005. "Ambiguity and Workarounds as Contributors to Medical Error." *Annals of Internal Medicine* 142 (2): 627–30.

Thomas, E. J., and R. L. Helmreich. 2002. "Will Airline Safety Models Work in Medicine?" In *Medical Errors: What Do We Know? What Do We Do?* edited by M. M. Rosenthal and K. M. Sutcliffe. San Francisco: Jossey-Bass.

Thomas, E. J., and L. A. Peterson. 2003. "Measuring Errors and Adverse Events in Health Care." *Journal of General Internal Medicine* 18 (1): 61–7.

Tucker, A. L., and A. C. Edmondson. 2003. "Why Hospitals Don't Learn from Failures: Organizational and Psychological Dynamics That Inhibit System Change." *California Management Review* 45 (2): 55–72.

van der Schaaf, T. W. 2002. "Medical Applications of Industrial Safety Science." *Quality and Safety in Health Care* 11: 205–6.

Weick, K. E. 1987. "Organizational Culture as a Source of High Reliability." *California Management Review* 29 (2): 112–27.

———. 1993a. "The Collapse of Sensemaking in Organizations: The Mann Gulch Disaster." *Administrative Science Quarterly* 38: 628–52.

———. 1993. "The Vulnerable System: An Analysis of the Tenerife Air Disaster." In *New Challenges to Understanding Organizations*, edited by K. H. Roberts, pp. 173–98. New York: Macmillan.

Weick, K. E., and K. H. Roberts. 1993. "Collective Mind in Organizations: Heedful Interrelating on Flight Decks." *Administrative Science Quarterly* 38: 357–81.

Weick, K. E., and K. M. Sutcliffe. 2001. *Managing the Unexpected: Assuring High Performance in an Age of Complexity*. San Francisco: Jossey-Bass.

Weick, K. E., K. M. Sutcliffe, and D. Obstfeld. 1999. "Organizing for High Reliability: Processes of Collective Mindfulness." *Research in Organizational Behavior* 21: 81–123.

Woolever, D. R. 2005. "The Impact of a Patient Safety Program on Medical Error Reporting." In *Advances in Patient Safety: From Research to Implementation*, vol 1, edited by K. Henriksen, J. Battles, D. I. Lewin, and E. Marks, pp. 307–316. Rockville, MD: Agency for Healthcare Research and Quality.

Young, G. J., M. P. Charns, K. Desai, S. F. Khuri, M. G. Forbes, W. Henderson, and J. Daley. 1998. "Patterns of Coordination and Clinical Outcomes: A Study of Surgical Services." *Health Services Research* 33 (5, part 1): 1211–36.

Zhan, C., E. Kelley, H. P. Yang, M. Keyes, J. Battles, R. J. Borotkanics, and D. Stryer. 2005. "Assessing Patient Safety in the United States: Challenges and Opportunities." *Medical Care* 43 (3 suppl): I-42–I-47.

Zhan, C., and M. R. Miller. 2003a. "Excess Length of Stay, Charges, and Mortality Attributable to Medical Injuries During Hospitalization." *Journal of the American Medical Association* 290 (14): 1868–74.

———. 2003b. "Administrative Data Based Patient Safety Research: A Critical Review." *Quality and Safety in HealthCare* 12: 58ii–63ii.