

Creating High Reliability in Health Care Organizations

Peter J. Pronovost, Sean M. Berenholtz, Christine A. Goeschel, Dale M. Needham, J. Bryan Sexton, David A. Thompson, Lisa H. Lubomski, Jill A. Marsteller, Martin A. Makary, and Elizabeth Hunt

Objective. The objective of this paper was to present a comprehensive approach to help health care organizations reliably deliver effective interventions.

Context. Reliability in healthcare translates into using valid rate-based measures. Yet high reliability organizations have proven that the context in which care is delivered, called organizational culture, also has important influences on patient safety.

Model for Improvement. Our model to improve reliability, which also includes interventions to improve culture, focuses on valid rate-based measures. This model includes (1) identifying evidence-based interventions that improve the outcome, (2) selecting interventions with the most impact on outcomes and converting to behaviors, (3) developing measures to evaluate reliability, (4) measuring baseline performance, and (5) ensuring patients receive the evidence-based interventions. The comprehensive unit-based safety program (CUSP) is used to improve culture and guide organizations in learning from mistakes that are important, but cannot be measured as rates.

Conclusions. We present how this model was used in over 100 intensive care units in Michigan to improve culture and eliminate catheter-related blood stream infections—both were accomplished. Our model differs from existing models in that it incorporates efforts to improve a vital component for system redesign—culture, it targets 3 important groups—senior leaders, team leaders, and front line staff, and facilitates change management—engage, educate, execute, and evaluate for planned interventions.

Key Words. Patient safety, quality, reliability, culture

In the years 1999 and 2001, landmark reports from the Institute of Medicine (IOM) made deficiencies in quality of care and patient safety inescapably visible to health care professionals and the public (Institute of Medicine 1999, 2001). What have we accomplished since these reports? Are we safer; and if so, how do we know? Many say we lack empiric evidence to demonstrate improved safety (Wachter 2004; Brennan et al. 2005; Leape and Berwick 2005), with few measures to broadly evaluate our progress with improvements.

Current publicly reported performance measures are likely insufficient for providers to evaluate safety. In many hospitals, these performance measures apply to < 10 percent of a hospitals' discharges (Jha et al. 2005). We need scientifically sound and feasible measures of patient safety.

In light of these challenges, health care has turned to "high-reliability organizations" (HRO) (e.g., aviation), who achieved a high degree of safety or reliability despite operating in hazardous conditions (Weick and Sutcliffe 2001). Exactly what does reliability mean in health care and how do we know if we are reliable? These answers remain elusive.

Reliability is often presented as a defect rate in units of 10 and generally represents the number of defects per opportunity for that defect. In health care, an opportunity for a defect usually translates to a population of patients at risk for the medical error or adverse event. For example, within a health care institution, failure to use evidence-based interventions may occur in five of 10 patients, or a catheter-related blood stream infection (CRBSI) in four of 1,000 catheter days (McGlynn et al. 2003; CDC 2004). A fundamental principle in measuring reliability is focusing on defects that can be validly measured as rates, which is not possible for most patient safety defects. Rates need a clearly defined numerator (defect) and denominator (population at risk) and must be devoid of reporting biases (see framework below).

In addition to valid measures, HRO and health care safety experts recognize that the context in which work occurs, called "organizational culture," has important influences on patient safety (Donald and Canter 1994; Hofmann and Stetzer 1996; Zohar 2000; Barling, Loughlin, and Kelloway 2002; Hofmann, Morgeson, and Gerras 2003; Sexton, Thomas, and Pronovost 2005). For example, the ability of staff to raise concerns or senior leaders to

Address correspondence to Peter J. Pronovost, M.D., Ph.D., Departments of Anesthesiology & Critical Care Medicine, Surgery, and Health Policy & Management, The Johns Hopkins University, 1901 Thames Street, 2nd floor, Baltimore, MD 21231. Sean M. Berenholtz, M.D., M.H.S., David A. Thompson, D.N.Sc., M.S., and Lisa H. Lubomski, Ph.D., are with the Department of Anesthesiology & Critical Care Medicine, The Johns Hopkins University, Baltimore, MD. Christine A. Goeschel, R.N., M.P.A., M.P.S., is with the MHA Keystone Center for Patient Safety & Quality, Lansing, MI. Dale M. Needham, M.D., Ph.D., is with The Johns Hopkins University, Baltimore, MD. J. Bryan Sexton, Ph.D., is with the Departments of Anesthesiology & Critical Care Medicine and Health Policy & Management, The Johns Hopkins University, Baltimore, MD. Jill A. Marsteller, Ph.D., M.P.P., is with the Johns Hopkins Bloomberg School of Public Health, Department of Health Policy & Management, Baltimore, MD. Martin A. Makary, M.D., M.P.H., is with The Johns Hopkins Bayview Medical Center; Department of Surgery A5 Center, Baltimore, MD. Elizabeth A. Hunt, M.D., M.P.H., is with the Johns Hopkins University, Department of Anesthesiology & Critical Care Medicine, Baltimore, MD.

listen and act on those concerns can influence safety. In health care, communication failures are a leading contributing factor in all types of sentinel events reported to the Joint Commission on the Accreditation of Health Care Organizations (www.jcaho.org), with poor communication often occurring between the caregivers who interact most often—physicians and nurses (Sexton, Helmreich, and Thomas 2000). Valid measures of safety climate constructs can be made by systematically eliciting frontline caregivers' perceptions of the organizations commitment to safety (i.e., "safety climate") using questionnaires (Sexton et al. 2004). Early evidence demonstrates that safety climate is responsive to interventions (Pronovost 2005). As such, strategies to improve reliability must occur in a culture that is conducive to change.

A clear framework to measure safety within a health care organization is lacking. Federal agencies, organizations, and some institutions have developed "score cards" or performance measurement reports. However, some have scores of measures and most measures either lack validity (e.g., overall hospital mortality) or target-specific patient populations (e.g., congestive heart failure), preventing generalizability of results to the entire organization (Thomas and Hofer 1999; Hayward and Hofer 2004; Lilford et al. 2004). A comprehensive approach to evaluate an organization's progress with patient safety efforts has not been clearly articulated.

In this paper, we describe a comprehensive approach for health care organizations to measure patient safety and then present an example of how this approach was applied to eliminate CRBSIs and improve safety culture in intensive care units (ICUs) in the state of Michigan (Pronovost and Goeschel 2005). Through this improvement example, we hope to highlight the importance of balancing the use of scientifically sound and feasible measures of patient safety with wisdom from front-line staff, noting that both are necessary and equally important.

HOW DO WE KNOW WE ARE SAFER?: A FRAMEWORK

Donabedian's model for measuring quality can also serve as a framework for measuring safety. In this model, *structure* (how care is organized) plus *process* (what we do) influences patient *outcomes* (the results achieved) (Donabedian 1966). We adapted this model to patient safety by adding a fourth element, *culture* (the context in which care is delivered). While most current measures of quality focus either on process or outcome elements, many safety measures involve the structure and culture of patient care delivery.

The recent focus on measuring safety has prompted consideration of new structural measures. Such measures can include *institutional variables*, such as how involved leaders are in patient safety efforts, or credentialing mechanisms to ensure staff competency. Other measures could be *task variables*, such as the presence of protocols, or *team variables* where staff lower on the hierarchy feel comfortable voicing concerns to team members higher up the hierarchical ranks (Pronovost, Angus et al. 2002).

A key challenge in measuring safety is clarifying what can and cannot be measured as a valid rate. To be a valid rate, the numerator (event or harm) and denominator (population at risk) should be clearly defined and measured with minimal bias. A surveillance system must be in place to accurately identify and measure both the numerator and denominator of the rate (Gordis 2004). Most safety parameters, such as information from patient safety reporting systems (PSRS), are difficult, if not impossible, to capture in the form of a valid rate. Such safety parameters are still useful, but not interpretable as rates. For example, surgical complications or medical errors that result in significant patient harm are important as a numerical count (i.e., numerator), but they are not likely valid rates as there is no clear denominator and reporting biases present for the numerator. Establishing safety indicators that can be measured as valid rates is a critical first step in monitoring and improving safety and reliability.

We have developed a framework (Table 1) for measuring patient safety that has been previously published (Pronovost, Holzmueller, Sexton et al. 2006). In our framework, we address the critical issue of appropriate use of rates to measure safety by stratifying measurements into two categories. One category uses valid rate-based measures that are readily available using existing hospital resources. This category addresses outcome and process measures, respectively: (1) how often do we harm patients? and (2) how often do we use evidence-based medicine? The second category captures indicators that are essential to patient safety, but not measurable as valid rates. This category addresses structural and context measures, respectively: (1) how do we know we learned from mistakes? and (2) how well have we created a culture of safety—measured with the Safety Attitudes Questionnaire?

INTERVENTIONS TO IMPROVE CULTURE AND LEARN FROM MISTAKES

To improve the nonrate-based measures that are not related to a specific discipline in health care settings, we implement the comprehensive unit based

Table 1: Framework for a “Score Card” for Patient Safety and Effectiveness

<i>Domain</i>	<i>Definition</i>	<i>Example</i>
How often do we harm patients?	Measures of health care-acquired infections using standardized definitions and measurement techniques	Catheter-associated blood stream infections
How often do we provide the interventions that patients should receive?	Measure the proportion of patients that receive evidence-based interventions using either previously validated process of care measures or a validated methodology to develop new measures	Proportion of mechanically ventilated patients receiving elevation of head-of-bed and prophylaxis for peptic ulcers and deep venous thrombosis Proportion of patients receiving appropriate sepsis and palliative care
How do we know we learned from defects?	What proportion of months does each patient care area/unit learn from mistakes	Proportion of months in which at least one sentinel event was reviewed and a policy was created/revised and/or staff awareness or use of that policy was measured
How well have we created a culture of safety?	Annual assessment of safety culture at the unit level within a health care institution	Percent of patient care areas in which 80% of staff report positive safety and teamwork climate

safety program (CUSP), which has demonstrated improvements in safety culture (Pronovost 2005). CUSP provides enough structure such that a health care organization can develop a broad strategy to improve safety, yet flexible enough to defer to the local concerns and wisdom of staff in individual care areas. As part of CUSP, a senior executive adopts a work area and actively participates in safety efforts with staff. Staff in each work area are asked to learn from one defect per month, and department and hospital leaders learn from one defect per quarter using a structured tool (Pronovost, Holzmueller, Martinez et al. 2006). The goal is to move away from just reporting and superficially reviewing multiple hazards to focusing intently on a few and mitigating the hazards (i.e., redesign the system in which work is performed). In addition, CUSP asks safety teams to implement tools, such as daily goals and morning briefings (Pronovost et al. 2003; Thompson et al. 2005) to help improve safety culture.

MODEL TO IMPROVE RELIABILITY

Our model to improve reliability focuses on the rate-based measures of safety; how often do we harm patients and how often do we use evidence-based

medicine. Rate-based measures are specific to a clinical area or discipline. Using the objective of eliminating CRBSI as an example, we describe the model below

1. *Identify interventions associated with an improved outcome in a specific patient population.* To a large extent, this has been accomplished with practice guidelines or summaries of clinical research evidence. For example, the Centers for Disease Control (CDC) and others have published guidelines for preventing CRBSIs (CDC 2004).

2. *Select interventions that have the biggest impact on outcomes and convert these into behaviors* (Grimshaw et al. 2001b; Michie and Johnston 2004). The team should focus on approximately five interventions that are supported by strong evidence, have the greatest potential benefit, and reflect patients' values and preferences. Recent recommendations to grade evidence into "do or do not do" will greatly facilitate this step (Atkins et al. 2004). In selecting interventions, it may be helpful, if not done in the evidence review, to make a table of each potential intervention with the strength of the evidence supporting its use, the strength of the relationship (e.g., a risk ratio) between the intervention and the outcome, and the barriers in implementing the intervention (Gordis 2004).

3. *Develop measures to evaluate reliability.* Here, we seek a scientifically sound and feasible rate-based measure that can either be an outcome or process element of safety. The measure(s) selected should be carefully considered. Both types of measures have strengths and weaknesses that have been published (Rubin, Pronovost, and Diette 2001; Lilford et al. 2004; Pronovost, Nolan et al. 2004). For example, if the intervention is a medication, we could measure if it was given, or what medication, dose, and/or when it was given. Several principles guide us in deciding which of these to measure. First, choose measures that are scientifically sound or supported by the evidence. If timing or dose of antibiotic administration is important, measure when the medication was given and dose given as two separate variables. Second, measure what is feasible, or easily collected with available resources. Third, if possible, measure where defects most commonly occurred. To do this, review each step in the process for a sample of patients and identify where defects most commonly occurred. For example, evidence suggests that steroids reduce mortality in septic shock patients (Annane et al. 2002). When we monitored use of steroids for this patient population, we found that failure to prescribe the medication was the most common defect. As a result, we developed a measure to evaluate whether patients with septic shock received steroids.

Development of measures typically requires significant resources and expertise in developing measures and specific clinical content (Garber 2005),

which few health care organizations will likely have available. As such, national measures should be developed and broadly shared among health care organizations.

In this case, the National Nosocomial Infection Surveillance System (NNIS) has standardized measures for CRBSI that are valid, reliable, and widely used, which prompted us to measure the outcome rather than the process. Attempts to measure the process proved neither valid nor feasible. Such a measure would require additional ICU staff—not available to us—to independently monitor the placement of all central venous catheters.

4. *Measure baseline performance.* This is the best test of whether the proposed measures can be feasibly collected. If baseline data cannot be collected with minimal bias, it is unlikely that these data can be collected after the intervention has been implemented. Moreover, without baseline data, an organization cannot assess if safety has improved. In addition to collecting data, a health care organization should create a database to evaluate data quality and missing data, store and analyze data, and produce reports. In our experience, few quality improvement projects create such a database.

5. *Ensure patients receive evidence-based interventions.* This effort is the biggest challenge. While steps 1–4 are generally performed by a team of researchers and clinicians with sufficient resources who may or may not personally implement the interventions, step 5 involves teams from the participating health care organization who will actually implement the interventions. These interventions must be tailored to address each participant's current system, culture, resources, and commitment. While there is no formula for system redesign, there are many tactics that appear effective for improving care (Grol et al. 1998; Cabana et al. 1999; Grol 2001; Pronovost, Wu et al. 2002; Pronovost, Weast et al. 2004; Pronovost and Berenholtz 2002; Bradley et al. 2005).

The change model we used to improve reliability (outlined in Table 2) was designed as a practical application of theories related to diffusion of innovation and behavior change (Grimshaw et al. 2001a; Greenhalgh et al. 2004; Michie et al. 2005). The change model includes four components: engage, educate, execute, and evaluate. Each component targets senior leaders, team leaders, and front-line staff.

Engaging and educating front-line staff is challenging and resource intensive. The execute component encourages staff to use HRO theory (i.e., standardize, create independent checks, and learn from mistakes) to ensure patients receive evidence-based interventions. Here, we encourage teams to first consider how they can standardize (including reducing complexity) what they do to reduce the risk of failure. Often this step includes creating a standard

Table 2: Strategy for Leading Change

	<i>Executive Leaders</i>	<i>Team Leaders</i>	<i>Staff</i>
Engage	<p>How do I make the world a better place? How do I create an organization that is safe for patients and rewarding for staff? How does this strategy fit our mission?</p>	<p>How do I make the world a better place? How do I create a unit that is safe for patients and rewarding for staff? How do I touch their hearts?</p>	<p>How do I make the world a better place? Do I believe I can change the world, starting with my unit? Can I help make my unit safer for patients and a better place to work? What do I need to do?</p>
Educate	<p>What do I need to do? What is the business case? How do I engage the board and medical staff? How can I monitor progress?</p>	<p>What do I need to do? What is the evidence? Are executive and medical staff aware of evidence, agree with it, able to implement it? Are there tools to help me develop a plan?</p>	<p>Why is this change important? How are patient outcomes likely to improve? How does my daily work need to change? Where do I go for support? How do we ensure we do it?</p>
Execute	<p>How do we ensure we do it? Do the Board and Medical Staff support the plan and have the skills and vision to implement? How do I know the team has sufficient resources, incentives and organizational support?</p>	<p>How do we ensure we do it? Do staff know the plan and do they have the skills and commitment to implement? How do we ensure we do it? Have we tailored this to our environment?</p>	<p>Can I be a better team member and team leader? How can I share what I know to make care better? Am I learning from defects?</p>
Evaluate	<p>How will I know I made a difference? Have resources been allocated to collect and use safety data? Is the work climate better? Are patients safer? How do I know?</p>	<p>How will I know I made a difference? Have I created a system for data collection, unit level reporting, and using data to improve it? Is the work climate better? Are patients safer? How do I know?</p>	<p>How will I know I made a difference? What is our unit level report card? Is the unit a better place to work? Is teamwork better? Are patients safer? How do I know?</p>

order set or protocol. Next, teams create independent checks (i.e., two or more persons recheck independent of the other[s]) for key processes. Finally, when defects occur, teams are encouraged to evaluate or learn the causes.

IMPROVING PATIENT SAFETY THROUGHOUT MICHIGAN ICUS

Project Overview

We applied this safety framework to improve safety in over 100 Michigan ICUs. This research study, called the Keystone ICU project, was based on a collaborative model (Ovretveit et al. 2002; Mills and Weeks 2004) between the Johns Hopkins University, Quality and Safety Research Group (QSRG), and the Michigan Health & Hospital Association (MHA), Keystone Center for Patient Safety & Quality. The project was designed as a prospective cohort study to evaluate the effects of implementing patient-safety interventions. The research was conducted from September 30, 2003 to September 30, 2005. It was funded by the U.S. Agency for Healthcare Research and Quality (AHRQ) and received Institutional Review Board approval from the Johns Hopkins University School of Medicine.

In June 2003, all Michigan hospitals with ICUs were invited to participate in the Keystone ICU project. To participate, hospitals had to assemble an ICU improvement team and send a written commitment to the project, signed by a hospital senior executive. At a minimum, the ICU improvement team included a senior executive, the ICU director and nurse manager, an ICU physician and nurse, and often a department administrator. Hospital senior executives were asked to ensure that the ICU physician and nurse could commit 20 percent of their time to the project. In addition, each team agreed to implement the patient-safety interventions, collect and submit the required data in a timely manner, attend the biannual 1.5-day conferences, and participate in monthly conference calls.

Study Goals and Methods

The overall objective of the study was to improve patient safety using the safety score card (Table 1), in participating ICUs. We will discuss two study objectives in this paper: improving culture and eliminating CRBSI.

Before implementing the intervention (January–March 2004), and 1 year after exposure (March–May 2005), participating ICUs assessed their safety culture in a pre–post design. The Safety Attitudes Questionnaire (SAQ)

(ICU version) (Pronovost and Sexton 2005) was administered to all caregivers who routinely had contact with ICU patients. This survey is reliable, sensitive to change (Gregorich, Helmreich, and Wilhelm 1990; Thomas et al. 2005), and elicits attitudes shown to predict important performance outcomes (Fou-shee 1984; Helmreich et al. 1986; Pronovost et al. 2005). The six domains of the SAQ are perceptions of management, job satisfaction, stress recognition, working conditions, teamwork climate, and safety climate. A more detailed report of assessing and improving safety climate will be presented elsewhere (unpublished data, assessing and improving safety climate in a statewide sample of ICUs). In this study, we report context of care issues related to patient safety and perceptions of leadership to demonstrate the impact of providers surfacing and addressing safety issues with hospital leaders. Response options for each item range from 1 (disagree strongly) to 5 (agree strongly).

Throughout the study, data on the number of CRBSIs and central line days were collected monthly from the hospital Infection Control Practitioner (ICP) using CDC NNIS system definitions and standards. A quarterly CRBSI rate was calculated as the number of infections per 1,000 central line days for each 3-month period. Each quarterly CRBSI rate was assigned to one of five time periods: preimplementation baseline, peri-implementation, and 0–3, 4–6, or 7–9 months postimplementation.

To reduce bias in data collection, we developed a manual of operations, which included explicit definitions for each process and outcome measure. Standardized data collection forms were developed, pilot tested, revised, and distributed to ICU teams and then converted into electronic format. Teams were trained to collect data via conference calls. ICUs received monthly and quarterly ICU performance reports and compared their performance with aggregate results from the other participating ICUs.

Teams focused first on improving ICU culture using CUSP because we believed that this change was necessary before teams could redesign care and improve reliability (Sexton, Helmreich, and Thomas 2000; Shortell et al. 2004b). For rate-based measures, we will discuss CRBSI.

Model to Improve Reliability: Toward Eliminating CRBSI

Identify and select interventions (steps 1 and 2). To reduce CRBSI, we summarized the nearly 100 page evidence summary into five behavior-specific interventions related to central line placement: (1) wash your hands, (2) use full-barrier precautions, (3) prepare the insertion site with chlorhexadine antiseptic, (4) avoid the femoral site for insertion, and (5) remove unnecessary lines.

Develop measures and collect baseline data (steps 3 and 4). Because NNIS has standardized measures for CRBSI, we opted to measure the outcome not the process. The median baseline rate of CRBSI was 4.2 per 1,000 catheter days.

Ensure patients receive interventions (step 5). Using our change model, we engaged ICU staff by providing an estimate of the number of deaths attributable to CRBSIs in their ICU. Indeed, harm was now visible. We educated staff by making the research evidence supporting the CRBSI intervention easily accessible in the form of original literature, concise evidence summaries, and slide presentations of the relevant literature. We accomplished engagement and education of front line ICU staff through conference calls, newsletters, and printed educational materials. Available resources limited our ability to create electronic learning tools, but this represents another potential aid for engaging and educating a large number of ICU staff. To implement the Keystone ICU project interventions, team leaders were encouraged to make a task list and associated time line for the interventions and then pilot test the interventions on a small sample of patients or caregivers before wide-scale implementation.

To execute the interventions and ensure patients reliably received these evidence-based interventions, we asked teams to standardize, create independent checks, and learn from mistakes. Complexity was reduced and standardization accomplished by creating a central line cart to store all necessary equipment and supplies for line insertion. Previously, caregivers went to eight different locations in the ICU to collect all necessary equipment. In addition, we created a checklist of the five interventions to reduce CRBSI and empowered nurses assisting with central line placement to ensure physician compliance with all five interventions under nonemergency conditions (Berenholtz et al. 2004). Finally, when a CRBSI occurred, the care team evaluated the case to identify whether it could have been prevented.

ICU teams partnered with their hospital infection control staff to implement the CRBSI intervention and monitor its impact. This approach centralized and standardized data collection and fostered local ownership and accountability for improving CRBSI.

In addition, we directly involved senior leadership from each participating hospital in specific tasks to help the project succeed. For example, strong evidence has shown that skin sterilization, specifically with chlorhexidine, before central venous catheter placement will reduce CRBSI by 50 percent (Mermel 2000). At the start of this project, 20 percent of Michigan hospitals had chlorhexidine routinely available in their ICU central line kits. Chief executive officers (CEOs) were sent a letter from the principal investigator (PJP) and project director (CG) outlining this evidence and asking them

to facilitate the availability and use of chlorhexadine in their hospitals. Within 6 weeks, 76 percent of participating hospitals had chlorhexadine in-house and by project end, all teams were using chlorhexadine (Pronovost and Goeschel 2005).

The senior leader's role is to provide teams with sufficient resources and incentives, and remove barriers (e.g., political) to the team's success. Unfortunately, we lack a formal mechanism to evaluate the extent to which teams perceive senior leaders are performing this role. To surface barriers for successful intervention implementation and provide feedback to senior leaders regarding these barriers, we surveyed ICU teams monthly using a "team checkup" survey. Specifically, we asked their perceptions about the adequacy of physician and senior leader support, time to implement the interventions, and support for data collection. Over half of the teams reported that senior and physician leaders, and insufficient time significantly deterred their progress. Senior leaders were given this survey data in a process to evaluate their leadership role in the project.

Impact

We obtained data from 99 of 107 ICUs in 2004 and 98 of 127 in 2005. Between the 2004 and 2005 administrations, ICU mergers, closings, splits, or failure to collect data in both years left 72 ICUs intact with 2004 and 2005 data. In 72 ICUs, we received 4,474 of 5,975 surveys (75 percent response rate) in 2004 and 3,876 of 5,965 (65 percent response rate) in 2005. Two-tailed paired sample *t*-tests showed that context of care items improved, and we report agreement with the following items at the respondent and ICU levels: "Patient safety is constantly reinforced as the priority in this ICU" $t(71) = 5.091$, $p < .001$, respondents pre 74 percent and post 80 percent, ICUs pre 58 percent and post 78 percent; "There is widespread adherence to clinical guidelines and evidence based criteria in this ICU" $t(71) = 7.041$, $p < .001$, respondents pre 59 percent and post 66 percent, ICUs pre 10 percent and post 25 percent; "The administration of this hospital is doing a good job" $t(71) = 3.449$, $p < .001$, respondents pre 36 percent and post 42 percent, ICUs pre 1 percent and post 3 percent; "Hospital administration supports my daily efforts" $t(71) = 3.417$, $p < .001$, respondents pre 33 percent and post 38 percent, ICUs pre 1 percent and post 3 percent. Figure 1 shows the distribution of agreement for each context of care item for 2004 and 2005.

At the start of the study, 107 ICUs agreed to participate and 98 ICUs collected CRBSI data using NNIS definitions. As described in Table 3, the

Figure 1: Distribution of context of care items in 2004 and 2005. The y-axis is the percent of respondents that agree (agree slightly or agree strongly), and the x-axis represents Michigan intensive care units (ICUs)

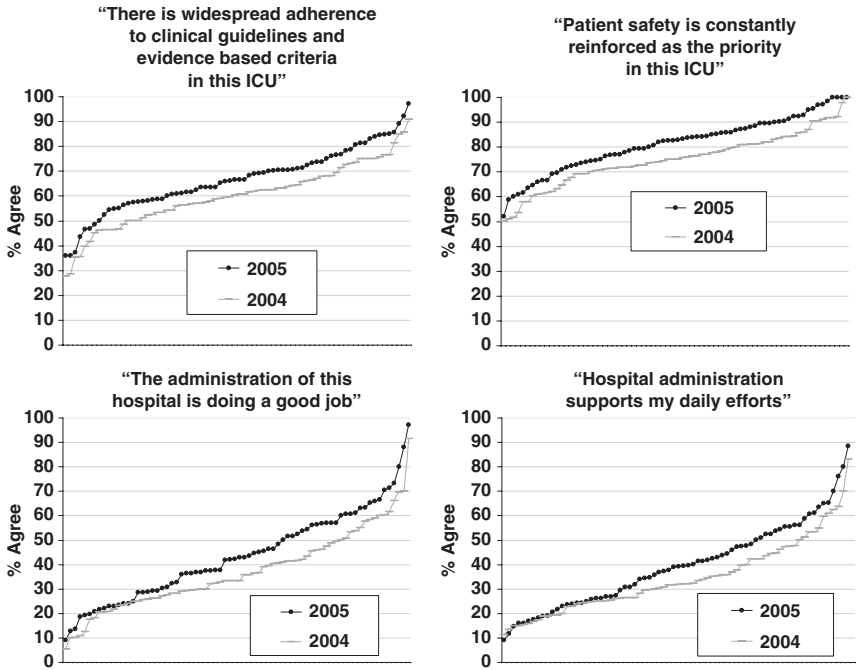


Table 3: Catheter-Related Blood Stream Infection Rates per ICU Month of Observation by Time Period*

Time Period*	ICU Months of Observation (%)	Proportion with Zero CRBSI (%)	p-value**
Preintervention baseline	203 (23)	59	Reference
Periintervention	218 (25)	66	.17
0-3 months postintervention	193 (22)	74	.002
4-6 months postintervention	145 (16)	74	.003
7-9 months postintervention	54 (6)	80	.005
Unknown	69 (8)	75	.016

* Time period is measured in relation to implementation of the CRBSI intervention.

**p-value for comparison with the proportion of ICU-months with zero CRBSI at preintervention baseline using two-sample test of proportion.

CRBSI, catheter-related blood stream infection; ICU, intensive care unit.

proportion of all ICU months of observation with zero CRBSI increased from 59 percent at baseline to 80 percent by 7–9 months postimplementation ($p = .005$, relative risk 0.50, 95 percent CI 0.29–0.87). The prevalence of small ICUs with a relatively low number of catheter line days made the goal of eliminating CRBSI easier. Consequently, we performed a sensitivity analysis of the results. In this analysis, we focused only on observations with ≥ 150 catheter line days per month (52 percent of the entire sample). As expected, the proportion of ICU months with zero CRBSI decreased (44 percent at baseline). However, the magnitude and significance of the improvement achieved by the intervention was unchanged (relative risk 0.53, 95 percent CI 0.30–0.91). For observations with < 150 catheter line days per month, the proportion of ICU months with zero CRBSIs at baseline was greater than 74 percent and the intervention also demonstrated evidence of benefit, but did not reach statistical significance because of the small number of observations in this strata in the 7–9 month postintervention time period (relative risk 0.19, 95 percent CI 0.03–1.32). Thus, the benefit of this intervention is beneficial at eliminating CRBSI across a range of ICU sizes.

COMMENT

We present a framework for improving reliability in health care that was associated with a significant reduction in CRBSIs across nearly 100 ICUs in Michigan and with significant improvements in safety culture (Grol 2001). Our framework differs in several important ways from existing models. First, it incorporates efforts to improve culture. Organizational culture is the lubrication that allows for system redesign and helps ensure the sustainability of changes (Sexton, Helmreich, and Thomas 2000; Shortell et al. 2004a). Second, we targeted three distinct groups to improve safety: senior leaders, team leaders, and front-line staff. Third, teams were given a manual of operations to facilitate change management—engage, educate, execute, and evaluate—for the planned interventions.

Nevertheless, there are limitations to the proposed framework. First, the resources required for this model likely exceed those available at any single hospital. Consequently, these programs are best implemented through a large consortium of hospitals (e.g., state-wide via a state hospital association). Second, measuring improvements in safety takes resources to both develop measures and collect data. Developing measures requires expertise not commonly present in most health systems, and the collection of data for many

measures is not yet a routine part of hospital operations. Third, we calculated CRBSI rates using the 1,000 catheter days standardized through NNIS, which does account for the majority of risks of exposure to a catheter, but does not account for an individual patient's risk of infection from the device. Fourth, like all models, this proposed framework requires empiric validation; efforts to improve effectiveness and efficiency should be a research priority.

CONCLUSION

Nearly all of health care lacks the ability to evaluate whether they are providing safer patient care. HRO provide insight into the context of care, often called culture, that influences reliability. In this paper, we outline a framework for health care organizations to improve reliability and describe application of this model to a large cohort of ICUs in Michigan. Use of this model was associated with a significant reduction in CRBSIs and improvement in culture. We look forward to empiric validation of this model.

ACKNOWLEDGMENT

Funding was provided in part from the Agency for Healthcare Research and Quality (AHRQ), grant #1UC1HS14246.

REFERENCES

- Annane, D., V. Sebillé, C. Charpentier, P. Bollaert, B. Francois, J. Korach, G. Capellier, Y. Cohen, E. Azoulay, G. Troche, P. Chaumet-Riffaut, and E. Bellissant. 2002. "Effect of Treatment with Low Doses of Hydrocortisone and Fludrocortisone on Mortality in Patients with Septic Shock." *Journal of the American Medical Association* 288: 862–71.
- Atkins, D., D. Best, P. A. Briss, M. Eccles, Y. Falck-Ytter, S. Flottorp, G. H. Guyatt, R. T. Harbour, M. C. Haugh, D. Henry, S. Hill, R. Jaeschke, G. Leng, A. Liberati, N. Magrini, J. Mason, P. Middleton, J. Mrukowicz, D. O'Connell, A. D. Oxman, B. Phillips, H. J. Schunemann, T. T. Edejer, H. Varonen, G. E. Vist, J. W. Williams, Jr., and S. Zaza. 2004. "Grading Quality of Evidence and Strength of Recommendations." *British Medical Journal* 328: 1490.
- Barling, J., C. Loughlin, and E. Kelloway. 2002. "Development and Test of a Model Liking Transformational Leadership and Occupational Safety." *Journal of Applied Psychology* 87: 488–96.

- Berenholtz, S. M., P. J. Pronovost, P. A. Lipsett, D. Hobson, K. Earsing, J. E. Farley, S. Milanovich, E. Garrett-Mayer, B. D. Winters, H. R. Rubin, T. Dorman, and T. M. Perl. 2004. "Eliminating Catheter-Related Bloodstream Infections in the Intensive Care Unit." *Critical Care Medicine* 32: 2014–20.
- Bradley, E. H., J. Herrin, J. A. Mattera, E. S. Holmboe, Y. Wang, P. Frederick, S. A. Roumanis, M. J. Radford, and H. M. Krumholz. 2005. "Quality Improvement Efforts and Hospital Performance: Rates of Beta-Blocker Prescription after Acute Myocardial Infarction." *Medical Care* 43: 282–9.
- Brennan, T., A. Gawande, E. Thomas, and D. Studdert. 2005. "Accidental Deaths, Saved Lives, and Improved Quality." *New England Journal of Medicine* 353: 1405–9.
- Cabana, M., C. Rand, N. Powe, A. Wu, M. Wilson, P. Abboud, and H. R. Rubin. 1999. "Why don't Physicians Follow Clinical Practice Guidelines? A Framework for Improvement." *Journal of the American Medical Association* 282: 1458–65.
- CDC. 2004. "National Nosocomial Infections Surveillance (Nnis) System Report, Data Summary from January 1992 through June 2004." *American Journal of Infection Control* 32: 470–85.
- Donabedian, A. 1966. "Evaluating the Quality of Medical Care." *Milbank Memorial Fund Quarterly* 3(Suppl): 166–206.
- Donald, I., and D. Canter. 1994. "Employee Attitudes and Safety in the Chemical Industry." *Journal of Loss Prevention in the Process Industries* 7: 203–8.
- Foushee, H. 1984. "Dyads and Triads at 25,000 Feet: Factors Affecting Group Process and Aircrew Performance." *American Psychologist* 39: 885–993.
- Garber, A. M. 2005. "Evidence-Based Guidelines as a Foundation for Performance Incentives." *Health Affairs (Millwood)* 24: 174–9.
- Gordis, L. 2004. *Epidemiology*, 3d edition. Philadelphia: Saunders.
- Greenhalgh, T., G. Robert, F. Macfarlane, P. Bate, and O. Kyriakidou. 2004. "Diffusion of Innovations in Service Organizations: Systematic Review and Recommendations." *Milbank Quarterly* 82: 581–629.
- Gregorich, S., R. Helmreich, and J. Wilhelm. 1990. "The Structure of Cockpit Management Attitudes." *Journal of Applied Psychology* 75: 682–90.
- Grimshaw, J. M., L. Shirran, R. Thomas, G. Mowatt, C. Fraser, L. Bero, R. Grilli, E. Harvey, A. Oxman, and M. A. O'Brien. 2001a. "Changing Provider Behavior: An Overview of Systematic Reviews of Interventions." *Medical Care* 39: II2–45.
- . 2001b. "Changing Provider Behavior: An Overview of Systematic Reviews of Interventions." *Medical Care* 39: II2–45.
- Grol, R. 2001. "Improving the Quality of Medical Care: Building Bridges among Professional Pride, Payer Profit, and Patient Satisfaction." *Journal of the American Medical Association* 286: 2578–85.
- Grol, R., J. Dalhuijsen, S. Thomas, C. Veld, G. Rutten, and H. Mokkink. 1998. "Attributes of Clinical Guidelines That Influence Use of Guidelines in General Practice: Observational Study." *British Medical Journal* 317: 858–61.
- Hayward, R., and T. Hofer. 2004. "Estimating Hospital Deaths Due to Medical Errors; Preventability Is in the Eye of the Reviewer." *Journal of the American Medical Association* 286: 415–20.

- Helmreich, R., H. Foushee, R. Benson, and W. Russini. 1986. "Cockpit Resource Management: Exploring the Attitude-Performance Linkage." *Aviation Space and Environmental Medicine* 57: 1198-200.
- Hofmann, D., F. Morgeson, and S. Gerras. 2003. "Climate as a Moderator of the Relationship between Leader-Member Exchange and Content Specific Citizenship: Safety Climate as an Exemplar." *Journal of Applied Psychology* 88: 170-8.
- Hofmann, D., and A. Stetzer. 1996. "A Cross-Level Investigation of Factors Influencing Unsafe Behaviors and Accidents." *Personnel Psychology* 49: 307-39.
- Institute of Medicine. 1999. *To Err Is Human: Building a Safer Health System*, edited by L. Kohn, J. Corrigan, and M. Donaldson. Washington, DC: National Academy Press.
- . 2001. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington, DC: National Academy Press.
- Jha, A., Z. Li, E. Orav, and A. Epstein. 2005. "Care in U.S. Hospitals—The Hospital Quality Alliance Program." *New England Journal of Medicine* 353: 265-74.
- Kohn, L., J. Corrigan, and M. Donaldson. 1999. *To Err Is Human: Building a Safer Health System. Institute of Medicine Report*. Washington, DC: National Academy Press.
- Leape, L., and D. Berwick. 2005. "Five Years after 'To Err Is Human': What Have We Learned?" *Journal of the American Medical Association* 293: 2384-90.
- Lilford, R., M. A. Mohammed, D. Braunholtz, and T. P. Hofer. 2004. "The Measurement of Active Errors: Methodological Issues." *Quality and Safety Health Care* 12 (suppl II): ii8-12.
- Lilford, R., M. Mohammed, D. Spiegelhalter, and R. Thomson. 2004. "Use and Misuse of Process and Outcome Data in Managing Performance of Acute Medical Care: Avoiding Institutional Stigma." *Lancet* 363: 1147-54.
- McGlynn, E., S. Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, and E. A. Kerr. 2003. "The Quality of Health Care Delivered to Adults in the United States." *New England Journal of Medicine* 348: 2635-4.
- Mermel, L. 2000. "Prevention of Intravascular Catheter-Related Infections." *Annals of Internal Medicine* 132: 391-402.
- Michie, S., and M. Johnston. 2004. "Changing Clinical Behaviour by Making Guidelines Specific." *British Medical Journal* 328: 343-5.
- Michie, S., M. Johnston, C. Abraham, R. Lawton, D. Parker, and A. Walker. 2005. "Making Psychological Theory Useful for Implementing Evidence Based Practice: A Consensus Approach." *Quality and Safety in Health Care* 14: 26-33.
- Mills, P. D., and W. B. Weeks. 2004. "Characteristics of Successful Quality Improvement Teams: Lessons from Five Collaborative Projects in the VHA." *Joint Commission Journal of Quality and Safety* 30: 152-6.
- Ovretveit, J., P. Bate, P. Cleary, D. Cretin, D. Gustafson, K. McInnes, H. McLeod, T. Molfenter, P. Plsek, G. Robert, S. Shortell, and T. Wilson. 2002. "Quality Collaboratives: Lessons from Research." *Quality and Safety in Health Care* 11: 345-51.
- Pronovost, P., D. Angus, T. Dorman, K. Robinson, T. Dremsizov, and T. Young. 2002. "Physician Staffing Patterns and Clinical Outcomes in Critically Ill

- Patients: A Systematic Review." *Journal of the American Medical Association* 288: 2151-62.
- Pronovost, P., and S. Berenholtz. 2002. A Practical Guide to Measuring Performance in the Intensive Care Unit. VHA Research Series, vol. 2. Irving, TX: VHA Inc.
- Pronovost, P., S. Berenholtz, T. Dorman, P. Lipsett, T. Simmonds, and C. Haraden. 2003. "Improving Communication in the ICU Using Daily Goals." *Journal of Critical Care* 18: 71-5.
- Pronovost, P., and C. Goeschel. 2005. "Improving ICU Care: It Takes a Team." *Healthcare Executive* (March/April): 14-22.
- Pronovost, P., C. Holzmueller, E. Martinez, C. Cafeo, D. Hunt, C. Dickson, M. Awad, and M. Makary. 2006. "A Practical Tool to Learn from Defects in Patient Care." *Joint Commission on Journal of Quality and Patient Safety* 32: 102-8.
- Pronovost, P., C. Holzmueller, J. Sexton, M. Miller, S. Berenholtz, A. Wu, T. Perl, R. Davis, D. Baker, L. Winner, and L. Morlock. 2006b. "How Will We Know Patients Are Safer? An Organization-Wide Approach to Measuring and Improving Patient Safety." *Critical Care Medicine* (in press).
- Pronovost, P., T. Nolan, S. Zeger, M. Miller, and H. Rubin. 2004. "How Can Clinicians Measure Safety and Quality in Acute Care?" *Lancet* 363: 1061-7.
- Pronovost, P. J., and B. J. Sexton. 2005. "Assessing Safety Culture: Guidelines and Recommendations." *Quality and Safety Health Care* 14: 231-3.
- Pronovost, P., B. Weast, K. Bishop, L. Paine, R. Griffith, B. Rosenstein, R. P. Kidwell, K. B. Haller, and R. Davis. 2004. "Senior Executive Adopt-a-Work Unit: A Model for Safety Improvement." *Joint Commission on Journal of Quality and Safety* 30: 59-68.
- Pronovost, P., B. Weast, B. Rosenstein, J. B. Sexton, C. Holzmueller, L. Paine, R. Davis, and H. R. Rubin. 2005. "Implementing and Validating a Comprehensive Unit-Based Safety Program." *Journal of Patient Safety* 1: 33-40.
- Pronovost, P., A. Wu, T. Dorman, and L. Morlock. 2002. "Building Safety into ICU Care." *Journal of Critical Care* 17: 78-85.
- Rubin, H., P. Pronovost, and G. Diette. 2001. "The Advantages and Disadvantages of Process-Based Measures of Health Care Quality." *International Journal of Quality Health Care* 13: 469-74.
- Sexton, J., R. Helmreich, and E. Thomas. 2000. "Error, Stress and Teamwork in Medicine and Aviation: Cross Sectional Surveys." *British Medical Journal* 320: 745-9.
- Sexton, J. B., E. Thomas, R. Helmeich, T. Neilands, K. Rowan, and K. Vella. 2004. "Frontline Assessments of Healthcare Culture: Safety Attitudes Questionnaire Norms and Psychometric Properties." Report number, Technical Report 04-01.
- Sexton, J., E. Thomas, and P. Pronovost. 2005. "The context of Care and the Patient Care Team: The Safety Attitudes Questionnaire." In *Building a Better Delivery System. A New Engineering Health Care Partnership*, edited by P. P. Reid, W. D. Compton, J. H. Grossman, and G. Fanjiang, pp. 119-23. Washington, DC: National Academies Press.
- Shortell, S., J. Marsteller, M. Lin, M. Pearson, S. Wu, P. Mendel, S. Cretin, and M. Rosen. 2004a. "The Role of Perceived Team Effectiveness in Improving Chronic Illness Care." *Medical Care* 42: 1040-8.

- . 2004b. “The Role of Perceived Team Effectiveness in Improving Chronic Illness Care.” *Medical Care* 42: 1040–8.
- Thomas, J., and T. Hofer. 1999. “Accuracy of Risk-Adjusted Mortality Rate as a Measure of Hospital Quality of Care.” *Medical Care* 37: 83–92.
- Thomas, E., J. Sexton, T. Neilands, A. Frankel, and R. Helmreich. 2005. “The Effect of Executive Walk Rounds on Nurse Safety Climate Attitudes. A Randomized Trial of Clinical Units.” *BMC Health Service Research* 5: 28.
- Thompson, D., C. Holzmueller, D. Hunt, C. Cafeo, B. Sexton, and P. Pronovost. 2005. “A Morning Briefing: Setting the Stage for a Clinically and Operationally Good Day.” *Journal of Quality and Patient Safety* 31: 476–9.
- Wachter, R. 2004. “The End of the Beginning: Patient Safety Five Years After ‘To Err Is Human’.” *Health Affairs Web Exclusive* (W4): 534–45.
- Weick, K., and K. Sutcliffe. 2001. *Managing the Unexpected: Assuring High Performance in an Age of Complexity*, 1st edition. San Francisco: Jossey-Bass.
- Zohar, D. 2000. “A Group-Level Model of Safety Climate: Testing the Effect of Group Climate on Microaccidents in Manufacturing Jobs.” *Journal of Applied Psychology* 85: 587–96.