Attitude is everything?: The impact of workload, safety climate, and safety tools on medical errors: A study of intensive care units

Johannes Steyrer
Michael Schiffinger
Clemens Huber
Andreas Valentin
Guido Strunk

Background: Hospitals face an increasing pressure toward efficiency and cost reduction while ensuring patient safety. This warrants a closer examination of the trade-off between production and protection posited in the literature for a high-risk hospital setting (intensive care).
Purposes: On the basis of extant literature and concepts on both safety management and organizational/safety culture, this study investigates to which extent production pressure (i.e., increased staff workload and capacity utilization) and safety culture (consisting of safety climate among staff and safety tools implemented by management) influence the occurrence of medical errors and if/how safety climate and safety tools interact.
Methodology/Approach: A prospective, observational, 48-hour cross-sectional study was conducted in 57 intensive care units. The dependent variable is the incidence of errors affecting those 378 patients treated throughout the entire observation period. Capacity utilization and workload were measured by indicators such as unit occupancy, nurse-to-patient/physician-to-patient ratios, levels of care, or NEMS scores. The safety tools considered include Critical Incidence Reporting Systems, audits, training, mission statements, SOPs/checklists, and the use of barcodes. Safety climate was assessed using a psychometrically validated four-dimensional questionnaire.

Key words: intensive care units, medical error, patient safety, safety climate, safety tools, workload

Johannes Steyrer, PhD, *is Director, Research Institute for Health Care Management and Health Care Economics, Vienna University of Economics and Business, Austria. E-mail: johannes.steyrer@wu.ac.at.*
Michael Schiffinger, PhD, *is Senior Research Fellow, Research Institute for Health Care Management and Health Care Economics, Vienna University of Economics and Business, Austria.*
Clemens Huber, PhD, MPharm, *is Research Assistant, Research Institute for Health Care Management and Health Care Economics, Vienna University of Economics and Business, Austria.*
Andreas Valentin, MD, MBA, *is Head of 2nd Medical Department for Intensive Care, General and Medical ICU, Rudolfstiftung Hospital, Vienna, Austria.*
Guido Strunk, PhD, *is Founder of Complexity-Research, Vienna, Austria, and Senior Researcher, Department of Business Administration and Economic Education, University of Technology, Dortmund, Germany.*

The authors have disclosed that they have no significant relationships with, or financial interest in, any commercial companies pertaining to this article.

DOI: 10.1097/HMR.0b013e318272935a

Health Care Manage Rev, 2012, 00(0), 00–00

Linear regression was employed to identify the effects of the predictor variables on error rate as well as interaction effects between safety tools and safety climate.
Findings: Higher workload has a detrimental effect on safety, whereas safety climate—unlike the examined safety tools—has a virtually equal opposite effect. Correlations between safety tools and safety climate as well as their interaction effects on error rate are mostly nonsignificant.
Practice Implications: Increased workload and capacity utilization increase the occurrence of medical error, an effect that can be offset by a positive safety climate but not by formally implemented safety procedures and policies.

Reducing the number of medical errors as a means of improving patient safety has been in the spotlight for over a decade now (e.g., Kohn, Corrigan, & Donaldson, 2000). At the same time, hospitals, like other high-hazard industries, face the basic dilemma of "production versus protection" (Reason, 1997, p. 4), that is, providing quick and (cost-)effective treatment while ensuring patient safety and avoiding medical errors. In addition, pressure on hospitals toward efficiency and cost reduction has risen over the last years (Hsieh, Clement, & Bazzoli, 2010), with potentially detrimental effects on safety (Hansez & Chmiel, 2010) as well as error-induced cost (van den Bos et al., 2011).

One central determinant of medical error prevention consists in the so-called safety culture (Naveh, Katz-Navon, & Stern, 2005), which is defined by a pioneer source as "the product of individual and group values, attitudes, perceptions, competencies and patterns of behavior that determine the commitment to, and the style and proficiency of, an organisation's safety management" (International Safety Advisory Group, 1991, p. 23). A related construct is safety climate, which refers primarily to "the shared perceptions of employees concerning the degree to which safety is a top priority for employees within the organization" (Stern, Katz-Navon, & Naveh, 2011, p. 57).

Referring to the distinction made by Reason (1997) between error management tools (p. 130) and safety climate in the sense of shared values and attitudes (p. 194), we distinguish between two elements of safety culture—the above-mentioned shared perceptions and attitudes concerning safety, hence termed *safety climate*, and tangible, formalized safety procedures and tools (e.g., Critical Incident Reporting Systems [CIRS], checklists, barcodes, safety audits, or mission statements), hence termed *safety tools*. This distinction between safety climate and safety tools is not always made in extant literature, with some authors seeing observable safety features as a part of safety climate both in theory on organizational culture and climate and empirical studies in health care and other fields (e.g., Katz-Navon, Naveh, & Stern, 2005; Neal, Griffin, & Hart, 2000; Ostroff, Kinicki, & Tamkins, 2003). Indeed, some safety tools like safety audits and mission statements are arguably quite closely related to the perception- and attitude-centered notion of safety climate, whereas others (e.g., checklists and barcodes) represent a much more tangible and procedural aspect of work design. One central distinction from safety climate, however, is that from a management perspective, all safety tools can be formally designed and implemented.

This difference applies to several similar distinctions in related fields, too, like that between hard and soft aspects of total quality management (e.g., Wilkinson, Redman, Snape, & Marchington, 1998) and/or organizational culture (e.g., the 7S model, with strategy, structure, and systems as "hard" aspects and skills, style, staff, and shared values as "soft" aspects; e.g., Waterman, Peters, & Phillips, 1980). It also relates to two different approaches to safety management in health care, where Neal et al. (2000) differentiate between *safety compliance*, which "involves adhering to safety procedures and carrying out work in a safe manner," and *safety participation* consisting of "helping coworkers, promoting the safety program within the workplace, demonstrating initiative, and putting effort into improving safety in the workplace" (p. 101). In a similar vein, Khatri, Baveja, Boren, and Mammo (2006) contrast *control-based management* with *commitment-based management*. The former is characterized, among other things, by an "emphasis on compliance/obedience" (p. 119) and a "bureaucratic, rule-based culture" (p. 125), aiming at regulation and standardization by a centralized quality department (Khatri, Brown, & Hicks, 2009). By contrast, commitment-based management attaches importance to communication and teamwork and makes "quality and safety issues permeate the entire organization" (Khatri et al., 2009, p. 318). In spite of all these dichotomies represented by the mentioned concepts, these aspects (climate vs. tools, soft vs. hard, control/compliance vs. commitment/participation) are arguably not mutually exclusive, which raises the question to what extent the potential effects of these approaches are independent, antagonistic, or mutually reinforcing.

On the basis of these considerations, our study investigates the following research questions:

1) How strongly does production pressure (reflected in higher capacity utilization and/or staff workload) affect error rate?
2) To what extent do safety tools (see above) affect error rate?
3) To what extent do shared perceptions and attitudes among frontline staff concerning safety (safety climate) affect error rate?

4) Do safety tools and safety climate interact in affecting error rate?

In investigating these questions, our study contributes to extant research in three ways. First, although the effect of safety culture/climate on patient safety has been previously examined, the influence of staff workload/capacity utilization as factors that are thought to adversely affect patient safety has hitherto rarely been investigated let alone directly compared with safety climate/safety tools as elements of safety culture that are thought to positively affect patient safety. Second, it distinguishes between safety climate and safety tools as two elements of safety culture that represent different yet complimentary approaches to safety management and relates them to extant concepts of organizational and safety culture. Third, it consequently examines the main and interaction effects of these two elements (safety climate and safety tools) on patient safety as opposed to the influence of production pressure in a particularly error-prone medical specialty (intensive care).

Framework and Hypotheses

Regarding the above-mentioned conflict between production and protection, there are several approaches that aim at improving both safety and efficiency via process optimization, like, e.g., Six Sigma or Lean Healthcare, and although some sources suggest that the conclusion regarding the overall benefit of Lean Healthcare implementation is somewhat sobering (Grove, Meredith, MacIntyre, Angelis, & Neailey, 2006; Radnor, Holweg, & Waring, 2012), there are indeed successful instances in health care of reconciling these contradictory goals (Radnor, 2011, p. 4).

One central determinant whether such initiatives are beneficial to patient safety or not is arguably whether the modified processes lead to a reduction in workload for frontline staff (e.g., via less red tape, smoother handovers, etc.) or rather increase workload and production pressure on staff in an effort to "do more with less." Several empirical findings suggest a tendency toward the latter in practice (Mehri, 2006; Schön, Bergquist, & Klefsjö, 2010), suggesting that, although efforts for increased efficiency need not necessarily lead to increased workload, in practice they apparently do (see also Rasmussen, 1997), and indeed the premise that production pressure increases the rate of medical error is corroborated by several studies (e.g., Hansez & Chmiel, 2010; Valentin et al., 2009). Hence, our first hypothesis:

H1: A higher degree of capacity utilization and/or workload is associated with a higher error rate.

Safety-related efforts, by contrast, contribute to increased safety margins (Rasmussen, 1997, p. 190) and should thus reduce the occurrence of errors. Although this applies to both tools and climate, their pathways of effectiveness are arguably different (although not mutually exclusive). Procedural tools like checklists and barcodes serve as "technical" barriers against errors. Those tools that are closer to the perception-related notion of safety climate (safety audits and mission statement) primarily represent an "official" and observable reflection of the prevailing safety climate. Finally, training and CIRS emphasize the learning aspect of a safety culture and mainly aim at a better recognition of and appropriate response to identified threats to safety and/or sources of error. Indeed, the notion that safety tools reduce medical error is supported by empirical findings (Shojania, Duncan, McDonald, & Wachter, 2002), specifically for CIRS (Valentin et al., 2009), audits and mission statements (Ovretveit, 1999), or standard operating procedures for handovers (Catchpole et al., 2007):

H2: Implementation of safety tools is associated with a lower error rate.

A good safety climate, on the other hand, establishes a priority for safety compared with efficiency goals (e.g., Zohar, 2000) and should thus contribute to less treatment errors as well, a notion supported by numerous studies (Clarke, 2006; Hofmann & Mark, 2006; Huang et al., 2010; Katz-Navon et al., 2005; Naveh et al., 2005; Neal & Griffin, 2006; Singer, Lin, Falwell, Gaba, & Baker, 2009; Stern et al., 2011; Vogus & Sutcliffe, 2007):

H3: A better safety climate is associated with a lower error rate.

Concerning the respective effectiveness of these two safety culture elements (safety tools vs. safety climate), several sources suggest that a formalized, control-based approach might not be as important a determinant of (patient) safety as an approach emphasizing shared views and social norms concerning safety among staff (Podgórski, 2010), which applies to the medical field, too (Khatri, Halbesleben, Petroski, & Meyer, 2007; Waring, 2009). In addition, compared with other high-risk industries, work processes in health care depend to a greater extent on the decisions and discretion of frontline staff, with less reliance on automation and standardization (Gaba, 2000), which further underlines the importance of an emphasis on safety among frontline staff for error prevention:

H4: Safety climate reduces medical error more than safety tools.

Finally, we argue that safety tools and safety climate reinforce each other. For the "technical barrier" type of tools (checklists and barcodes), a good safety climate enhances compliance with safety procedures, strengthening their effectiveness. Safety tools that focus on learning (training and CIRS) provide increased knowledge and awareness about important threats and errors, adding to the attitudinal aspect represented by climate. Finally, safety tools that represent the observable complement of a good safety climate may degenerate into mere rhetoric and pretense in case of a weak safety climate (Stern et al., 2011):

H5: Safety climate and safety tools interact in a mutually reinforcing way in their influence on error rate.

Method

Study Design

The study was conducted as a joint research project by the Austrian Center for Documentation and Quality Assurance in Intensive Care, together with the Research Institute for Healthcare Management and complexity-research.com of Vienna University of Economics and Business (WU). We drew on a sample of ICUs recruited by cooperating associations already in 2004 (apart from the Austrian Center for Documentation and Quality Assurance in Intensive Care, these were the German Interdisciplinary Federation of Intensive Care Medicine and Emergency Medicine and the Swiss Society of Intensive Care Medicine), which have an explicit interest in patient safety and prior experience with the recording of medical errors for scientific studies (Valentin et al., 2006, 2009). Required sample size was estimated beforehand based on expected effect sizes and validity of measures. Participating units could choose between two possible dates on which to begin a 48-hour observation period.

To avoid problems associated with the use of archive data (Clarke, 2006), this study used a prospective approach to error measurement. Medical errors that actually or potentially caused harm to a patient were recorded over a 48-hour period (which included change of shift for nurses and physicians), with safety climate measured during that time, too. To approach the problem of variation across specialties, we focused on a particular medical field: intensive care units (ICUs). Our choice was motivated by the fact that intensive care is particularly prone to error. The combination of complexity and potential for great harm makes it even more fraught with risk than other medical fields (Valentin et al., 2009). A study of ICUs therefore seemed especially well suited to investigating the research questions set out above.

Sample

The final study sample consisted of 378 patients from 57 ICUs in as many (predominantly larger urban) hospitals. The data obtained during observation are summarized in Table 1, in which the values of continuous variables are expressed as mean ± standard deviation and those of categorical variables as frequencies and percentages. It is noteworthy that the overall mean NEMS score of 35.5 (±11.5) score points per patient suggests a high level of treatment intensity.

Figures for the safety climate dimensions are based on responses from 549 nurses and 185 physicians, representing response rates of 41.4% and 35.2%, respectively. Response rate standard deviation across ICUs is 17.7 %. ICUs with less than five respondents for the safety climate survey were excluded from the analyses. The mean age of respondents was 37.5 (±9.2) years, and their mean professional experience was 14.4 (±8.7) years. Some 17% of respondents had managerial functions. The following results are based on those 378 patients who remained in an ICU throughout the entire observation period. An analysis of patients with an observation time of at least 12 hours (795 in total) led to similar results and is therefore not reported.

Data Collection

A study manual and all questionnaires were made available in German on the study Web site (www.sifim.org). Detailed handbooks as well as simulated training sessions were provided to all participants. Additional support and supervision was offered during data collection. Data collection itself was highly standardized in order to reduce the risk of errors. For each patient observed, basic demographic characteristics (age and gender) were recorded by staff on a form, along with the start and end of the individual observation period. Throughout that time, the form remained at the patient position and was used to record the occurrence of any of the predefined medical errors relating to the patient concerned, irrespective of the staff member recording them. Every contributor to the questionnaire could see all previous entries, making duplicate reporting highly unlikely. In addition, each ICU staff member was asked to fill in the safety climate questionnaire available on the study Web site. For both patients and medical personnel, data acquisition and reporting were anonymous. Last but not least, the characteristics of individual ICUs were reported by a designated coordinator, who was also responsible for briefing the unit staff, facilitating data collection, and transmitting information to the study database.

Measures

Concerning the operationalization of "production pressure," our measure of *capacity utilization* was constructed from three separate indicators expressing the relation between patient numbers and resources, as follows:

- unit occupancy, defined as the ratio of "bed usage" (sum of all hours of patient observation) to "bed potential" (number of ICU beds times 48, the observation period duration in hours);
- nurse-to-patient ratio, defined as "total nurse working hours" (mean number of nurses times 48) divided by "bed usage";
- physician-to-patient ratio, defined by analogy with 2.

Our *workload* measure was also composite and based on three indicators, in this case all were concerned with the level of care provided per patient. They were

- the number of medications received by patients in relation to nurse numbers;
- the number of tubes, catheter, probes, lines, and/or drains inserted into patients, again in relation to nurses;
- the so-called Nine Equivalents of Nursing Manpower Use Score (NEMS; Reis Miranda, Moreno, & Iapichino, 1997), consisting of basic monitoring, intravenous medication, mechanic airway, supplementary airway, singular vasoactive medication, multiple vascoactive medication, dialysis, and specific interventions.

The composite variable showed no linear correlation with error rates, so its natural logarithm was used as a linear correction. The results obtained in both cases were similar.

On the "protection" side, the safety tools we included correspond to those most commonly used in German ICUs, the use of which was measured by means of six categorical (yes/no) indicators for each ICU: existence of a *mission statement* stressing patient safety, regular performance of *safety audits*, implementation of *CIRS*, regular *training* to improve patient safety, and implementation of SOP checklists and/or *barcodes* or electronic tools to avoid medication errors; the latter two being safety tools directly related to daily medical tasks.

Concerning the measurement of *safety climate*, besides operationalizations of safety climate in general (e.g., Burke, Sarpy, Tesluk, & Smith-Crowe, 2002; Zohar, 2000), several instruments specifically referring to patient safety have been developed (e.g., Colla, Bracken, Kinney, & Weeks, 2005; Singer et al., 2007). However, no reliable and valid instrument in German was available, and a literal translation makes little sense from a methodological point of view (Hambleton, Merenda, & Spielberger, 2005; Muñiz & Bartram, 2007). This led to the development of the Vienna Safety Climate Questionnaire (VSCQ) in an iterative process using data from a total of 1,968 respondents and including multidimensional scaling (n = 61), item selection and exploratory and confirmatory factor analysis (n = 954), an additional confirmatory factor analysis sample recruited after final item selection (n = 761), and validity (for all samples). This process was completed before the current study was conducted.

Despite going beyond mere translation, the VSCQ is based on existing instruments for measuring safety climate (e.g., Colla et al., 2005; Flin, Burns, Mearns, Yule, & Robertson, 2006; Schutz, Counte, & Meurer, 2007; Singer et al., 2007; Weingart, Farbstein, Davis, & Phillips, 2004). Accordingly, the scales included in this study are closely related to extant dimensions of safety climate. For instance, taking the Patient Safety Climate in Healthcare Organizations survey by Singer and colleagues (2007) as a reference, *Management Commitment to Patient Safety* largely corresponds to "Senior managers' engagement," *Organizational Learning* to "Learning," *Communication and Cooperation Regarding Patient Safety* to "Fear of shame" and "Fear of blame," and *Attitude Toward Safety Management* to "Overall emphasis on safety" and "Unit safety norms."

Despite considerable intercorrelations of the VSCQ scales (see Table 3), a confirmatory factor analysis based on three samples (total n = 2,608, with the sample used in this study included) supports the posited dimensions and item allocation, with RMSEA $\leq$ .05 and $\chi^2/df < 3.6$ in all cases (Browne & Cudeck, 1993). The questionnaire has been successfully validated (Steyrer, Latzke, Pils, Vetter, & Strunk, 2011), and the performance of the VSCQ on relevant psychometric criteria (see Table 2) is so satisfactory that it is now listed in the catalogue of safety culture instruments used in member states of the European Union compiled by the European Network for Patient Safety.

Although safety climate was ascertained individually, it represents a construct on group and/or organizational level (here, ICU); the same ultimately applies to error rate (Neal & Griffin, 2006), with this study not focusing on the probability of an individual patient encountering a medical error but rather on the performance of the whole ICU and its relationship to ICU-wide safety climate and safety tools. All analyses were therefore conducted at ICU level, with aggregated safety climate scale values. Table 1 shows the aggregation indices ICC(1), ICC(k) (notation follows McGraw, Kenneth, & Wong, 1996), and the interquartile range for $r_{WG(J)}$ for all scales. Although using the uniform distribution as reference for the $r_{WG(J)}$ calculations makes them upper bound values (LeBreton & Senter, 2008), these values as well as the ICC(k) values are well above the commonly cited threshold of 0.7, and ICC(1) values are well above 0.05 (Bliese, 2000; James, Demaree, & Wolf, 1993), suggesting that analysis on an aggregate level is appropriate here.

As regards our dependent variable, a *medical error* was defined, in line with previous studies on patient safety in ICUs, as an event that harmed or could have harmed a patient, whether by omission or commission (Thomas et al., 2000). We focused on intensive care routine processes and recorded errors in seven distinct categories based on insights from two previous studies (Valentin et al., 2009; Valentin et al., 2006): administration of medication (wrong dose, time and medication, wrong means of administration, or missed administration) and unplanned dislodgement of airways, arterial lines, central venous catheters, urinary catheters, enteral nutrition probes, or drains. The *error rate* was the ratio of patients affected by errors in a given ICU to the total number of patients in that unit.

Analysis

In several other studies on the relationship between safety climate and medical errors, the occurrence of errors represents a count variable of rare events, which is why the authors use Poisson or negative binomial models to analyze their data

Table 1

Characteristics of 378 patients from 57 intensive care units

	Count (%) or mean ± *SD*
Patients	
Age (in years)	60.3 ± 19.1
Women (%)	141 (37.3)
Men (%)	237 (62.7)
Number of tubes, lines, drains, etc., per patient	4.7 ± 2.1
Number of medications applications) per patient	49.4 ± 27.2
NEMS score (points) per patient	35.5 ± 11.5
ICUs	
Number (%) by type of unit	
Mixed	25 (43.9)
Medical	16 (28.1)
Surgical	12 (21.1)
Pediatric	2 (3.5)
Other	2 (3.5)
Number (%) by unit size (beds)	
<8	13 (22.8)
8–12	27 (47.4)
>12	17 (29.8)
Number (%) by hospital size (beds)*	
<400	17 (29.8)
400–900	23 (40.4)
>900	14 (24.6)
Number of physicians	12.6 ± 7.1
Number of nurses	41.6 ± 26.3
Number of patients (during a 48-hour period)	6.6 ± 4.7
Safety climate total score	49.7 ± 19.0
Unit occupancy (0–1)	0.79 ± 0.19
Physicians to patient ratio (0–1)	0.35 ± 0.25
Nurses to patient ratio (0–1)	0.80 ± 0.43
Total number of tubes, etc., per nurse	8.40 ± 4.60
Total number of medications per nurse	66.79 ± 35.02
Total NEMS score (points) per patient	68.10 ± 35.60
ICUs with Critical Incident Reporting Systems	33 (57.9)
ICUs with regular safety audit	11 (19.3)
ICUs with barcodes or other electronic measures designed to prevent medication errors	14 (24.6)
ICUs with checklists	51 (89.5)
Error rate (number of patients with errors / all patients)	0.44 ± 0.32

*Data from 3 ICUs missing.

(e.g., Hofmann & Mark, 2006; Katz-Navon et al., 2005; Singer et al., 2009). The distribution of our criterion variable, by contrast, does not correspond to a count of rare events, with a median and mean of 0.44 (*SD*, ±0.32), a merely slightly positive skewness value (0.23), and even a negative excess kurtosis (−0.94), which is why OLS regression and correlations are employed in this study (as for instance in Naveh et al., 2005).

On the basis of the correlation results presented in Table 3, we chose the predictors for the regression analyses, with a focus on keeping the models parsimonious and minimizing multicollinearity. To test H4, we conducted Steiger's Z test (Steiger, 1980), comparing the (bivariate) effect of safety climate on error rate to that of the safety tools.

Findings

Bivariate analysis reveals a high intercorrelation between capacity utilization and workload (0.72) and of both with error rate (capacity utilization, 0.25; workload, 0.40), which is why we only retained workload as a predictor for the following regression analyses. Similarly, although all safety climate scales are significantly correlated with error rate, the high intercorrelations of the VSCQ scales prompted us to choose *Attitude Toward Safety Management* as safety climate proxy for the regression analyses, because this VSCQ dimension corresponds most closely to the definition of safety climate presented above.

The intercorrelations concerning the use of safety tools are generally much lower, so we conducted separate analyses for each safety tool to assess the respective impact of workload, safety attitudes, and each safety tool on error rate. We ran additional analyses including the number of safety tools as a predictor, but this was merely correlated with the individual safety tools and (like mission statements, albeit more weakly) with two safety climate scales (management commitment and organizational learning) but not with either capacity utilization, workload, or error rate. On the basis of the finding by Katz-Navon et al. (2005), which posits that in contrast to "optimal" level of detail of safety procedures both scarcely and overly detailed safety procedures lead to a higher error rate, we checked for a curvilinear relationship between number of tools and error rate, too, but found none.

Table 4 presents the results of the regression analyses, with error rate always as the dependent variable and workload, *Attitude Toward Safety Management*, its interaction with the respective safety tool, and the safety tool main effect as predictors. Both correlation and regression results support H1 and H3 (workload increases the error rate, safety climate reduces it) but strongly contradict H2: None of the examined safety tools significantly reduces error rate. More precisely, there was no significant relationship with error rate for any of the safety tools, except for a marginally significant (two-tailed $p < .10$) main effect of safety audits opposite to the predicted direction.

Consequently, the regression results seem to support H4 as well (safety climate reduces the error rate more than safety tools). To properly test for significance, we compared

Table 2

Sample items and scale properties for the four Vienna Safety Climate Questionnaire scales (different *n* applies to consistency and validity criteria as opposed to aggregation indices)

	n = 1,968		*n* = 734	
1. Management commitment to patient safety: ''Management encourages the staff to report incidents''/''Most superiors know that they act as role models regarding safety issues''	Items	10	ICC(1)	.19
	α	.91	ICC(*k*)	.81
	Validity	.40	$r_{WG(J)}$	.83
			(IQR)	.92
2. Organizational learning: ''After critical incidents, we undertake major efforts to investigate the cause''/''We investigate critical incidents to draw new conclusions for our actions''	Items	10	ICC(1)	.18
	α	.86	ICC(*k*)	.80
	Validity	.43	$r_{WG(J)}$	.83
			(IQR)	.92
3. Communication and cooperation regarding patient safety: ''If someone notices that a colleague makes a mistake it is no problem to address it''/''Mistakes are personalized and individual persons are charged'' (−)	Items	10	ICC(1)	.17
	α	.87	ICC(*k*)	.78
	Validity	.37	$r_{WG(J)}$	.80
			(IQR)	.90
4. Attitude toward safety management: ''Process quality and error management are rather lip service than lived practice'' (−)/'' Sometimes it seems like a waste of time, which is taken for safety efforts without much effect'' (-)	Items	10	ICC(1)	.15
	α	.82	ICC(*k*)	.76
	Validity	.39	$r_{WG(J)}$	.83
			(IQR)	.90

the correlation coefficients of safety climate and safety tools with error rate; in all instances, the difference in favor of safety climate is significant at the 5% level (one-tailed; Steiger's Z ranges between 2.05 for checklists and 3.26 for CIRS). To check whether this difference in effect might be a statistical artifact due to the limited variation of a dichotomous variable as opposed to a scale variable, we performed a median split on all safety climate variables and entered them as well as the safety tools variables into an ANOVA. The respective eta-square values yield the same picture as the effect size *r*, suggesting that the observed effects are not a result of different scale properties.

Table 3

Overview of means, standard deviations, and correlations (intensive care unit level)

	Mean (*SD*)	1	2	3	4	5	6	7	8	9	10	11	12
1. Management commitment to patient safety	−0.02 (.58)	—											
2. Organizational learning	−0.02 (.51)	.85	—										
3. Communication and cooperation regarding patient safety	−0.00 (.54)	.74	.74	—									
4. Attitude toward safety management	−0.00 (.53)	.81	.81	.86	—								
5. Critical incident reporting systems	0.58 (.50)	.08	.07	.05	.05	—							
6. Regular safety audit	0.19 (.40)	.17	.06	.10	.12	-.03	—						
7. Barcodes	0.25 (.43)	−.08	−.08	−.06	−.12	−.01	−.07	—					
8. SOP checklists	0.89 (.31)	.22	.17	.07	.10	−.06	.17	.06	—				
9. Training	0.46 (.50)	.24	.25	.12	.12	−.00	.09	.13	.20	—			
10. Mission statement	0.60 (.50)	.36	.37	.19	.19	.02	.04	.05	.19	.54	—		
11. Capacity utilization	0.63 (.42)	−.19	−.20	−.05	−.28	.00	−.05	.26	−.02	−.02	.12	—	
12. Workload	−1.80 (.45)	−.04	−.14	−.03	−.23	.04	.04	.24	.10	.10	.04	.72	—
13. Error rate	0.44 (.32)	−.29	−.36	−.26	−.44	.16	.13	.19	−.09	−.08	−.07	.25	.40

Note. n = 57. Correlations ≥ .27 are significant at the 5% level; ≥ .34 significant at the 1% level (two tailed).

Table 4

Regression model results (standardized beta coefficients)

Dependent variable: Error rate	CIRS	Audits	Training	Mission statement	Barcodes	Checklists
Workload	.31**	.27**	.32**	.30**	.37**	.33**
Safety climate (Attitude Toward Safety Management)	−.37**	−.43**	−.35**	−.36**	−.28*	−.35**
Safety tools (see respective column)	.16***	.21*	−.06	−.02	.10	−.08
Safety climate × Safety tools	−.06	−.18***	−.07	−.15	.24*	.03
R^2	.32	.34	.30	.31	.34	.30

Note. CIRS = critical incident reporting systems.

**p < .01. *p < .05. †p < .10 (one tailed).

Concerning H5, the results do not support our assumption. There was just one merely marginally significant interaction with audits in the predicted direction (one-tailed $p < .08$) and even a "reversed" interaction with barcodes opposite to our assumption (two-tailed $p < .07$), suggesting that barcodes even contribute to an increased error rate in case of a good safety climate.

Limitations

Although this study has some strengths, for example, a prospective designs in which accidents were measured following the measurement of safety climate (Clarke, 2006), considering safety climate at the clinical area level rather than at the hospital level (Sexton et al., 2006), an absence of common method variance, and the introduction of a now recognized German instrument for measuring safety climate that can relate to tried and tested instruments for measuring safety climate in English (e.g., Schutz et al., 2007; Singer et al., 2007), it has some shortcomings that should be considered when interpreting the results, too. First, it is not based on a random sample: Units were self-selecting in that they chose to respond to the call for participation and had previous experience with external studies on medical error in ICUs (Valentin et al., 2009; Valentin et al., 2006), although not in connection with safety climate and/or tools. As a result, the possibility of sample bias cannot be completely excluded, but such a bias should rather consist in more accurate reporting of errors and a better (and/or bounded below) safety climate compared with "average" ICUs, which would not weaken our results. Moreover, despite an ex ante research design controlling for a range of variables and designing the project based on state-of-the-art studies in the areas of error and safety culture measurement, which enabled us to include a great number of potential error rate predictors, some degree of distortion from omitted variables cannot be entirely ruled out. Concerning staff workload during the observation period, we were unable to record the precise number and time of individual treatments (i.e., the exact distribution of workload) and therefore have to assume that these were approximately evenly distributed.

Other conceivable causes of distortion lie in the methods we used to measure safety performance and safety climate. Our metric of the former was based on a limited set of error types. Although these were selected for being the most important ones according to previous studies in ICUs (Valentin et al., 2006, 2009), the inclusion of other types might have produced different results. In addition, error incidence was recorded by hospital staff; therefore, it is possible that some over- or under-reporting occurred, although the high values recorded make under-reporting appear somewhat implausible. Similarly, the VSCQ relies on employees responding in a valid and truthful manner, something that cannot be guaranteed; on the other hand, the obtained values do not hint at any biases like social desirability or similar and the data were regularly verified and checked for any apparent inconsistencies during the execution of the study.

Discussion and Practice Implications

Our findings for hospital care in ICUs support the notion of a production versus protection trade-off in two ways. First, although it is arguably possible to implement measures that increase both efficiency and safety (as mentioned above in connection with Lean Healthcare), simply aiming at "doing more with less" obviously comes with a safety penalty, as reflected in the considerable effect of staff workload on error rate. Second, safety culture can apparently thwart the detrimental effects of increased workload on safety quite effectively. However, corroborating the findings by Khatri et al. (2007), a management approach relying on safety tools and procedures to compensate for the erosion of safety margins by increased production pressure might rather abet than abate medical errors.

Regarding any interaction between safety climate and safety tools, our findings were generally weak. For the two "climate-related" tools (mission statements and safety audits), only the implementation of a mission statement was significantly correlated with some of the safety climate dimensions, and interactions with safety climate, although in the predicted direction, failed to reach statistical significance for our sample. The same applies to an even larger extent to the safety tools aimed at learning and increased awareness (training and CIRS), where no interaction was found at all. Concerning training as ascertained in our study, we argue that it consists mainly in off-the-job "ground school" lessons, representing individual "single-loop learning," which is just a precursor of "double-loop learning" that takes place and becomes effective on an organizational level (Argyris, 1982). CIRS, on the other hand, represents such double-loop learning but still failed to interact with safety climate in the predicted way. One possible explanation for this is that CIRS in hospitals is still in a somewhat exploratory phase, especially that staff might not (yet) trust it to be positively anonymous and nonpunitive, because it is administered and evaluated by management. A similar observation was made in another high-hazard industry, where the Aviation Safety Reporting System, now a valued and effective safety tool, initially failed to work until its anonymity and nonpunitive nature were clearly and credibly established by introducing a neutral organization functioning as a broker, with the supervisory authorities having strictly no access to any clues concerning the identity of the reporting persons (Stolzer, Halford, & Goglia, 2011, p. 57f).

By contrast, checklists and barcodes address daily work routines in a highly structured and rigid manner. Here, the marginally significant interaction found in our data for use of barcodes (i.e., it lowers the error rate in case of a poor safety climate but contributes to an augmentation of errors in ICUs with a good safety climate) may hint at staff not using the implemented technology in case of a weak safety climate, therefore detecting less errors, whereas in case of a sound safety climate, staff may be more likely to use the available technology because they recognize its role in promoting safety and, as a result, detect more errors. Another possible explanation relating to the finding by Katz-Navon et al. (2005) that overly detailed safety procedures increase the number of medical errors could be that, in units with a firm safety attitude, the implementation of barcodes represents an additional burden (despite falling slightly short of statistical significance for two-tailed testing, barcodes are the only safety tools with an appreciable relationship to workload and capacity utilization) but actually serves as a barrier against errors in the absence of such an emphasis on safety. The failure to find any comparable (or contrasting) effect for checklists might be largely rooted in the fact that almost all participating ICUs used checklists (see bottom of Table 1), making this safety tool more of a constant than a variable.

In conclusion, besides underscoring the importance of a sound safety climate to prevent a decline in patient safety resulting from increased production pressure, our findings support the view that safety is created at the "sharp end" (Dekker, 2006), consequently strongly cautioning against a "shortcut" approach of trying to achieve safety via procedures and policies implemented top-down.

Acknowledgments

We thank the ICU teams and ICU study coordinators who participated in this study and contributed in an open-minded and dedicated fashion. This study was conducted as a joint research project and financed by the Austrian Center for Documentation and Quality Assurance in Intensive Care and the Interdisciplinary Group of Management and Organizational Behavior of the Vienna University of Economics and Business. The study was supported by the German Interdisciplinary Federation of Intensive Care Medicine and Emergency Medicine (DIVI) and the Swiss Society of Intensive Care Medicine. Last but not least, our thanks go to two anonymous reviewers and the editor for numerous valuable hints and suggestions, which have contributed to equally numerous improvements throughout the article.

References

Argyris, C. (1982). *Reasoning, learning, and action: Individual and organizational*. San Francisco, CA: Jossey-Bass.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. J. Klein & S. W. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349–381). San Francisco, CA: Jossey-Bass.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing equation model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Burke, M. J., Sarpy, S. A., Tesluk, P. E., & Smith-Crowe, K. (2002). General safety performance: A test of a grounded theoretical model. *Personnel Psychology, 55*, 429–457.

Catchpole, K. R., de Leval, M. R., McEwan, A., Pigott, N., Elliott, M. J., McQuillan, A., … Goldman, A. J. (2007). Patient handover from surgery to intensive care: Using Formula 1 pit-stop and aviation models to improve safety and quality. *Paediatric Anaesthesia, 17*(5), 470–478.

Clarke, S. (2006). The relationship between safety climate and safety performance: A meta-analytic review. *Journal of Occupational Health Psychology, 11*(4), 315–327.

Colla, J. B., Bracken, A. C., Kinney, L. M., & Weeks, W. B. (2005). Measuring patient safety climate: A review of surveys. *Quality and Safety in Health Care, 14*(5), 364–366.

Dekker, S. (2006). *The field guide to understanding human error*. Farnham, UK: Ashgate.

Flin, R., Burns, C., Mearns, K., Yule, S., & Robertson, E. M. (2006). Measuring safety climate in health care. *Quality and Safety in Health Care, 15*(2), 109–115.

Gaba, D. M. (2000). Structural and organizational issues in patient safety: A comparison of health care to other high-hazard industries. *California Management Review, 43*(1), 83–102.

Grove, A. L., Meredith, J. O., MacIntyre, M., Angelis, J., & Neailey, K. (2006). UK health visiting: Challenges faced during lean implementation. *Leadership in Health Services, 23*(3), 204–218.

Hambleton, R. K., Merenda, P. F., & Spielberger, C. D. (2005). *Adapting educational and psychological tests for cross-cultural assessment*. Mahwah, NJ: Lawrence Erlbaum.

Hansez, I., & Chmiel, N. (2010). Safety behavior: Job demands, job resources, and perceived management commitment to safety. *Journal of Occupational Health Psychology, 15*(3), 267–278.

Hofmann, D. A., & Mark, B. (2006). An investigation of the relationship between safety climate and medication errors as well as other nurse and patient outcomes. *Personnel Psychology, 59*, 847–869.

Hsieh, H. M., Clement, D. G., & Bazzoli, G. J. (2010). Impacts of market and organizational characteristics on hospital efficiency and uncompensated care. *Health Care Management Review, 35*(1), 77–87.

Huang, D. T., Clermont, G., Kong, L., Weissfeld, L. A., Sexton, J. B., Rowan, K. M., & Angus, D. C. (2010). Intensive care unit safety culture and outcomes: A US multicenter study. *International Journal for Quality in Healthcare, 22*(3), 151–161.

International Safety Advisory Group. (1991). *Safety culture*. Vienna, Austria: International Atomic Energy Agency.

James, L. R., Demaree, R. G., & Wolf, G. (1993). An assessment of within-group interrater agreement. *Journal of Applied Psychology, 78*, 306–309.

Katz-Navon, T., Naveh, E., & Stern, Z. (2005). Safety climate in health care organizations: A multidimensional approach. *Academy of Management Journal, 48*(6), 1075–1089.

Khatri, N., Baveja, A., Boren, S. A., & Mammo, A. (2006). Medical errors and quality of care: From control to commitment. *California Management Review, 48*(3), 115–141.

Khatri, N., Brown, G. D., & Hicks, L. L. (2009). From a blame culture to a just culture in health care. *Health Care Management Review, 34*(4), 312–322.

Khatri, N., Halbesleben, J. R., Petroski, G. F., & Meyer, W. (2007). Relationship between management philosophy and clinical outcomes. *Health Care Management Review, 32*(2), 128–139.

Kohn, L. T., Corrigan, J. M., & Donaldson, M. S. (2000). *To err is human: Building a safer health system*. Washington, DC: National Academy Press.

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11*(4), 815–852.

McGraw, K. O., Kenneth, O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30–46.

Mehri, D. (2006). The darker side of lean: An insider's perspective on the realities of the Toyota production system. *Academy of Management Perspectives, 20*(2), 21–42.

Muñiz, J., & Bartram, D. (2007). Improving international tests and testing. *European Psychologist, 12*(3), 206–219.

Naveh, E., Katz-Navon, T., & Stern, Z. (2005). Treatment errors in healthcare: A safety climate approach. *Management Science, 51*(6), 948–960.

Neal, A., & Griffin, M. A. (2006). A study of the lagged relationships among safety climate, safety motivation, safety behavior, and accidents at the individual and group levels. *Journal of Applied Psychology, 91*(4), 946–953.

Neal, A., Griffin, M. A., & Hart, P. M. (2000). The impact of organizational climate on safety climate and individual behavior. *Safety Science, 34*, 99–109.

Ostroff, C., Kinicki, A. J., & Tamkins, M. M. (2003). Organizational culture and climate. In W. C. Borman, D. R. Ilgen, & R. J. Klimoski (Eds.), *Handbook of psychology: Industrial and organizational psychology* (Vol. 12, pp. 565–593). New York, NY: Wiley.

Ovretveit, J. (1999). Total quality management in European healthcare. *International Journal of Health Care Quality Assurance, 13*, 74–79.

Podgórski, D. (2010). The use of tacit knowledge in occupational safety and health management systems. *International Journal of Occupational Safety and Ergonomics, 16*(3), 283–310.

Radnor, Z. J. (2011). Implementing lean in health care: Making the link between the approach, readiness, and sustainability. *International Journal of Industrial Engineering and Management, 2*(1), 1–12.

Radnor, Z. J., Holweg, M., & Waring, J. J. (2012). Lean in healthcare: The unfilled promise? *Social Science and Medicine, 74*, 364–371.

Rasmussen, J. (1997). Risk management in a dynamic society: A modelling problem. *Safety Science, 27*(2–3), 183–213.

Reason, J. T. (1997). *Managing the risks of organizational accidents*. Farnham, UK: Ashgate.

Reis Miranda, D., Moreno, R., & Iapichino, G. (1997). Nine equivalents of nursing manpower use score (NEMS). *Intensive Care Medicine, 23*(7), 760–765.

Schön, K., Bergquist, B., & Klefsjö, B. (2010). The consequences of Six Sigma on job satisfaction: A study at three companies in Sweden. *International Journal of Lean Six Sigma, 1*(2), 99–118.

Schutz, A. L., Counte, M. A., & Meurer, S. (2007). Development of a patient safety culture measurement tool for ambulatory health care settings: Analysis of content validity. *Health Care Management Science, 10*, 139–149.

Sexton, J. B., Helmreich, R. L., Neilands, T. B., Rowan, K., Vella, K., Boyden, J., … Thomas, E. J. (2006). The Safety Attitudes Questionnaire: Psychometric properties, benchmarking data, and emerging research. *BMC Health Services Research, 6*, 44–54.

Shojania, K. G., Duncan, B. W., McDonald, K. M., & Wachter, R. M. (2002). Safe but sound: Patient safety meets evidence-based medicine. *Journal of the American Medical Association, 288*(4), 508–513.

Singer, S. J., Lin, S., Falwell, A., Gaba, D. M., & Baker, L. (2009). Relationship of safety climate and safety performance in hospitals. *Health Services Research, 44*(2 Pt. 1), 399–421.

Singer, S. J., Meterko, M., Baker, L., Gaba, D., Falwell, A., & Rosen, A. (2007). Workforce perceptions of hospital safety culture: Development and validation of the patient safety climate in healthcare organizations survey. *Health Services Research, 42*(5), 1999–2021.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*, 245–251.

Stern, Z., Katz-Navon, T., & Naveh, E. (2011). The effect of safety management systems on continuous improvement of patient safety: The moderating role of safety climate and autonomy. *Quality Management Journal*, *18*(1), 54–67.

Steyrer, J., Latzke, M., Pils, K., Vetter, E., & Strunk, G. (2011). Development and validation of a patient safety culture questionnaire in acute geriatric units. *Gerontology*, *57*(6), 481–489.

Stolzer, A. J., Halford, C. D., & Goglia, J. J. (2011). *Implementing safety management systems in aviation*. Farnham, UK: Ashgate.

Thomas, E. J., Studdert, D. M., Burstin, H. R., Orav, E. J., Zeena, T., Williams, E. J., … Brennan, T. A. (2000). Incidence and types of adverse events and negligent care in Utah and Colorado. *Medical Care*, *38*(3), 261–271.

Valentin, A., Capuzzo, M., Guidet, B., Moreno, R. P., Dolanski, L., Bauer, P., … Sentinel Events Evaluation Study Investigators. (2006). Patient safety in intensive care units: Results from the multinational Sentinel Events Evaluation (SEE) study. *Intensive Care Medicine*, *32*(10), 1591–1598.

Valentin, A., Capuzzo, M., Guidet, B., Moreno, R. P., Metnitz, B., Bauer, P., … Sentinel Events Evaluation Study Investigators. (2009). Errors in administration of parenteral drugs in intensive care units: Multinational prospective study. *British Medical Journal*, *338*, b814.

van den Bos, J., Rustagi, K., Gray, T., Halford, M., Ziemkiewicz, E., & Shreve, J. (2011). The $17.1 billion problem: The annual cost of measurable medical errors. *Health Affairs*, *30*(4), 596–603.

Vogus, T. J., & Sutcliffe, K. M. (2007). The impact of safety organizing, trusted leadership, and care pathways on reported medication errors in hospital nursing units. *Medical Care*, *45*(10), 997–1002.

Waring, J. J. (2009). Constructing and re-constructing narratives of patient safety. *Social Science and Medicine*, *69*, 1722–1731.

Waterman, R. H. Jr., Peters, T. J., & Phillips, J. R. (1980). Structure is not organization. *Business Horizons*, *23*(3), 14–26.

Weingart, S. N., Farbstein, K., Davis, R. B., & Phillips, R. S. (2004). Using a multihospital survey to examine the safety culture. *Joint Commission Journal on Quality and Patient Safety*, *30*(3), 125–132.

Wilkinson, A., Redman, T., Snape, E., & Marchington, M. (1998). *Managing with total quality management*. Basingstoke, UK: Macmillan.

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*(4), 587–596.