OPEN

The Correlation Between Neonatal Intensive Care Unit Safety Culture and Quality of Care

Jochen Profit, MD, MPH,† Paul J. Sharek, MD,†‡§ Xin Cui, PhD, MPH,*† Courtney C. Nisbet, RN, MS,†§ Eric J. Thomas, MD, MPH,|| Daniel S. Tawfik, MD,¶ Henry C. Lee, MD, MS,*† David Draper, PhD,** and J. Bryan Sexton, PhD††‡‡‡*

Objectives: Key validated clinical metrics are being used individually and in aggregate (Baby-MONITOR) to monitor the performance of neonatal intensive care units (NICUs). The degree to which perceptions of key components of safety culture, safety climate, and teamwork are related to aspects of NICU quality of care is poorly understood. The objective of this study was to test whether NICU performance on key clinical metrics correlates with caregiver perceptions of safety culture.

Study Design: Cross-sectional study of 6253 very low-birth-weight infants in 44 NICUs. We measured clinical quality via the Baby-MONITOR and its nine risk-adjusted and standardized subcomponents (antenatal corticosteroids, hypothermia, pneumothorax, healthcare-associated infection, chronic lung disease, retinopathy screen, discharge on any human milk, growth velocity, and mortality). A voluntary sample of 2073 of 3294 eligible professional caregivers provided ratings of safety and teamwork climate using the Safety Attitudes Questionnaire. We examined NICU-level variation across clinical and safety culture ratings and conducted correlation analysis of these dimensions.

Results: We found significant variation in clinical and safety culture metrics across NICUs. Neonatal intensive care unit teamwork and safety climate ratings were correlated with absence of healthcare-associated infection ($r = 0.39$ [$P = 0.01$] and $r = 0.29$ [$P = 0.05$], respectively). None of the other clinical metrics, individual or composite, were significantly correlated with teamwork or safety climate.

Conclusions: Neonatal intensive care unit teamwork and safety climate were correlated with healthcare-associated infections but not with other quality metrics. Linkages to clinical measures of quality require additional research.

Key Words: Baby-MONITOR, composite measure, safety culture, teamwork, neonatal intensive care unit

Abbreviations: CPQCC = California Perinatal Quality Care Collaborative, HAI = healthcare-associated infection, NICU = neonatal intensive care unit, NNP = neonatal nurse practitioner, RN = registered nurse, RCP = respiratory care practitioner, SAQ = Safety Attitudes Questionnaire, VON = Vermont Oxford Network, VLBW = very low-birth-weight

(*J Patient Saf* 2020;16: e310–e316)

Establishing a culture of safety in health care has been a national health policy priority. The Joint Commission includes the establishment of a culture of safety as a critical component for achieving highly reliable and safe care[1] and requires hospitals to measure and monitor safety culture in an ongoing fashion.[2] A culture of safety has been defined as "individual and group values, attitudes, perceptions, competencies, and patterns of behavior that determine the commitment to, and the style and proficiency of, an organization's health and safety management."[3] Maintaining a

From the *Perinatal Epidemiology and Health Outcomes Research Unit, Division of Neonatology, Department of Pediatrics, Stanford University School of Medicine and Lucile Packard Children's Hospital; †California Perinatal Quality Care Collaborative; ‡Center for Quality and Clinical Effectiveness, Lucile Packard Children's Hospital; §Division of General Pediatrics, Department of Pediatrics, Stanford University, Palo Alto, California; ||University of Texas at Houston - Memorial Hermann Center for Healthcare Quality and Safety, University of Texas Medical School, Houston, Texas; ¶Division of Pediatric Critical Care Medicine, Department of Pediatrics, Stanford University School of Medicine; Stanford; **Department of Applied Mathematics and Statistics, Baskin School of Engineering University of California, Santa Cruz, California; and ††Department of Psychiatry, Duke University School of Medicine, and ‡‡Duke Patient Safety Center, Duke University Health System, Durham, North Carolina.

Correspondence: Jochen Profit, MD, MPH, Perinatal Epidemiology and Health Outcomes Research Unit, Division of Neonatology, Department of Pediatrics, Stanford University School of Medicine, MSOB Rm x115, 1265 Welch Rd, Stanford, CA 94305 (e-mail: profit@stanford.edu).

The authors disclose no conflict of interest.

This work was supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (R01 HD084679-01, co-PI: J.B.S. and J.P., and K24 HD053771-01, PI: E.J.T.).

J.P. had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

J.P. acquired funding for this study, conceptualized and designed the study, selected data for inclusion in analyses, analyzed the data, assisted with interpretation of the results, drafted the initial manuscript, and approved the final manuscript as submitted. P.J.S. helped conceptualize and design the study, was the local lead for CPQCC NICUs, helped select data for inclusion in the survey and analyses, assisted with interpretation of the results, revised the manuscript, and approved the final manuscript as submitted. X.C. assisted with designing the analysis and interpretation of the results, revised the manuscript, and approved the final manuscript as submitted. C.N. coordinated data collection among CPQCC member NICUs, de-identified data, assisted with interpretation of the results, revised the manuscript, and approved the final manuscript as submitted. E.J.T. helped acquire funding, conceptualized and designed the study, selected data for inclusion in the survey and analyses, assisted with interpretation of the results, revised the manuscript, and approved the final manuscript as submitted. D.S.T. helped with interpretation of the results, revision of the manuscript, and approved the final manuscript as submitted. H.C.L. helped with conceptualizing the paper, designing the analysis, and interpreting the results. He revised the manuscript and approved the final manuscript as submitted. D.D. created the statistical methodology underlying the Baby-MONITOR, helped with interpreting the results and revising the manuscript, and approved the final manuscript as submitted. J.B.S. helped acquire funding for this study, conceptualized and designed the study, selected data for inclusion in the survey and analyses, and assisted with interpretation of the results. He revised the initial manuscript and approved the final manuscript as submitted.

What's Known: Establishing a strong culture of safety is a health policy priority. Teamwork and safety climate, two well-established dimensions of safety culture, vary significantly among NICUs. The contribution of this variation to differences in NICU quality of care delivery is unknown.

What This Study Adds: NICU teamwork and safety climate correlated significantly with healthcare-associated infections. However, other metrics of quality of care did not correlate. Caution should be applied in equating efforts to improve safety culture with expectations for better quality of care delivery.

Supplemental digital contents are available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.journalpatientsafety.com).

culture of safety is a strategy for preventing patient harm. Safety culture surveys assess caregiver perceptions of unit norms. In other areas of healthcare, caregiver perceptions of safety culture, especially its best-studied subdomains, safety climate and teamwork, have been shown to vary widely and were linked to clinical outcomes, including healthcare-associated infections (HAIs) and mortality.[4–8]

The Safety Attitudes Questionnaire (SAQ) was found to be psychometrically robust[9] in the neonatal intensive care unit (NICU) and demonstrated significant variation in safety culture.[10] Preterm infants in the NICU setting are fragile and require complex and prolonged intensive care. This makes them vulnerable to lapses in teamwork and patient safety. However, the relationship between NICU safety culture ratings and clinical metrics of quality has not been established. In addition, the evidence from other areas of healthcare demonstrates relatively weak links between safety culture and clinical outcomes,[11] along with concerns about reporting bias.[12,13]

Neonatal intensive care is a complex and multidimensional activity, which the measurement of its quality should reflect. In previous work, we developed the Baby-MONITOR (Measure Of Neonatal InTensive care Outcomes Research), a composite indicator of quality of care provided to very low-birth-weight (VLBW, <1500 g) infants.[14] A panel of experts selected 9 of 28 potential metrics in a modified Delphi experiment.[15] This selection process was subsequently sanctioned by a sample of clinical neonatologists.[16] Both groups identified the same nine clinical metrics. These metrics routinely collected by NICUs that are members of the California Perinatal Quality Care Collaborative (CPQCC) and the Vermont Oxford Network (VON), which collects data on nearly 85% of the VLBW infants born in the United States. Each measure is risk-adjusted, standardized, equally weighted, and averaged. The Baby-MONITOR has face validity[15,16] and has been shown to be robust to variations in methods of construction.[14] Several of the measures of the Baby-MONITOR more narrowly represent safety domains, including infections, antenatal steroids, hypothermia on admission, pneumothorax, and retinopathy screening. Others could be defined in more quality domains requiring teamwork for high performance (chronic lung disease, growth velocity, and any human milk feedings at discharge). Overlap exists between safety and quality and several measures require both behaviors for high performance.

Based on the associations of quality of care delivery with health worker assessments of safety and teamwork climate in other areas of medicine, we hypothesized that caregiver assessments of teamwork and safety climate would correlate with clinical metrics of quality for VLBW infants in the NICU setting, using the Baby-MONITOR and its individual subcomponents.

METHODS

Design

This cross-sectional study combined registry data routinely submitted by NICU members of the CPQCC[17] with a voluntary sample of healthcare workers, participating in two simultaneous quality improvement initiatives, organized by the CPQCC focused on Delivery Room Management. More than 90% of NICUs in California are members of the CPQCC. We used the CPQCC clinical data to compute risk-adjusted scores for each subcomponent of the Baby-MONITOR. We then correlated these scores against health care worker assessments of teamwork and safety climate. Clinical data were obtained between January 1, 2010, and December 31, 2012, and safety culture survey data were collected between June and September 2011 from 44 participant NICUs. Thus, we selected the clinical data to cover approximately 1.5 years before and after the timing of the survey, providing a large clinical sample for estimation of quality of care delivery to VLBW infants in California NICUs.

The CPQCC assures high data quality through several mechanisms. It trains local NICU personnel to abstract clinical data. Annual training sessions help promote accuracy and uniformity in data abstraction. In addition, each record has range and logic checks both at the time of data collection and at data closeout, with auditing of records with excessive missing data. Definitions align with those specified for members of the VON.

Sample

Infants

Our goal for this study was to create a relatively homogenous and unbiased sample of VLBW infants for comparison across NICUs.[15] To ensure that patient outcomes reflected the quality of care of the NICU under observation, we excluded infants who died before 12 hours after delivery and those with severe congenital anomalies. We also restricted the analysis to infants born after 24 completed weeks of gestation to avoid systematic bias based on decisions to withhold resuscitation at the threshold of viability.[18] We used multiyear analyses because of the small number of VLBW infants cared for in some institutions.

Health Care Workers

A cross-sectional anonymous survey of all NICU healthcare worker perceptions of teamwork was performed among a voluntary sample of NICUs participating in a quality improvement collaborative organized by the CPQCC.[17] We offered to analyze and provide feedback on a survey of safety culture to all 61 NICUs who participated in the improvement initiative, 44 of which accepted. For all units, we used a paper-based version of the SAQ instrument to investigate safety and teamwork climate. Staff with a 0.5 full time equivalent or greater time commitment to the NICU for at least the four weeks before survey administration were invited to participate. Details of the survey and its administration can be found in the eAppendix, http://links.lww.com/JPS/A193.[19] Our response rate was 62.9% (2073/3294), with a range across the 44 hospitals of 21.7% to 100% (mean [SD] = 69.7% [19.8%]).

The study was approved by the institutional review board at Stanford University.

Metrics

Quality of Care

Details of the Baby-MONITOR have been published elsewhere.[14,15] In brief, an expert panel selected nine metrics of quality for inclusion in the composite, including the following[1]: antenatal corticosteroid use,[2] hypothermia (<36°C) during the first hour after delivery,[3] nonsurgically induced pneumothorax,[4] HAI,[5] chronic lung disease (oxygen requirement at 36 weeks' gestational age),[6] timely eye exam (retinopathy of prematurity screening as recommended by the American Academy of Pediatrics),[7] discharge on any human breast milk,[8] growth velocity,[9] and mortality during the birth hospitalization.[9]

Each of the metrics is scored so that a higher score indicates a better outcome. All metrics, except for timely eye exam as a process measure, are individually risk adjusted for severity of illness at the time of birth. To further classify NICU performance on each quality measure, we used a method developed by Draper and Gittoes.[20] For each NICU and for each subcomponent of the Baby-MONITOR, a z score was computed as the observed rate minus the expected rate divided by its estimated standard error.

TABLE 1. Description of Sample

Characteristics	Level	n	%
Survey respondents			
NICU level (n = 44)			
CCS level	Regional	10	23
	Community	28	64
	Intermediate	5	11
	Non-CCS	1	2
Respondent level (n = 2073)			
Female		1697	85
Typical shift	Days	894	48
	Evenings	79	4
	Nights	602	32
	Variable	293	16
Job position	Physician	204	10
	Fellow physician	31	2
	NNP	35	2
	RN	1464	72
	RCP	286	14
	Others	21	1
Working experience in specialty	<6 mo	20	1
	6–11 mo	27	1
	1–2 y	74	4
	3–4 y	192	10
	5–10 y	476	24
	11–20 y	538	27
	≥21 y	643	33
Clinical sample (n = 6253)			
Gestational age, wk	25–27	2134	34
	28–29	1682	27
	≥30	2437	39
Female		3091	49
Prenatal care		6057	97
Multiple gestation		1713	27
Cesarean delivery		4695	75
SGA		1144	18
Maternal age, y	≤19	578	9
	20–29	2640	42
	30–39	2588	41
	≥40	447	7
Apgar 5 min	≤3	299	5
	4–6	1047	17
	≥7	4878	78
Outborn		761	12
Baby-MONITOR metrics of quality of care			
No HAI		5463	91
No chronic lung disease		4302	79
Any human breast milk at discharge		3728	66
High growth velocity		2654	55
Survival		5789	94
Antenatal steroid use		4524	90
No hypothermia on admission		5598	90
No pneumothorax		6033	97
Timely retinopathy exam		3883	96

CCS, California Children's Services; Non-CCS: NICUs not certified by the California Children's Services Program; NNP, neonatal nurse practitioner; RCP, respiratory care practitioner; SGA, small for gestational age.

These standardized *z* scores are approximately normally distributed with the mean of 0 and standard deviation of 1 when no quality differences are present.

Safety Culture

Of the several safety culture survey instruments in the literature, the SAQ is widely used and has good psychometric properties.[21] The SAQ contains 30 items that load on the following six domains: teamwork climate, safety climate, job satisfaction, perceptions of management, stress recognition, and working conditions. Each item is rated on a five-point Likert scale ranging from 1 (disagree strongly) to 5 (agree strongly). Positions included attending physicians, fellow physicians, neonatal nurse practitioners, registered nurses (RNs), respiratory care practitioners, and others.

Safety culture scale scores were calculated at the NICU level as follows: first, we created for each scale item a binary variable that was 1 if respondents "strongly" or "slightly" agreed with the item

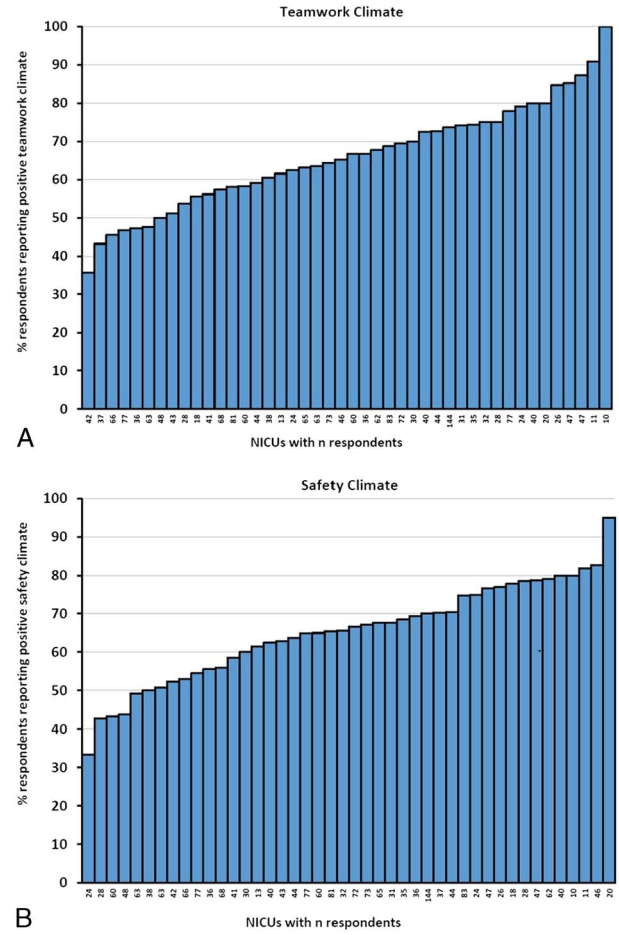

FIGURE 1. A, Percent of positive responses for the teamwork climate in 44 NICUs. Percent of positive teamwork climate is the percent of respondents responding in the "agree slightly" to "agree strongly" range across the positively worded items. B, Percent of positive responses for the safety climate in 44 NICUs. Percent of positive safety climate is the percent of respondents responding in the "agree slightly" to "agree strongly" range across the positively worded items. Numbers in x-axis, number of respondents in each NICU.

and 0 otherwise, and then, we computed the means of these dichotomous variables.[9,21] We call this "percentage agree" or "percentage reporting good 'safety climate' or 'teamwork climate.'"

Analyses

We used summary statistics such as frequencies, percentages, means (standard deviations), and graphs to describe demographics and our three variables: Baby-MONITOR (with subcomponents), safety climate, and teamwork climate. For each NICU, we computed *z* scores and percentiles for the Baby-MONITOR and its subcomponents, as well as percent positive rates at the scale and item level for the safety climate and teamwork scales. We then used Pearson correlation coefficient to test for correlations between the clinical and safety culture metrics.

Statistical analyses were performed using SAS (Version 9.4; SAS Institute, Inc, Cary, NC).

RESULTS

Table 1 shows the characteristics of survey respondents and the clinical sample. A total of 6253 VLBW infants in 44 NICUs met the inclusion criteria. Of these NICUs, 10 (22.7%) were designated as regional NICUs, 28 (63.6%) as community NICUs, and 5 (11.4%) as intermediate NICUs as defined by the California Department of Healthcare Services. These designations are roughly equivalent with designations by the American Academy of Pediatrics as level 4, 3, and 2, respectively.[22] Most respondents were RNs and female. The distribution of job positions among respondents mirrored the distribution of eligible participants across participating NICUs. Providers were quite experienced, with the largest number of respondents (643 [33%]) having worked more than 20 years in their specialty. Clinical characteristics are slightly better than all California estimates.[14]

Figures 1A and B show the percent of positive responses for the teamwork and safety climate scales, respectively. The mean (SD, range) percent positive response was 77.6% (6.2, 64.9%–89.6%) for teamwork and 77.0% (5.5, 66.2%–86.6%) for safety climate. Variation between NICUs was significant for both scales (analysis of variance; $P < 0.001$). Table 2 shows the distribution of responses across all respondents for the teamwork and safety climate scales and items. For all items, the response scores of the top 4 NICUs were significantly different from the bottom 4 NICUs.

The NICU Baby-MONITOR scores ranged from −2.5 to 1.7 standard units, indicating significant and clinically meaningful variation (eFigure 1 in the eAppendix shows NICU-level variation, http://links.lww.com/JPS/A193).

eTable 1 in the eAppendix, http://links.lww.com/JPS/A193, exhibits results at the NICU level, including the percent positive responses for the teamwork and safety climate scales and the observed minus expected scores in standard deviation units and percentiles for the Baby-MONITOR and each of its subcomponents. We found significant variation in performance across the composite and its subcomponents. The widest variation of Baby-MONITOR scores and its subcomponents between the top and bottom performing NICU was found in growth velocity, with observed minus expected performance ranging from −7.3 to 10.7 standard units. By definition, a difference of 1.96 standard units implies statistical significance; these variations are large in clinical terms.

Table 3 shows the Pearson correlation coefficients between teamwork and safety climate with the Baby-MONITOR and its

TABLE 2. Teamwork and Safety Climate Item Response Distributions*

	Overall	Bottom 4 NICUs	Top 4 NICUs	Difference Between the Top and Bottom 4 NICUs
	Mean (SD)	Mean (SD)	Mean (SD)	
Teamwork climate				
It is easy for personnel here to ask questions when there is something they do not understand.	83.79 (5.35)	73.43 (1.15)	91.67 (0.72)	‡
I have the support I need from others in this NICU to care for patients.	85.73 (4.88)	77.91 (1.73)	95.00 (2.22)	‡
Nurse input is well received in this NICU.	77.24 (8.35)	61.93 (3.89)	91.11 (4.27)	‡
In this NICU, it is difficult to speak up if I perceive a problem with patient care.	28.07 (7.45)	14.28 (3.64)	40.48 (0.63)	‡
Disagreements in this NICU are appropriately received.	69.07 (8.51)	55.15 (1.14)	85.06 (2.22)	‡
The physicians and nurses here work together as a well-coordinated team.	77.51 (10.48)	55.85 (2.16)	92.84 (1.79)	‡
Safety climate				
The culture in this NICU makes it easy to learn from the errors of others.	67.61 (7.80)	55.14 (1.84)	80.65 (1.93)	‡
Medical errors are handled appropriately in this NICU.	83.71 (6.97)	71.16 (2.66)	95.29 (3.30)	†
I know the proper channels to direct questions regarding patient safety in this NICU.	87.73 (4.15)	80.93 (1.11)	94.54 (1.01)	‡
I am encouraged by others in this NICU, to report any patient safety concerns I may have.	79.96 (5.68)	70.03 (0.51)	89.15 (1.40)	‡
I receive appropriate feedback about my performance.	74.32 (7.12)	62.51 (2.23)	86.30 (2.61)	‡
I would feel safe being treated here as a patient.	80.94 (8.20)	67.50 (0.58)	96.32 (1.68)	§
In this NICU, it is difficult to discuss errors.	35.04 (8.75)	16.32 (7.11)	49.12 (3.78)	†

*Scale score for a respondent = ((mean of the item from NICUs − 1) × 25).

†$P < 0.01$.

‡$P < 0.001$.

§$P < 0.0001$.

TABLE 3. Correlations Between Teamwork and Safety Climate and Baby-MONITOR and Its Subcomponents

	Teamwork Climate		Safety Climate	
	r (95% CI)	*P*	*r* (95% CI)	*P*
No HAI	**0.39 (0.10 to 0.61)**	**0.01**	0.29 (−0.01 to 0.55)	0.05
No chronic lung disease	0.17 (−0.13 to 0.45)	0.27	0.18 (−0.13 to 0.45)	0.25
Any human breast milk at discharge	0.02 (−0.28 to 0.32)	0.89	0.09 (−0.22 to 0.38)	0.57
High growth velocity	−0.05 (−0.35 to 0.25)	0.74	−0.20 (−0.47 to 0.10)	0.19
Survival	0.15 (−0.15 to 0.43)	0.33	0.24 (−0.06 to 0.50)	0.12
Antenatal steroid use	−0.05 (−0.36 to 0.27)	0.78	0.01 (−0.30 to 0.33)	0.93
No hypothermia on admission	−0.06 (−0.35 to 0.25)	0.72	−0.16 (−0.44 to 0.14)	0.30
No pneumothorax	0.25 (−0.05 to 0.51)	0.11	0.07 (−0.23 to 0.36)	0.63
Timely retinopathy exam	−0.07 (−0.37 to 0.23)	0.64	−0.16 (−0.44 to 0.15)	0.32
Baby-MONITOR score	0.14 (−0.17 to 0.42)	0.38	0.02 (−0.28 to 0.31)	0.92

n = 44 NICUs.

95% CI, 95% confidence interval; *r*, Pearson correlation coefficient.

subcomponents. Only HAI exhibited a statistically significant relationship (Figs. 2A, B). None of the other components of the Baby-MONITOR score or the composite were significantly correlated with teamwork and safety climate. eTables 2A and B in the eAppendix, http://links.lww.com/JPS/A193, show the correlations between teamwork climate items, safety climate items, and Baby-MONITOR and its subcomponents. Both positive and negative correlations were observed with no predominant pattern.

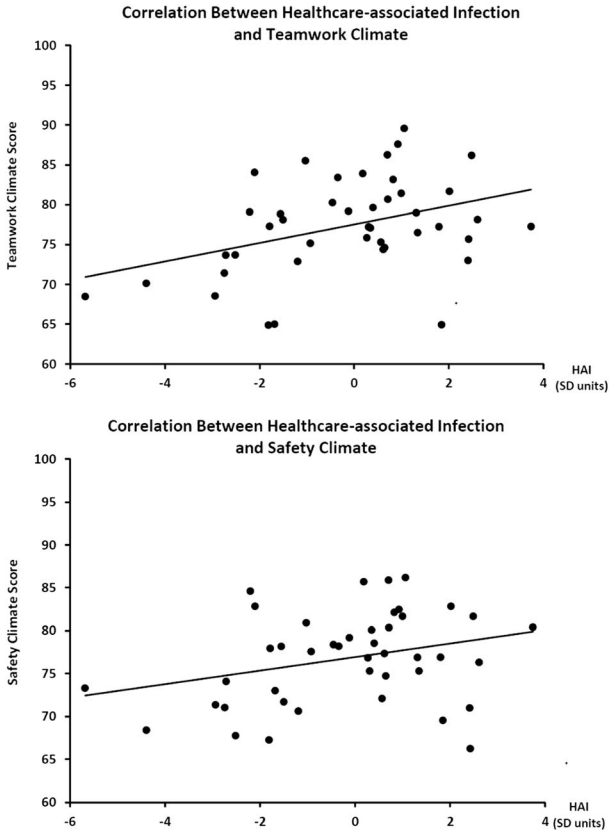

FIGURE 2. A, Correlation between HAI and teamwork climate. Each dot represents one NICU. B, Correlation between HAI and safety climate. Each dot represents one NICU. x-axis, HAI in SD units; y-axis, safety climate score.

DISCUSSION

This study extends findings from the healthcare literature demonstrating substantial variability in safety culture and clinical metrics of quality and outcomes. Our findings also reinforce results from studies in adult intensive care settings,[23] and our own previous work, which revealed links between safety or teamwork climate and infection-related outcomes. Healthcare workers in NICUs who report that "I would feel safe being treated here" work in units with lower infection rates. The reverse of this association is also true, whereby NICUs with high infection rates have fewer healthcare workers who report "It is easy to ask questions when there is something they don't understand" or "Physicians and nurses work together as a well-coordinated team." However, we were surprised that none of the other clinical outcomes were significantly correlated with safety or teamwork climate. Overall, our findings reflect a weaker than expected correlation of metrics of quality with teamwork and safety climate.

A priori, we did expect some metrics to yield lower correlations, particularly because some of them may be outside the direct purview of many frontline providers in the NICU (antenatal corticosteroid therapy, pneumothorax, hypothermia on admission, and timely eye exam). For example, obstetricians provide antenatal steroids, specialized delivery room teams may be accountable for pneumothoraces (at least those that occur in the delivery room) or hypothermia on admission, and the ophthalmology team may have a separate system for tracking infants in need for a retinopathy exam, with variable input by frontline NICU staff. Therefore, when frontline staff responded to the survey, the processes associated with these outcomes may not have been at the forefront of their minds. Nevertheless, we wanted to test for correlations with these metrics, given that they had previously been identified as key metrics of NICU quality; knowing how these relate to safety and teamwork metrics is important information for leadership.

We did expect other subcomponents of the Baby-MONITOR to correlate with teamwork and safety climate. Specifically, based on the adult literature and clinical rationale, we expected NICU-level safety culture ratings to be associated with HAI, mortality, chronic lung disease, any human breast milk at discharge, and growth velocity. Each of these metrics requires multidisciplinary

care teams to work together effectively, with flat hierarchies, good communication, and the freedom to speak up if an unsafe situation is identified.[24,25] However, in this sample, the HAI rate was the only metric associated with safety and teamwork climate. It may be that the perceptions of unit teamwork and safety climate are associated with concrete behaviors that help prevent infections. For example, in one study, units with higher teamwork and safety climate ratings exhibited better hand hygiene.[8]

Potentially, the care processes linked with other outcomes are less concretely defined than with infection and not as readily recognizable to survey respondents. For example, growth velocity requires a more prolonged and complex set of interventions and has no sentinel outcome (e.g., infection). Although teamwork is critical in ensuring optimal nutritional support, providers may not as readily associate distinct behaviors with this outcome. We speculate that connecting actions and outcomes in the minds of providers may potentially be a critical intervention for improvement and one that could be monitored in its success through repeated assessments of safety culture and clinical outcomes. When these connections are not made, efforts to improve teamwork and safety behaviors may not directly result in better quality of care delivery.

The previous safety culture literature may also be subject to publication bias.[12,13] Few randomized controlled studies exist to demonstrate causal relationships between safety culture and clinical outcomes. Thus, our findings are important and have practical implications. They add to a growing literature highlighting[11] the fact that providers and managers need to be careful not to confuse efforts to improve safety culture with expectations for broad-based quality improvement. It is important to recognize that the evidence for links between safety culture and clinical outcomes is still being developed and that many things may influence clinical outcomes beyond what a safety culture survey can measure. Other studies have shown that interventions to improve teamwork (e.g., TeamSTEPPS) may improve teamwork behaviors without necessarily improving clinical outcomes.[25] Although such activities may be necessary to create a favorable contextual environment for effective implementation of standardized evidence-based care delivery, they are not sufficient. Neonatal intensive care units still have to do the hard work of establishing care delivery mechanisms that optimize care outcomes. It they fail to do this, high safety culture ratings may merely reflect nice people providing suboptimal care on a range of outcomes.

Our results must be viewed within the context of the study design. Our cross-sectional study design is hypothesis generating. In addition, because all the analyses are conducted at the NICU level, our sample of 44 NICUs is relatively small to detect statistical significance, making the size and direction of the correlation coefficients more informative in this context. It is also important to understand that the strength of the correlations is not unusual with regard to institutional level variables. In a previous paper, we found just slightly higher correlations when we correlated clinical outcomes with one another.[26] Here, we correlated dimensions (teamwork and safety climate), which are more distant in their relation to clinical outcomes. The correlations overall likely indicate that the Baby-MONITOR and dimensions of safety culture measure different aspects of quality of care delivery. By tracking both, institutions may gain insights about different components of service delivery that promote high-quality care and operational excellence.

Neonatal intensive care units participating in this study were not randomly chosen. Rather, they participated in the collaborative quality improvement effort for specific reasons. This may have introduced systematic bias into our analysis, the direction of which is not easily ascertained. Future studies will need to confirm our findings in larger samples and different healthcare settings. Without our knowledge, NICUs may have been engaged in a variety of quality and safety efforts that may have influenced respondent perceptions. In addition, the culture of safety survey information was gathered over the short timeframe of June 2011 through September 2011, which might not accurately reflect the safety culture scores throughout the three years used to evaluate the clinical outcomes (January 2010 to December 2012), potentially biasing our results toward the null. Any self-report survey may be subject to reporting bias, however, our relatively large sample size and response rate compare favorably with similar studies of safety culture assessments in the literature. In addition, we used some negatively valenced items, such as "In this NICU, it is difficult to speak up if I perceive a problem with patient care," checked the psychometrics for this sample, and reported substantial variability between the NICUs.[19] Finally, compared with other safety culture tools, the SAQ, and especially the safety and teamwork climate scales, perform favorably in terms of psychometrics, clinical applicability, and responsiveness to interventions.

CONCLUSIONS

This study reveals significant correlation between HAI rates and NICU teamwork and safety climate. However, other metrics of quality predicted to correlate with teamwork and safety climate did not. Caution is needed in equating efforts to improve safety culture with expectations for broad-based quality improvement.

ACKNOWLEDGMENTS

The authors thank the providers from the following NICUs who participated by sharing their assessments: Alta Bates Hospital, Berkeley; Antelope Valley Hospital, Lancaster; Arrowhead Regional Medical Center, Colton; California Pacific Medical Center-California, San Francisco; Cedars Sinai Medical Center, Los Angeles; Children's Hospital Central California, Madera; Children's Hospital of Orange County, Orange; Community Regional Medical Center - Fresno, Fresno; Doctors Medical Center-Modesto, Modesto; El Camino Hospital-Mountain View, Mountain View; Garfield Medical Center, Monterey Park; Good Samaritan Hospital of Santa Clara Valley–San Jose, San Jose; Harbor UCLA Medical Center, Torrance; Hoag Memorial Presbyterian Hospital, Newport Beach; Hollywood Presbyterian, Los Angeles; Huntington Memorial Hospital, Pasadena; John Muir Medical Center, Walnut Creek; Kaweah Delta District Hospital, Visalia; Mercy Medical Center Redding, Redding; Mercy San Juan Hospital, Carmichael; Methodist Hospital of Southern California, Arcadia; Miller Children's at Long Beach Memorial Medical Center, Long Beach; Orange County Global Medical Center, Santa Ana; Pomona Valley Regional Medical Center, Pomona; Providence Saint Joseph's Medical Center, Burbank; Providence Tarzana Medical Center, Tarzana; Rady Children's Hospital at Palomar, Escondido; Redlands Community Hospital, Redlands; Regional Medical Center of San Jose, San Jose; Riverside County Regional Medical Center, Moreno Valley; Salinas Valley Memorial Healthcare System, Salinas; Scripps Memorial Hospital-La Jolla, La Jolla; Sequoia District Hospital, Redwood City; Sharp Chula Vista Medical Center, Chula Vista; Sharp Mary Birch Hospital for Women, San Diego; St. Bernadine Medical Center, San Bernardino; St. Jude Medical Center, Fullerton; Sutter Medical Center-Sacramento, Sacramento; Tri-City Medical Center, Oceanside; UC Davis Medical Center, Sacramento; UC San Francisco Medical Center, San Francisco; Valley Presbyterian Hospital, Van Nuys. The authors also thank the local leaders and staff of the CPQCC NICUs who contributed their efforts to this study.

REFERENCES

1. Chassin MR, Loeb JM. High-reliability health care: getting there from here. *Milbank Q*. 2013;91:459–490.
2. The Joint C. *2009 Comprehensive Accreditation Manual for Hospitals (CAMH): The Official Handbook*. Chicago: Joint Commission Resources; 2008.
3. Study Group on Human Factors HaSCoGB. *Organising for Safety: Third Report of the ACSNI (Advisory Committee on the Safety of Nuclear Installations). Sudbury*. England: HSE Books; 1993.
4. Thomas EJ, Sexton JB, Helmreich RL. Discrepant attitudes about teamwork among critical care nurses and physicians. *Crit Care Med*. 2003; 31:956–959.
5. Sexton JB, Holzmueller CG, Pronovost PJ, et al. Variation in caregiver perceptions of teamwork climate in labor and delivery units. *J Perinatol*. 2006;26:463–470.
6. Kho ME, Carbone JM, Lucas J, et al. Safety Climate Survey: reliability of results from a multicenter ICU survey. *Qual Saf Health Care*. 2005;14: 273–278.
7. Modak I, Sexton JB, Lux TR, et al. Measuring safety culture in the ambulatory setting: the safety attitudes questionnaire—ambulatory version. *J Gen Intern Med*. 2007;22:1–5.
8. Daugherty EL, Paine LA, Maragakis LL, et al. Safety culture and hand hygiene: linking attitudes to behavior. *Infect Control Hosp Epidemiol*. 2012;33:1280–1282.
9. Profit J, Etchegaray J, Petersen LA, et al. The Safety Attitudes Questionnaire as a tool for benchmarking safety culture in the NICU. *Arch Dis Child Fetal Neonatal Ed*. 2012;97:F127–F132.
10. Profit J, Etchegaray J, Petersen LA, et al. Neonatal intensive care unit safety culture varies widely. *Arch Dis Child Fetal Neonatal Ed*. 2012;97: F120–F126.
11. DiCuccio MH. The relationship between patient safety culture and patient outcomes: a systematic review. *J Patient Saf*. 2015;11: 135–142.
12. Weaver SJ, Lubomksi LH, Wilson RF, et al. Promoting a culture of safety as a patient safety strategy: a systematic review. *Ann Intern Med*. 2013; 158(5 Pt 2):369–374.
13. Ioannidis JP, Munafò MR, Fusar-Poli P, et al. Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention. *Trends Cogn Sci*. 2014;18:235–241.
14. Profit J, Kowalkowski MA, Zupancic JA, et al. Baby-MONITOR: a composite indicator of NICU quality. *Pediatrics*. 2014;134:74–82.
15. Profit J, Gould JB, Zupancic JA, et al. Formal selection of measures for a composite index of NICU quality of care: Baby-MONITOR. *J Perinatol*. 2011;31:702–710.
16. Kowalkowski M, Gould JB, Bose C, et al. Do practicing clinicians agree with expert ratings of neonatal intensive care unit quality measures? *J Perinatol*. 2012;32:247–252.
17. Gould JB. The role of regional collaboratives: the California Perinatal Quality Care Collaborative model. *Clin Perinatol*. 2010;37:71–86.
18. Peerzada JM, Richardson DK, Burns JP. Delivery room decision-making at the threshold of viability. *J Pediatr*. 2004;145:492–498.
19. Profit J, Sharek PJ, Amspoker AB, et al. Burnout in the NICU setting and its relation to safety culture. *BMJ Qual Saf*. 2014;23:806–813.
20. Draper D, Gittoes M. Statistical analysis of performance indicators in UK higher education. *J Roy Statist Soc Ser A*. 2004;167:449–474.
21. Sexton JB, Helmreich RL, Neilands TB, et al. The Safety Attitudes Questionnaire: psychometric properties, benchmarking data, and emerging research. *BMC Health Serv Res*. 2006;6:44.
22. Levels of neonatal care. *Pediatrics*. 2012;130:587–597.
23. Sexton JB, Berenholtz SM, Goeschel CA, et al. Assessing and improving safety climate in a large cohort of intensive care units. *Crit Care Med*. 2011; 39:934–939.
24. Dietz AS, Pronovost PJ, Mendez-Tellez PA, et al. A systematic review of teamwork in the intensive care unit: what do we know about teamwork, team tasks, and improvement strategies? *J Crit Care*. 2014;29:908–914.
25. Weaver SJ, Lyons R, DiazGranados D, et al. The anatomy of health care team training and the state of practice: a critical review. *Acad Med*. 2010; 85:1746–1760.
26. Profit J, Zupancic JA, Gould JB, et al. Correlation of neonatal intensive care unit performance across multiple measures of quality of care. *JAMA Pediatr*. 2013;167:47–54.