

Teamwork in the Operating Room

Frontline Perspectives among Hospitals and Operating Room Personnel

J. Bryan Sexton, Ph.D.,* Martin A. Makary, M.D., M.P.H.,† Anthony R. Tersigni, Ed.D.,‡ David Pryor, M.D.,§ Ann Hendrich, M.S., F.A.A.N.,|| Eric J. Thomas, M.D., M.P.H.,# Christine G. Holzmüller, B.L.A.,** Andrew P. Knight, M.A.,†† Yun Wu, M.A.S.,‡‡ Peter J. Pronovost, M.D., Ph.D.§§

Background: The Joint Commission on Accreditation of Healthcare Organizations is proposing that hospitals measure culture beginning in 2007. However, a reliable and widely used measurement tool for the operating room (OR) setting does not currently exist.

Methods: OR personnel in 60 US hospitals were surveyed using the Safety Attitudes Questionnaire. The teamwork climate domain of the survey uses six items about difficulty speaking up, conflict resolution, physician–nurse collaboration, feeling supported by others, asking questions, and heeding nurse input. To justify grouping individual-level responses to a single score at each hospital OR level, the authors used a multilevel confirmatory factor analysis, intraclass correlations, within-group interrater reliability, and Cronbach's α . To detect differences at the hospital OR level and by caregiver type, the authors used multivariate analysis of variance (items) and analysis of variance (scale).

Results: The response rate was 77.1%. There was robust evidence for grouping individual-level respondents to the hospital OR level using the diverse set of statistical tests, e.g., Comparative Fit Index = 0.99, root mean squared error of approximation = 0.05, and acceptable intraclass correlations, within-group interrater reliability values, and Cronbach's α = 0.79. Teamwork climate differed significantly by hospital ($F_{59, 1,911} = 4.06$, $P < 0.001$) and OR caregiver type ($F_{4, 1,911} = 9.96$, $P < 0.001$).

Conclusions: Rigorous assessment of teamwork climate is possible using this psychometrically sound teamwork climate scale. This tool and initial benchmarks allow others to compare their teamwork climate to national means, in an effort to focus more on what excellent surgical teams do well.

TEAMWORK in the operating room (OR) is an important component of OR efficiency, quality of care, and patient safety. One principle in the 1999 Institute of Medicine report was to “promote effective team functioning” to create safe hospital systems.¹ Errors in the OR can have catastrophic consequences for patients, families, caregivers, and entire institutions. Breakdowns in communication and collaboration among OR team members have resulted in retained sponges, mismatched blood transfusions, and wrong-extremity nerve blocks.²⁻⁴ In addition, the Joint Commission on Accreditation of Healthcare Organizations recently reported breakdowns in communication as a leading root cause of wrong-site surgeries, operative and postoperative events, sentinel events, delays in treatment, and medication errors.^{|||}⁵

As such, anesthesiology and surgical departments now find themselves under increasing pressure to prevent these negative outcomes. The result has been a growing pool of team-training interventions and consultants that may improve the quality and efficiency of our clinical work through better teamwork. However, these training mechanisms are generally devoid of any reliable metrics to demonstrate improvements in teamwork.

The aviation industry has demonstrated important ties between teamwork and performance.⁶ This teamwork–safety association surfaced after investigating commercial aviation accidents and exposing cockpit crew member's reluctance to speak up and question the actions of a captain. From this came surveys to assess frontline pilots' perceptions of cockpit climate, which were also used to predict pilot performance.^{6,7} We applied this aviation model in medicine to assess teamwork climate and found similar intimidation and reluctance to voice concerns in the OR.⁸ Fortunately, there is emerging evidence demonstrating that teamwork climate in the OR is responsive to interventions, specifically preoperative briefings. Using a model adapted from commercial aviation, OR teams have conducted a brief discussion just before skin incision at the time of a surgical “time-out” to briefly review names and roles of all team members, the operative plan, familiarity with the procedure, and potential issues for the case. Two separate studies

This article is accompanied by an Editorial View. Please see: Orkin FK, Weeks WB: No patient left behind. ANESTHESIOLOGY 2006; 105:869-70.

* Assistant Professor, The Johns Hopkins University, Department of Anesthesiology and Critical Care Medicine, and Johns Hopkins Bloomberg School of Public Health, Department of Health Policy and Management. † Assistant Professor, The Johns Hopkins University, Department of Surgery, and Johns Hopkins Bloomberg School of Public Health, Department of Health Policy and Management. ‡ President and CEO, Ascension Health System, § Senior Vice President, Clinical Excellence, || Vice President, Clinical Excellence Operations, Ascension Health System, St. Louis, Missouri. # Associate Professor, The University of Texas, Houston Medical School, Department of Medicine and Center of Excellence for Patient Safety Research and Practice, Houston, Texas. ** Senior Research Coordinator, ‡‡ Biostatistician, The Johns Hopkins University, Department of Anesthesiology and Critical Care Medicine. †† Doctoral Candidate, Management Department, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania. §§ Professor, The Johns Hopkins University, Departments of Anesthesiology and Critical Care Medicine and Surgery, and Johns Hopkins Bloomberg School of Public Health, Department of Health Policy and Management.

Received from The Johns Hopkins University, School of Medicine and the Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland. Submitted for publication May 26, 2005. Accepted for publication June 6, 2006. Supported by grant Nos. UCHS014246 and IPOIHS1154401 from the Agency for Healthcare Research and Quality, Rockville, Maryland.

Address correspondence to Dr. Makary: 4940 Eastern Avenue, Building A-5, Baltimore, Maryland 21224. mmakary1@jhmi.edu. Individual article reprints may be purchased through the Journal Web site, www.anesthesiology.org.

||| Joint Commission on Accreditation of Healthcare Organizations: Sentinel event statistics—December 31, 2005. Available at: www.jointcommission.org/SentinelEvents/Statistics. Accessed May 5, 2006.

performed preoperative briefings and assessed caregiver collaboration using teamwork climate items before and after briefings. Each study found significant improvement in the post-teamwork climate.^{9,10}

To date, there are no psychometrically sound tools to measure culture that are widely used in the OR setting. Nevertheless, teamwork climate is a recognized marker of performance,^{5,11} and hospitals are encouraged to start measuring culture in the year before the new Joint Commission on Accreditation of Healthcare Organizations requirement. The need for empirically sound tools grows stronger as the Joint Commission on Accreditation of Healthcare Organizations requirement date approaches.

The primary aims of this study were (1) to test the psychometric properties (*i.e.*, reliability and validity) of a teamwork climate scale in the OR setting and (2) to provide some initial benchmarking information on OR teamwork climate by hospital and caregiver type. Secondary objectives of this study were to examine differences in perceptions of OR teamwork by hospital and caregiver type.

Materials and Methods

Our survey, the Safety Attitudes Questionnaire (SAQ),^{12,13} is a refinement of the Intensive Care Unit Management Attitudes Questionnaire.^{8,14} The latter was adapted from the Flight Management Attitudes Questionnaire¹⁵ and its predecessor, the Cockpit Management Attitudes Questionnaire.¹⁶ These surveys are reliable and sensitive to change,¹⁷ and the elicited attitudes are shown to predict performance.¹⁸⁻²⁰ There is a 25% overlap in item content between the SAQ and the Flight Management Attitudes Questionnaire. We improved content validity and created an OR version of the SAQ after reviewing the literature on teamwork in the OR, conducting focus groups and asking OR caregivers to review the survey for content relevance.

The six domains of the SAQ are teamwork climate, safety climate, job satisfaction, perceptions of management, stress recognition, and working conditions. These domains are based on previous research in aviation and medicine.^{8,14,21,22} Relative to other culture survey instruments being used in health care, the SAQ seems to be the leading tool given its content and characteristics of validity and reliability.²³ Results from the teamwork climate domain are reported in this article. Teamwork climate reflects how caregivers from the same work unit perceive quality of collaboration between personnel. Six of the 30 SAQ domain items define teamwork climate.

Content validity of the OR teamwork climate items was established through behavioral observations of OR teamwork during surgical procedures,²⁴ which highlighted specific behaviors contributing to observer rat-

ings of overall OR teamwork and drew attention to frequent communication breakdowns between caregivers. Teamwork behaviors observed included assertion (speaking up when there is a problem), conflict resolution (resolving disagreements), inquiry (asking questions), and interdisciplinary coordination (*e.g.*, collaboration between physicians and nurses). These themes also surfaced in focus groups of caregivers²⁵ and during behavioral observations of neonatal resuscitations.²⁶ These same collaboration themes are well represented in the OR teamwork climate items: Nurse input about patient care is well received in the OR; in the ORs here, it is difficult to speak up if I perceive a problem with patient care; disagreements in the ORs here are resolved appropriately (*i.e.*, not *who* is right but *what* is best for the patient); I have the support I need from other personnel to care for patients; it is easy for personnel in the ORs here to ask questions when there is something that they do not understand; and the physicians and nurses here work together as a well-coordinated team. Response options for each item ranged from 1 (disagree strongly) to 5 (agree strongly).

The SAQ (Operating Room Version) was administered to all OR caregivers in a Catholic health system comprised of 60 hospitals in 16 states in July and August of 2004. No one was excluded, and OR caregivers included surgeons, anesthesiologists, surgical technicians, certified registered nurse anesthetists, and OR nurses. For clarification of terminology, these five OR caregiver types were aggregated within each hospital and referred to as the "hospital OR level," whereas each type of caregiver was aggregated among all hospitals as the "caregiver level." Random sampling was not performed because of small sample sizes in caregiver types within a hospital that would threaten the representativeness of the data. Instead, we surveyed all hospitals and sought as high a response rate as possible within each caregiver type. Surveys were administered during preexisting departmental and staff meetings, together with a pencil and return sealable envelope to maintain confidentiality. Individuals not captured in preexisting meetings received a hand-delivered survey, pencil, and return envelope. No personal identifying information was tracked beyond caregiver type and hospital.

Statistical Analysis

Teamwork climate scale scores were calculated by taking the average of the six items (one item was reverse scored because of negative wording). Internal reliability was assessed using Cronbach's α .²⁷ A multilevel confirmatory factor analysis was performed on the six-item scale to test how well the items hung together as a construct.²⁸ Multilevel confirmatory factor analysis accounts for caregivers nested within hospital ORs and corrects the between group covariance matrix to prevent between-group bias.^{28,29}

Table 1. Response Rates and Respondent Demographics by Caregiver Type

| Caregiver Type | Response Rate (Returned/Admin) | Age, Mean (SD), yr | % Female (n) | Years of Experience in Position, Mean (SD) | Years Working at Current Hospital, Mean \pm SD |
|---------------------|--------------------------------|--------------------|---------------|--|--|
| Surgeon | 73% (222/305) | 48.3 (9.92) | 8.6% (19) | 17.4 (9.41) | 12.3 (9.20) |
| Surgical technician | 78% (564/728) | 37.8 (11.58) | 73.7% (417) | 11.2 (11.17) | 7.9 (9.45) |
| Anesthesiologist | 77% (170/220) | 45.8 (9.31) | 12.7% (21) | 15.8 (8.18) | 10.6 (8.60) |
| CRNA | 67% (121/181) | 44.6 (10.71) | 50.0% (63) | 14.7 (12.32) | 9.5 (9.35) |
| OR nurse | 79% (1,058/1,335) | 43.3 (10.85) | 89.0% (942) | 13.9 (10.04) | 10.7 (8.69) |
| Total | 77% (2,135/2,769) | 42.6 (11.3) | 68.5% (1,462) | 13.7 (10.47) | 10.0 (9.08) |

Characteristics of respondents and response rates by operating room (OR) caregiver type.

CRNA = certified registered nurse anesthetist.

To accurately assess culture or climate constructs at the hospital OR level, individual perceptions must show high agreement within units (hospital ORs) and high variance between units.³⁰ We calculated intraclass correlation coefficients (ICCs) and the within-group interrater reliability (r_{wg}) statistic to justify aggregation of caregivers within their hospital ORs. The ICC(1) is a measure of between group variability.³¹ ICC(2) indicates the reliability of the group means (note that the ICC(2) relates respondents to each other and is distinct from the scale internal reliability, which relates the six items to each other). To calculate ICC(1) and ICC(2), a one-way analysis of variance was performed on individual-level responses, with intensive care unit as the independent variable. The within-group interrater reliability (r_{wg}) statistic is a measure of consensus that should meet or exceed 0.70 to provide an accurate index of within-group agreement.³² This test of consensus is important to determine whether OR caregivers within a hospital sufficiently share perceptions of teamwork climate, which would justify aggregation of individual responses to the hospital OR level.

To test for differences in perceptions of teamwork, we focused on teamwork climate by the caregiver level and the hospital OR level. Using multivariate analysis of variance, we tested for differences between OR caregivers and differences between hospitals with respect to each teamwork climate item. We then used analysis of variance to test the same groups for differences on the teamwork climate scale scores. Teamwork climate scale scores were computed by taking the average of the six items (one item was reverse scored because of its negative wording). In addition to the means used in multivariate analysis of variance, analysis of variance, and internal scale reliabilities, we also report the percent agreement (agree slightly plus agree strongly) for items and scale scores of each position and hospital. For each item, we call this “% agree,” and for the scale, we call it “% reporting good teamwork.” All statistical analyses were performed using SPSS version 13.0 (SPSS Inc., Chicago, IL) and *Mplus* version 2.01 (Muthén & Muthén, Los Angeles, CA).

Results

There were 2,135 respondents out of 2,769 questionnaires administered in 60 hospitals (222 surgeons, 1,058 OR nurses, 564 surgical technicians, 170 anesthesiologists, and 121 certified registered nurse anesthetists), for an overall response rate of 77.1% (range across hospitals, 57–100%). OR nurses had the highest response rate (79%), and certified registered nurse anesthetists had the lowest (67%) (table 1). The average respondent was 43 yr old with 10 yr of experience at the current hospital. Surgeons (8.6% female) and anesthesiologists (12.7% female) were predominantly male.

Teamwork Climate Scores

Here we describe how caregiver responses to six teamwork climate items were used to generate a representative and reliable teamwork climate score. A multilevel confirmative factor analysis was conducted to test the validity of the teamwork climate scale at the hospital OR level. A single, latent multilevel structure fit the data well (Comparative Fit Index = 0.99, Tucker-Lewis Index = 0.98, root mean square error of approximation = 0.05, standardized root mean square residual_{within} = 0.02, standardized root mean square residual_{between} = 0.02). All items had standardized loadings greater than 0.80 on the between factor. The six items also showed acceptable internal consistency overall and by caregiver type (overall α = 0.79, surgeon α = 0.78, anesthesiologist α = 0.78, certified registered nurse anesthetist α = 0.81, OR nurse α = 0.79, surgical technician α = 0.77). Overall, these results provide support for the reliability and validity of the teamwork climate scale score as an assessment of teamwork at the hospital OR level.

Analyses on the scale score provided strong support for the legitimacy of aggregating individual-level responses to the hospital OR level (*i.e.*, perceptions of teamwork within the ORs of a given hospital cluster together consistently, indicating specificity to that particular hospital). The ratio of total variance accounted for by group membership using the teamwork climate construct was in the range of acceptable values for reliability of a single assessment of a group mean, ICC(1) = 0.14, $P < 0.001$. The reliability of the group mean for

Table 2. Teamwork Climate Items and Descriptives

| Teamwork Climate Scale Item | Overall | Surgeons | Anesthesiologists | Surgical Technicians | CRNAs | OR Nurses |
|--|--------------------------------|-------------------------------|-------------------------------|--------------------------------|--------------------------------|-------------------------------|
| I have the support I need from other personnel to care for our patients | 3.94 (1.005); 76.6 (52.5–100) | 4.15 (0.917); 82.7 (20.0–100) | 4.06 (0.887); 78.1 (28.6–100) | 3.94 (1.001); 68.4 (16.7–100) | 4.00 (0.896); 80.5 (45.5–100) | 3.87 (1.046); 72.9 (26.7–100) |
| It is easy for personnel in this clinical area to ask questions when there is something that they do not understand | 3.96 (1.078); 75.8 (44.4–100) | 4.31 (0.870); 85.9 (50.1–100) | 4.10 (0.923); 74.9 (40.0–100) | 3.89 (1.136); 69.8 (40.0–100) | 3.93 (0.926); 67.1 (42.9–93.3) | 3.90 (1.109); 72.2 (33.3–100) |
| Nurse input is well received in this clinical area | 3.96 (1.021); 76.2 (45.2–100) | 4.35 (0.781); 83.9 (57.1–100) | 4.11 (0.894); 71.8 (25.0–100) | 4.01 (0.948); 74.2 (38.1–100) | 3.85 (0.933); 72.6 (34.8–100) | 3.84 (1.102); 71.2 (35.7–100) |
| In this clinical area, it is difficult to speak up if I perceive a problem with patient care (reverse scored) | 3.48 (1.252); 56.6 (24.1–94.1) | 3.87 (1.201); 67.3 (0–100) | 3.70 (1.219); 61.3 (22.2–100) | 3.39 (1.225); 53.2 (20.0–90.9) | 3.43 (1.287); 59.7 (18.2–85.7) | 3.42 (1.261); 53.7 (12.5–100) |
| Disagreements in this clinical area are appropriately resolved (<i>i.e.</i> , not who is right, but what is best for the patient) | 3.32 (1.289); 52.6 (12.5–100) | 4.00 (0.995); 69.1 (20.0–100) | 3.59 (1.132); 54.5 (12.5–100) | 3.22 (1.341); 46.0 (0–90.0) | 3.33 (1.215); 55.1 (8.7–80.0) | 3.19 (1.298); 46.7 (9.7–100) |
| The doctors and nurses here work together as a well-coordinated team | 3.73 (1.089); 66.0 (36.8–100) | 4.10 (0.957); 79.8 (40.0–100) | 3.97 (0.956); 73.6 (40.0–100) | 3.66 (1.118); 60.3 (20.0–90.0) | 3.70 (1.070); 70.9 (22.7–100) | 3.65 (1.101); 65.1 (20.0–100) |

Teamwork climate scale items and descriptive statistics for benchmarking teamwork climate data. Data are presented as mean score on a 1–5 scale, where 5 = agree strongly (SD); overall percent agreement (minimum and maximum percent agreement). Note that the item “In this clinical area, it is difficult to speak up if I perceive a problem with patient care” is reverse scored here so that the data in the table show the percentage of respondents who answered this item in the positive direction, which was to “Disagree.”

CRNA = certified registered nurse anesthetist; OR = operating room.

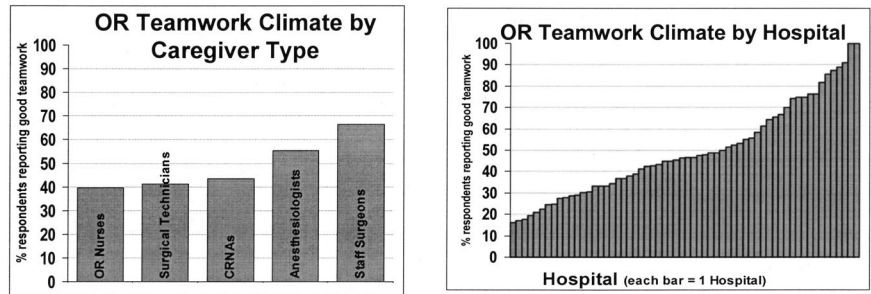
the construct was also within the range of acceptable values, $ICC(2) = 0.86$. The average within-group interrater reliability (r_{wg}) across units was 0.78, and above the standard 0.70 threshold. Together, these indices provide substantial justification for the legitimacy of OR teamwork climate as a collective-level construct.

Multivariate analysis of variance of the six items yielded two significant omnibus F results. An omnibus F for hospital OR level of $F_{354, 10,737} = 1.54$, $P < 0.001$, indicated that OR caregivers perceived teamwork issues differently as a function of the hospital in which they work. An omnibus F for OR caregiver of $F_{24, 6,266} = 3.28$, $P < 0.001$, indicated that OR caregivers perceive teamwork issues differently as a function of their role. OR caregivers responded differently to items about difficulty speaking up, conflict resolution, physician–nurse collaboration, feeling supported by others, asking questions

regarding uncertainties, and heeding nurse input. Table 2 displays the teamwork climate items and initial benchmarking information overall and by OR caregiver. Overall, only 53% of caregivers thought that disagreements were appropriately resolved. Less than half (47%) of OR nurses reported good conflict resolution, relative to 69% of surgeons and 55% of anesthesiologists.

After testing for differences on the teamwork climate items at the caregiver and hospital OR levels, we tested for differences on the scale scores. Analysis of variance demonstrated differences in teamwork climate scale scores between OR caregivers, $F_{4, 1,911} = 9.96$, $P < 0.001$; surgeons ($\bar{x} = 78.38$); anesthesiologists ($\bar{x} = 73.02$); certified registered nurse anesthetists ($\bar{x} = 67.53$); OR nurses ($\bar{x} = 66.14$); and surgical technicians ($\bar{x} = 67.17$); and also by hospitals, $F_{59, 1,911} = 4.06$, $P < 0.001$. Figure 1 shows the percent reporting good team-

Fig. 1. The graph on the left shows the percentage of respondents reporting good teamwork climate by operating room (OR) caregiver type. The graph on the right shows the percentage of OR respondents reporting good teamwork climate in each hospital. CRNA = certified registered nurse anesthetist.



work (the scale score equivalent of “agree slightly” or “agree strongly”) by OR caregiver and by hospital OR level.

Discussion

Frontline OR caregivers provided assessments of their teamwork climate that were empirically and psychometrically sound and varied among caregiver types and hospitals. This resulted in a valid and reliable teamwork climate score that is representative of OR caregiver responses in a single hospital. From the perspective of frontline caregivers, some hospitals had very good OR teamwork climate, whereas others were relatively poor. OR teamwork climate seemed to be in the eye of the beholder for two reasons: (1) Variability was greater between hospital OR levels than within; and (2) a consistent pattern occurred across the 60 hospitals studied, revealing that caregiver types had discrepant attitudes about teamwork with each other.

Overall, surgeons and anesthesiologists seemed more satisfied with physician–nurse collaboration than nurses did. OR nurses and certified registered nurse anesthetists did not reciprocate the high ratings of teamwork climate given by physicians. Our data suggest that this global difference in frontline caregiver assessments may be due to several specific issues. Relative to physicians, nurses were less positive about speaking up, feeling supported by others, physician–nurse collaboration, conflict resolution, and heeding nurse input. The origins of these discrepant attitudes are not fully understood. Clearly, there are fundamental differences between nurses and physicians, including status, authority, gender, training, and patient care responsibilities. Discussions with respondents during survey feedback presentations revealed that nurses often described good collaboration as “having their input respected,” whereas physicians often described good collaboration as having nurses “who anticipate their needs and follow instructions.”

Historically, physicians and nurses have different expectations and styles of communicating. Nurses are trained to communicate more holistically, relaying the “story” of the patient, whereas physicians are trained to succinctly communicate the “headlines.”³³ Differences in communication expectations and techniques may

have roots in medical and nursing educational cultures. Nevertheless, recent evidence from critical care demonstrates that providing structure to physician and nurse communications is an effective method for improving shared understanding of the situation, which in turn is associated with subsequent reductions in duration of stay.³⁴

The Joint Commission on Accreditation of Healthcare Organization’s recent identification of culture as an important component of a hospital system has spurred many hospitals to find valid methods to measure culture. The SAQ elicits caregiver attitudes that can assess culture in the OR. Specifically, we demonstrated how the interpersonal component of the OR work environment can be measured using the teamwork climate scale of the SAQ. This psychometrically sound assessment provides benchmarks for anesthesiology and surgery departments and hospitals seeking to compare their teamwork climate to national means and can serve as a baseline measure for evaluating interventions. Results consistently showed the appropriateness of grouping data at the hospital OR level (OR caregivers within each hospital) and demonstrated substantial differences between hospitals in perceptions of teamwork. Future research should explore the reasons for this variability because it may indicate that some hospitals with higher teamwork climate scores are engaging in potential best practices that could be identified and disseminated to the perioperative community at large. For example, previous research³⁵ compared hospitals with better-than-average and worse-than-average clinical outcomes and found that site visits helped to identify differences in their coordination of care activities.

Although some caregivers may be skeptical about the benefits of structured communications, preoperative OR briefings and postoperative debriefings (to review and learn lessons from the case for future patient outcomes) may improve teamwork and predictability in the OR and hence improve performance. Verbalizing the plan and setting expectations is not a novel idea for teams at work in high-risk environments. In commercial aviation, where captains and first officers are responsible for hundreds of lives at a time, standard practice is to conduct predeparture briefings to set expectations and establish an open environment for communication during the

flight.⁷ Preliminary evidence suggests that preoperative OR briefings are associated with improved teamwork climate, reductions in wrong-site/wrong-procedure surgeries, early reporting of equipment issues, and reduced operational costs.⁹ Taking a few minutes to brief the entire surgical team on the operative plan, familiarity with the procedure, potential issues, and names and roles is a quick and practical method to improve outcomes.

Limitations

The teamwork climate results reported here represent findings from 60 hospitals in one system. Organizations and researchers wishing to use this information for benchmarking purposes should be aware of two limitations. First, representative response rates are critical for interpreting climate research. As such, the ability of others to obtain the representativeness we achieved here would require methodologic rigor and support from senior leadership. Second, although originally designed to be a baseline assessment, many of the hospitals had already implemented specific interventions aimed at improving patient safety. Consequently, although the results identify significant opportunities for improvement, the overall distribution across the 60 hospitals may be higher than expected for a true baseline assessment.

We propose this distribution of 60 hospitals as an initial step in benchmarking OR teamwork with some justification. First, an ideal baseline benchmarking distribution in an industry that is continuously evolving and reacting to evidence is difficult. Indeed, anesthesiology has focused on improving patient safety for several decades. As such, any scores on teamwork climate should be viewed as a measure of current climate rather than "baseline." Second, many hospitals may find this benchmarking distribution useful because they are considering or already engaging in patient safety improvement efforts, such as executive walkrounds,³⁶ comprehensive unit-based safety programs,²⁰ or briefings. Third, we have collected OR data from an additional 38 hospitals using the SAQ, and only one fell outside (lower than) the currently reported distribution of teamwork climate scores ranging from 16% to 100% positive. Fourth, the sample included data from 16 states. Last, and albeit anecdotally, during our frequent SAQ results feedback sessions, hospital leaders, managers, and frontline caregivers found it extremely interesting and helpful to see and compare their teamwork climate score with that of other hospitals. That is, the comparison data foster powerful culture conversations that are reflective and may be helpful.

Another limitation to the study is in accounting for all relevant sources of variability in the OR teamwork climate. We found solid evidence that perceptions of teamwork vary among OR caregivers and among hospitals. However, in the current study we did not capture vari-

ability in teamwork perceptions among different surgical specialties or different surgeons within a specialty in a single hospital. We opted not to track that level of detail here but will do so in future administrations of the SAQ. Nevertheless, findings from this study demonstrate that investigating teamwork climate at the caregiver level and at the hospital OR level is justifiable and provides unique diagnostic information about OR issues and discipline-specific issues within an OR.

There may also be limitations in convergent and content validity in the current study. Because there are no widely used teamwork climate instruments in the OR for use as comparators, it was not possible to achieve convergent validity. Although still preliminary, OR briefing research⁹ has identified two phenomena that one would expect to be negatively related to teamwork, such as rates of annual nurse turnover and wrong-site surgeries. One could hypothesize that turnover and wrong-site surgery rates would decrease as teamwork climate increases.

Comprehensive content validity could be a limitation because we assessed the construct using only six items related to interpersonal dynamics in the OR. However, our goal was survey brevity with scale thoroughness, which required compromises in content that we expect will continue to evolve and improve over time. Nevertheless, previous observations of effective teamwork behaviors in the OR demonstrate that our six teamwork climate items are legitimate areas of inquiry. In addition, research from commercial aviation and nuclear power plants³⁷ found these same behaviors (*e.g.*, speaking up, resolving conflicts, asking questions, interdisciplinary collaboration) to be important in managing workload, preventing errors, and solving problems in safety-critical industries.

Our current findings add a new metric for assessing teamwork in the OR and add much detail to more general reports of discrepant attitudes about collaboration between physicians and nurses.³⁸ These findings in the OR setting are similar to teamwork climate findings in intensive care units.¹⁴ We identified specific teamwork-related attitudes that can be targeted to potentially improve clinical and operational outcomes such as nurse turnover rates, duration of stay, burnout rates, surgical case duration, wrong-site surgeries, burns, and postoperative infections, to name a few.

Attitudes about teamwork are associated with error reduction behaviors in aviation,¹⁹ patient outcomes in intensive care units,³⁸⁻⁴⁰ and nurse turnover in the OR.⁹ Good teamwork is associated with better job satisfaction⁴¹ and less sick time taken from work.⁴² National Aeronautics and Space Administration researchers have also found that teamwork can counteract some of the detrimental effects of fatigue on performance. Recent work in pediatric surgery suggests that good teamwork-related behaviors may lead to better patient outcomes.⁴³

Similarly, recent observational work during neonatal resuscitations demonstrated links among teamwork behaviors, independent assessments of resuscitation quality, and compliance with evidence-based medicine.⁴⁴ Discrepant attitudes about teamwork may also be a significant source of nurse dissatisfaction with their profession⁴⁵ that has led to our critical nursing shortage.⁴⁶ These discrepancies in perceptions of teamwork climate by OR caregivers do not extend to safety climate, but the significant variability in teamwork climate at the hospital OR level is consistent in safety climate.⁴⁷

Given the emerging evidence about the importance of teamwork, its relationship to outcomes, and the Joint Commission on Accreditation of Healthcare Organizations' future culture assessment requirement, we must address the lack of teamwork climate data for managers, department chairs, and hospital executives.

In conclusion, using the SAQ, we measured teamwork climate in the ORs of 60 hospitals. The analyses demonstrated the psychometric soundness of the teamwork climate scale, which provides a robust consensus score that is internally consistent overall and across OR caregiver types. Teamwork climate varies widely among hospitals, emphasizing the need to measure culture and tailor improvements efforts. Recent studies of OR briefings seem to be a promising area of research for improving teamwork climate in the OR, but they are an initial step in what will be a long journey to develop and refine a variety of tools to understand and improve OR teamwork. We could and should focus on what surgical teams do well in an attempt to build constructively on what we learn from investigating defects. Given that teamwork climate is important, measurable, and responsive to interventions, teamwork climate assessments from front-line caregivers using the SAQ could become a routine measure of patient safety.

References

1. Institute of Medicine: Creating safety systems in health care organizations, To Err Is Human: Building a Safer Health System, Chapter 8, Report from Committee on Quality of Health Care in America Edition. Edited by Kohn L, Corrigan J, Donaldson M. Washington, D.C., National Academy Press, 1999, pp 155-202
2. Gawande A, Studdert D, Orav E, Brennan T, Zinner M: Risk factors for retained instruments and sponges after surgery. *N Engl J Med* 2003; 348:229-35
3. Gibbs V, Auerbach A: The retained surgical sponge, Making Health Care Safer: A Critical Analysis of Patient Safety Practices. Evidence report/technology assessment. No. 43. Edited by Shojania K, Duncan B, McDonald K, Wachter R. Rockville, Maryland, Agency for Healthcare Research and Quality, 2001, pp 255-7. AHRQ publication 01-E058
4. Edmonds C, Liguori G, Stanton M: Two cases of a wrong-site peripheral nerve block and a process to prevent this complication. *Reg Anesth Pain Med* 2005; 30:99-103
5. Joint Commission on Accreditation of Healthcare Organizations: Sentinel Events: Evaluating Cause and Planning Improvement, 2nd edition. Oakbrook Terrace, Illinois, Joint Commission on Accreditation of Healthcare Organizations, 1998
6. Helmreich R, Sexton J: Group interaction under threat and high workload, Group Interaction in High Risk Environments, Chapter 1. Edited by Dietrich R, Childress T. Aldershot, United Kingdom, Ashgate, 2004, pp 9-23
7. Helmreich R, Wilhelm J, Gregorich S, Chidester T: Preliminary results from evaluation of cockpit resource management training: Performance ratings of flight crews. *Aviat Space Environ Med* 1990; 61:576-9

8. Sexton J, Helmreich R, Thomas E: Error, stress and teamwork in medicine and aviation: Cross sectional surveys. *BMJ* 2000; 320:745-9
9. DeFontes J, Surbida S: Preoperative safety briefing project. *Permanente J* 2004; 8:21-7
10. Makary M, Sexton J, Freischlag J, Holzmueller C, Millman E, Pronovost P: Teamwork in the operating room: Teamwork in the eye of the beholder. *J Am Coll Surg* 2006; 202:746-52
11. Sexton J, Holzmueller C, Pronovost P, Thomas E, McFerran S, Nunes J, Thompson D, Knight A, Penning D, Fox H: Variation in caregiver perceptions of teamwork climate in labor and delivery units. *J Perinatol* 2006; 26:463-70
12. Sexton J, Thomas E: Measurement: Assessing a safety culture, Achieving Safe and Reliable Healthcare: Strategies and Solutions, Chapter 7. Chicago, Health Administration Press, 2004, pp 115-27
13. Sexton J, Thomas E, Pronovost P: The context of care and the patient care team: The Safety Attitudes Questionnaire, Building a Better Delivery System. A New Engineering/Health Care Partnership. Edited by Reid P, Compton W, Grossman J, Fanjiang G. Washington, D.C., The National Academies Press, 2005, pp 119-23
14. Thomas E, Sexton J, Helmreich R: Discrepant attitudes about teamwork among critical care nurses and physicians. *Crit Care Med* 2003; 31:956-9
15. Helmreich R, Merritt A, Sherman P, Gregorich S, Wiener E: The Flight Management Attitudes Questionnaire (FMAQ). Austin, University of Texas, 1993, pp 93-4
16. Helmreich R: Cockpit management attitudes. *Hum Factors* 1984; 26:583-9
17. Gregorich S, Helmreich R, Wilhelm J: The structure of cockpit management attitudes. *J Appl Psychol* 1990; 75:682-90
18. Foushee H: Dyads and triads at 25,000 feet: Factors affecting group process and aircrew performance. *Am Psychologist* 1984; 39:885-993
19. Helmreich R, Foushee H, Benson R, Russini W: Cockpit resource management: Exploring the attitude-performance linkage. *Aviat Space Environ Med* 1986; 57:1198-200
20. Pronovost P, Weast B, Rosenstein B, Sexton JB, Holzmueller C, Paine L, Davis R, Rubin H: Implementing and validating a comprehensive unit-based safety program. *J Patient Safety* 2005; 1:33-40
21. Sexton J, Klinec J: The link between safety attitudes and observed performance in flight operations, Proceedings of the Eleventh International Symposium on Aviation Psychology. Columbus, The Ohio State University, 2001, pp 7-13
22. Helmreich RL, Merritt AC: Culture at Work in Aviation and Medicine: National, Organizational, and Professional Influences. Aldershot, United Kingdom, Ashgate, 1998, p 176
23. Colla JB, Bracken AC, Kinney LM, Weeks WB: Measuring patient safety climate: A review of surveys. *Qual Saf Health Care* 2005; 14:364-6
24. Sexton J, Marsch S, Helmreich R, Betzendorfer D, Kocher T, Scheidegger D: Jumpseating in the operating room. *J Hum Perform Extreme Environ* 1996; 1:2-36
25. Thomas E, Sherwood G, Mulhollem J, Sexton BJ, Helmreich RL: Working together in the neonatal intensive care unit: Provider perspectives. *J Perinatol* 2004; 24:552-9
26. Thomas E, Sexton J, Helmreich R: Translating teamwork behaviors from aviation to healthcare: Development of behavioral markers for neonatal resuscitation. *Qual Saf Health Care* 2004; 13:i57-64
27. Cronbach L: Coefficient alpha and the internal structure of tests. *Psychometrika* 1951; 16:297-333
28. Dyer N, Hanges P, Hall R: Applying multilevel confirmatory factor analysis techniques to the study of leadership. *Leadership Q* 2005; 16:149-67
29. Muthen B: Multilevel covariance structure analysis. *Sociol Methods Res* 1994; 22:376-98
30. Klein K, Bliese P, Kozlowski S, Dansereau F, Gavin M, Griffin M, Hoffman D, James L, Yammarino F, Bligh M: Multilevel analytical techniques: Commonalities, differences and continuing questions, Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions. Edited by Klein K, Kozlowski S. San Francisco, Jossey-Bass/Pfeiffer, 2000, pp 512-53
31. Bliese P: Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis, Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions. Edited by Klein K, Kozlowski S. San Francisco, Jossey-Bass/Pfeiffer, 2000, pp 349-81
32. James L, Demaree R, Wolf G: Estimating within-group interrater reliability with and without response bias. *J Appl Psychol* 1984; 69:85-98
33. Sexton BJ: The Better the Team the Safer the World: Golden Rules of Group Interaction. Ladenburg, Germany, 2004, pp 1-58
34. Pronovost P, Berenholtz S, Dorman T, Lipsett P, Simmonds T, Haraden C: Improving communication in the ICU using daily goals. *J Crit Care* 2003; 18:71-5
35. Young G, Charns M, Daley J, Forbes M, Henderson W, Khuri S: Best practices for managing surgical services: The role of coordination. *Health Care Manage Rev* 1997; 22:72-81
36. Thomas E, Sexton J, Neilands T, Frankel A, Helmreich R: The effect of executive walk rounds on nurse safety climate attitudes: A randomized trial of clinical units. *BMC Health Serv Res* 2005; 5:28
37. Sexton J, Thomas E: Measurement: Assessing a safety culture, Achieving Safe and Reliable Healthcare: Strategies and Solutions, Chapter 7. Chicago, Health Administration Press, 2004, pp 115-27
38. Baggis J, Schmitt M, Mushlin A, Mitchell P, Eldredge D, Oakes D, Hutson

- AD: Association between nurse-physician collaboration and patient outcomes in three intensive care units. *Crit Care Med* 1999; 27:1991-8
39. Shortell S, Zimmerman J, Rousseau D, Gillies R, Wagner D, Draper E, Knaus W, Duffy J: The performance of intensive care units: Does good management make a difference? *Med Care* 1994; 32:508-25
40. Knaus W, Draper E, Zimmerman J: An evaluation of outcome from intensive care in major medical centers. *Ann Intern Med* 1986; 10:410-8
41. Posner B, Randolph W: Perceived situation moderators of the relationship between role ambiguity, job satisfaction and effectiveness. *J Soc Psychol* 1979; 109:237-44
42. Kivimaki M, Sutinen R, Elovainio M, Vahtera J, Rasanen K, Toyry S, Ferrie J, Firth-Cozens J: Sickness absence in hospital physicians: 2 year follow-up study on determinants. *Occup Environ Med* 2001; 58:361-6
43. de Leval M, Carthey J, Wright D, Farewell V, Reason J: Human factors and cardiac surgery: A multicenter study. *J Thorac Cardiovasc Surg* 2000; 119:661-72
44. Thomas E, Sexton J, Lasky R, Helmreich R, Crandell S, Tyson J: Teamwork and quality during neonatal care in the delivery room. *J Perinatol* 2005; 26:163-9
45. Aiken L, Clarke S, Sloane D, Sochalski J, Busse R, Clark H, Giovannetti P, Hunt J, Rafferty A, Shamian J: Nurses' reports on hospital care in five countries. *Health Affairs* 2000; 20:43-53
46. Bednash G: The decreasing supply of registered nurses: Inevitable future or call to action? *JAMA* 2000; 283:2985-7
47. Makary M, Sexton J, Freischlag J, Millman E, Pryor D, Holzmuller C, Pronovost P: Patient safety in surgery. *Ann Surg* 2006; 243:628-35