The Association Between Sensemaking During Physician Team Rounds and Hospitalized Patients' Outcomes

Luci K. Leykum, M.D., M.B.A., M.Sc.[1,2,3], *Hannah Chesser, M.D.*[4], *Holly J. Lanham, Ph.D.*[1,2,3], *Pezzia Carla, Ph.D.*[2,5], *Ray Palmer, Ph.D.*[6], *Temple Ratcliffe, M.D.*[1,2], *Heather Reisinger, Ph.D.*[7], *Michael Agar, Ph.D.*[8], *and Jacqueline Pugh, M.D.*[1,2]

[1]South Texas Veterans Health Care System, San Antonio, TX, USA; [2]Department of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA; [3]McCombs School of Business, University of Texas at Austin, San Antonio, TX, USA; [4]School of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA; [5]University of Dallas, Dallas, TX, USA; [6]Department of Family and Community Medicine, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA; [7]Iowa City VA Health Care System, Iowa City, IA, USA; [8]Ethknoworks LLC, Santa Fe, NM, USA.

BACKGROUND: Sensemaking is the social act of assigning meaning to ambiguous events. It is recognized as a means to achieve high reliability. We sought to assess sensemaking in daily patient care through examining how inpatient teams round and discuss patients.

OBJECTIVE: Our purpose was to assess the association between inpatient physician team sensemaking and hospitalized patients' outcomes, including length of stay (LOS), unnecessary length of stay (ULOS), and complication rates.

DESIGN: Eleven inpatient medicine teams' daily rounds were observed for 2 to 4 weeks. Rounds were audiotaped, and field notes taken. Four patient discussions per team were assessed using a standardized Situation, Task, Intent, Concern, Calibrate (STICC) framework.

PARTICIPANTS: Inpatient physician teams at the teaching hospitals affiliated with the University of Texas Health Science Center at San Antonio participated in the study. Outcomes of patients admitted to the teams were included.

MAIN MEASURES: Sensemaking was assessed based on the order in which patients were seen, purposeful rounding, patient-driven rounding, and individual patient discussions. We assigned teams a score based on the number of STICC elements used in the four patient discussions sampled. The association between sensemaking and outcomes was assessed using Kruskal-Wallis sum rank and Dunn's tests.

KEY RESULTS: Teams rounded in several different ways. Five teams rounded purposefully, and four based rounds on patient-driven needs. Purposeful and patient-driven rounds were significantly associated with lower complication rates. Varying the order in which patients were seen and purposefully rounding were significantly associated with lower LOS, and purposeful and patient-driven rounds associated with lower ULOS. Use of a greater number of STICC elements was associated with significantly lower LOS (4.6 vs. 5.7, p=0.01), ULOS (0.3 vs. 0.6, p=0.02), and complications (0.2 vs. 0.5, p=0.0001).

CONCLUSIONS: Improving sensemaking may be a strategy for improving patient outcomes, fostering a shared understanding of a patient's clinical trajectory, and enabling high reliability.

KEY WORDS: complexity science; sensemaking; length of stay; complication rates; inpatient teams.
J Gen Intern Med 30(12):1821–7
DOI: 10.1007/s11606-015-3377-4

Received November 27, 2014
Accepted April 15, 2015
Published online May 27, 2015

INTRODUCTION

Hospitalized patients' quality of care remains variable, despite significant efforts put into patient safety and quality improvement.[1,2] Consistent decreases in adverse events have not been achieved.[3] In this context, increased attention has been placed on high reliability, a concept that includes achieving consistently high performance levels,[4] not only through reducing failures, but also through improving recognition and action when failures occur.[5] High reliability organizations are characterized by real-time operational awareness, recognition of task complexity, use of near-misses to identify improvement targets, deference to expertise, and resilience among individuals throughout the organization.

High reliability has been studied in the organizational literature.[6–8] Sensemaking has emerged as a critical activity for achieving and maintaining high reliability.[9,10] Sensemaking is defined as a social act of assigning meaning to ambiguous events or data, and acting based on that meaning: "Sensemaking is a diagnostic process directed at constructing plausible interpretations of ambiguous cues that are sufficient to sustain action."[11,12] At the organizational level, this may include how individuals interpret and react to changes in competitive landscapes, new strategic initiatives, or internal reporting structure. In healthcare, sensemaking occurs at the organizational level, but also in the context of delivering care to individual patients. Through sensemaking, we assimilate data and contextual cues to reach conclusions that allow us to act. Effective provider sensemaking is one path to achieving high reliability in health care.

Clinical reasoning is often used to describe how clinicians come to an understanding of a patient. It has been defined as "the cognitive operations allowing clinicians to observe, collect, and analyze information that ultimately leads to an action."[13] However, just as clinical reasoning may be considered broader in scope than decision making, sensemaking may be considered broader than clinical reasoning. Sensemaking encompasses not only understanding individual patients, but also making sense of the competing tasks required by a group of patients.

How is sensemaking achieved in patient care? Providers must not only diagnose and treat patients, but also anticipate complications, recognize changes in a patient's course, and communicate effectively during care transitions. For example, Ghaferi examined differences between low and high-surgical mortality hospitals, expecting differences in complications to be a primary cause of mortality differences.[14] Instead, mortality rates were associated with how quickly and effectively providers recognized and managed complications—termed "failure to rescue." To better understand the ways providers make sense of what is happening with their patients in real-time, we examined sensemaking among inpatient medical physician teams, a work context characterized by high task interdependence and highly distributed cognition. While there have been limited observational studies of physician rounds, they have focused on how and where time is spent, and not on the substance of actions or discussions.[15] We observed how teams made sense of their daily tasks based on how they rounded, as well as how they discussed individual patients each day on rounds. We then assessed the association between patterns of team sensemaking and patient outcomes of length of stay (LOS), unnecessary length of stay (ULOS), and complication rates.

METHODS

Setting and Study Participants

This study was conducted at the Audie L. Murphy Veterans Affairs Hospital (ALMVAH) in the South Texas Veterans Health Care System (STVHCS) and University Hospital (UH), the primary teaching hospitals of the University of Texas Health Science Center at San Antonio (UTHSCSA). ALMVAH is the 220-bed acute care facility for STVHCS. UH is the 614-bed level-I trauma center for the Bexar County health district. This study was approved by the UTHSCSA Institutional Review Board, the STVHCS Research and Development Committee, and the UHS Research Committee.

As part of a study of inpatient medicine team relationships,[16] we observed inpatient teams over nine month-long periods. On two teams, the attendings changed halfway through the month. Because of resultant differences in team characteristics, these half-months were analyzed as separate teams. We rounded with teams daily, and one to two members of the research team took field notes and audiotaped rounds. We also observed any initial discussions of team expectations. Teams were comprised of one attending physician, one post-graduate-year (PGY)-2 or PGY-3 resident, and two interns. Two to three medical students, and occasionally a sub-intern, pharmacist, or Doctor of Pharmacy student were also team members. Patients with any acute medical illnesses were admitted to these teams, excepting new onset cardiac conditions.

We purposefully sampled teams, focusing on attending characteristics (sampling a range of years of experience and involvement in clinical, educational, and administrative activities), and time of year.[16–18] We obtained physician information based on our knowledge of the faculty pool. We sampled across the year excluding July or August to allow resident team members to settle in to their roles.

For this analysis, we used audiotaped recordings and field notes regarding team member discussions prior to rounds, and the order of teams' movements during rounds, as detailed below. We collected data regarding each patient admitted to the teams, including discharge diagnosis and comorbidities. We also collected daily census and number of admissions and discharges to assess team workload.

Assessing Sensemaking

We assessed sensemaking in two ways: first, by how teams made sense of daily rounding tasks, and second, how they made sense of individual patients' clinical courses.

To assess sensemaking of daily rounding tasks, we examined how teams rounded on their patients, as this would reflect their assessment and prioritization of tasks. We categorized each team's rounding practices in three ways: (1) whether teams explicitly discussed how they should round, termed "purposeful rounding," (2) the order in which physician teams saw patients, termed "rounding order," and (3) whether teams rounded in a patient-driven manner.

We noted whether an explicit discussion of the order in which patients should be seen occurred, which we considered evidence of "purposeful rounding." If there was a patient-driven reason for starting in a certain order, such as a patient worsening or being ready for discharge, we considered that "patient-driven." We categorized the order in which patients were seen using the rounding orders described in Table 1. We assigned one rounding order to each team daily, and summed

Table 1 Rounding Orders Observed

Rounding order	Definition
Gravity	Starting at the highest floor and moving down
Geography	Starting on a particular unit and moving to contiguous units
Intern by intern	Rounding on one intern's patients first and then the other intern's
Running the board	Rounding in the team room, discussing patients in the order they are written on the board, or in which their card appears in the attending's stack
Patient-driven	Prioritizing specific patients to see based on their clinical needs

the total number of rounding orders each team used over the course of our observations.

To assess sensemaking of individual patients, we listened to audiotaped discussions of four patient encounters per team. We purposefully sampled for encounters of complex clinical scenarios, but to limit potential confounding by variation in clinical diagnoses, sought consistency between scenarios across teams. To accomplish this, we first identified two discussions of complex patients on post-call days. The post-call day, or day after teams admit their highest potential patient number, is when team census is highest, and when much of the initial diagnosis and treatment is formulated, with diagnostic results returning, and treatment responses assessed. Post-call discussions provided insight into how teams made sense of patients' illnesses as they were immediately unfolding. We identified patients based on admitting diagnosis, including cirrhosis-related illness, pneumonia, or exacerbations of heart failure or chronic obstructive pulmonary disease. We then identified the subset of patients with age-adjusted Charlson Comorbidity scores between 6 and 10.[19] From this group, we randomly selected two patient discussions per team.

The second group of two patient encounters for each team related to complex patients with complications, or new problems that developed in the hospital. We assessed discussions that occurred on the days complications were recognized, as they could provide insight into how teams recognized, assessed, and made sense of complications. Again, to have consistency in the clinical scenarios, we focused on a narrow range of frequent complications, including acute kidney injury, hospital-acquired infection, or change in mental status. We first identified patients on each team that developed a complication. We then identified the subset with the same range in Charlson Comorbidity score, randomly selecting two patients per team. For one team, there was only a single patient who experienced a complication that met our criteria.

We used the Situation, Task, Intent, Concern, Calibrate (STICC) framework to assess sensemaking of patient discussions.[20] This framework was designed by sensemaking researchers to assess communication failures leading to inpatient adverse events, making it appropriate for assessing how teams made sense of what was happening during discussions on daily rounds. STICC elements, operational definitions, and examples are shown in Table 2. Two members of the research team independently listened to each discussion to assess presence of STICC elements.

We assigned each team a score based on the number of STICC elements reflected in each patient care discussion. Each time an element was noted in a discussion, the team was given a point. Thus, teams could receive 0 to 5 points for each patient discussion based on the number STICC elements utilized. Zero would reflect no STICC elements, and 5 would reflect all. The points for each element were then summed for a total of up to 20 points. For the team that had only three encounters that met our criteria, we imputed a total score based on those of the other encounters.

Assessing Patient Outcomes

Complications were defined as the development of a new problem, or a clinical deterioration after at least 24 h of stability. Specifically, new problems included newly identified clinical issues that were not part of the initial presentation and required additional action, such as an increased creatinine, fever, or fall. Clinical deteriorations included escalation of care in a patient previously stable or improving, such as a transfer to an increased level of care. These were assessed based on daily patient discussions, with chart review of problem lists in daily notes, laboratory studies, or diagnostic results performed if the discussion was unclear. LOS was determined based on the number of days the patient was in the hospital, including admission and discharge days. This was generally well-reflected in each day's field notes, but chart review was performed to confirm each patient's dates. ULOS was defined as days a patient remained hospitalized after being medically stable for discharge, based on the teams' daily patient discussions.

To avoid the possibility that care provided by another team would influence either LOS or ULOS, we included only patients whose entire hospitalization occurred under the care of the observed teams. Because we observed some teams for half-months, the longest possible LOS for patients cared for only by those teams was 13 days. Thus, we excluded patients with LOS longer than 13 days.

Table 2 STICC Elements and Definitions

Element	Definition	Examples
Situation	Discussion of "here is what we are dealing with." Working diagnosis	"What do you think is the most likely diagnosis?" "So what do you think is going on with the patient?"
Task	Assessment of "what we are going to do." Specific next steps should be explicitly discussed.	"So what we need to do is get him pooping. Give him lactulose 30 cc every 1–2 h to start."
Intent	Explicit, concrete discussion of why the team is embarking on a specific diagnostic or therapeutic plan.	"Why are we starting different antibiotics in this patient with pneumonia?"
Concern	Discussion of "what we need to keep our eye on" or "what we need to look out for." Should be specific to the patient, not only general to the disease.	"What types of side effects do we need to look for?" "What in-hospital complications is this patient at highest risk for?"
Calibrate	"Talk to me." Discussion regarding what the team might be missing, what is unclear or not yet understood. If-then contingency statements.	"What aspects of this patient don't quite fit together?" "Hearing this now, do you see anything we may have missed?"

We included all complications that occurred while patients were under the care of the teams we observed. Thus, patients who were transferred to or from other teams were included. However, to minimize the likelihood that the complications reflected another team's care, only complications occurring after 24 h were included for transferred patients.

Statistical Analysis

We assessed the association between our measures of sensemaking and patient outcomes. We first assessed each variable's distribution.[21,22] Because some had a skewed distribution, we used the Kruskal-Wallis sum rank test for between group comparisons, utilizing Dunn's test for pair-wise comparisons if overall significance occurred.[23]

Because there are not acceptable methods for covariate adjustments using the Kruskal-Wallace method, we compared patient characteristics and workload for teams with varying levels of sensemaking using a general linear modeling approach to ensure that there were not systematic differences that could impact our results. All models were run using SAS software.[24] Agreement and kappa were calculated using STATA 13.1.[25]

RESULTS

We observed 1941 discussions of 576 patients over 207 days. Two hundred and ninety-eight patients were included in our LOS and ULOS analyses. Two hundred and sixty-seven were excluded because they received care from another team, and 11 because their LOS exceeded 13 days. Three hundred and ninety-eight patients were included in our analysis of complications. Patient and workload characteristics for each team are shown in Table 3.

Five teams had explicit discussions of their rounding order. The discussions occurred relatively infrequently, from three to five times per team over the course of the observations. We dichotomized teams into two groups, categorizing teams as "purposeful rounders" if they ever discussed their rounding order, and as not if they never discussed it.

We observed five rounding orders. Based on the distribution of rounding orders used by the teams, we stratified teams into low, medium, and high, with low being one rounding order used during the entire period of observation, medium being two or three rounding orders, and high reflecting at least four. Teams whose rounding orders were more varied had significantly shorter mean rounding times than teams that did not [92.8 min (12.2) vs. 119.0 (20.9), $p=0.05$]. Four teams rounded in a patient-driven manner. Patient-driven rounding was always purposeful, typically because a patient was deemed "sick." The number of times we observed patient-driven rounding was infrequent, from one to three times per team. We dichotomized teams based on whether or not they ever rounded in this way. Rounding practices are summarized in Table 3.

The association between rounding orders, purposeful or patient-driven rounds, and patient outcomes, is shown in Table 4. Teams with more rounding orders had significantly lower complication rates ($p=0.02$). Purposefully rounding teams had significantly lower LOS, ULOS, and complications. Finally, teams that rounded in a patient-driven fashion also had

Table 3 Patient Characteristics, Workload Metrics, and Observed Rounding Practices for Each Team

	Team										
	1	**2**	**3(1)***	**3(2)***	**4(1)***	**4(2)***	**5**	**6**	**7**	**8**	**9**
Mean Age-adjusted Charlson Score (SD)	6.26[ab] (2.97)	7.09[ab] (2.86)	7.10[ab] (3.56)	7.14[ab] (3.17)	6.71[ab] (2.45)	6.40[ab] (2.99)	4.24[b] (2.72)	6.77[ab] (2.72)	8.71[a] (20.52)	5.04[b] (2.78)	6.60[ab] (2.87)
Mean Daily Census (SD)	14.11[b] (1.94)	15.80 [a] (1.76)	9.87[f] (0.69)	11.41[de] (1.85)	11.87[d] (2.81)	8.92[g] (0.91)	16.35[a] (2.18)	15.55[a] (1.78)	12.72[c] (1.73)	12.70[c] (1.91)	9.25[fg] (1.43)
Mean Daily Admissions (SD)	2.37[cde] (0.75)	2.78[bc] (1.69)	1.42[g] (0.78)	2.29[cde] (0.96)	2.55[bcd] (1.83)	1.55[gf] (0.55)	3.61[a] (2.05)	2.54[bcde] (1.28)	3.08[ab] (1.07)	2.79[bc] (1.13)	2.87[bc] (1.57)
Number rounding patterns	1	2	4	2	3	1	2	2	4	4	3
Purpose-ful rounding	N	N	Y	N	N	N	Y	N	Y	Y	Y
Patient-driven rounding	N	N	Y	N	N	N	N	N	Y	Y	Y
Rounding time (mean & std dev)	111.1 (34.8)	114.5 (32.6)	106.8 (37.3)	93.4 (33.7)	116.7 (79.2)	103.2 (32.8)	142.8 (28.6)	151.4 (63.5)	85.4 (30.4)	80.1 (19.9)	98.8 (46.5)
Situation	4	2	4	2	4	3	4	4	4	4	4
Task	1	2	4	2	1	2	4	3	4	4	4
Intent	2	1	3	1	2	3	3	2	4	4	4
Concern	1	0	2	2	0	0	3	1	4	2	2
Calibrate	0	0	2	0	0	1	2	0	4	1	2
STICC score	8	5	15	7	7	9	16	10	20	15	16

*(1) and (2) denote observations of half-month teams

Letters by numbers are from a one-way Anova with Dunnett's test for multiple comparisons. Means with the same letters are not significantly different

Table 4 Number of Observed Rounding Orders, Use of Purposeful or Patient-Driven Rounds, and Patient Outcomes

Outcome Mean (std dev) Median	Number of observed rounding orders			Purposeful rounds		Patient-driven rounds	
	Low	Med	High	No	Yes	No	Yes
LOS (in days)	5.6 (4.5) 4	5.3 (4.4) 4	4.5 (3.7) 4	5.7 (4.7) 4	4.6 * (3.7) 3	5.7 (4.6) 4	4.3* (3.5) 3
ULOS (in days)**	0.6 (1.5)	0.5 (1.8)	0.2* (0.6)	0.6 (1.7)	0.3* (1.2)	0.6 (1.8)	0.2 (0.8)
Complication rates (per patient per day)**	0.5 (0.8)	0.4 (0.8)	0.2* (0.6)	0.5 (0.9)	0.2* (0.6)	0.5 (0.9)	0.2* (0.6)

**significant at p<0.05*
*** Median ULOS and complication rates were 0 for all groups*
LOS length of stay, ULOS Unnecessary length of stay

significantly lower LOS ($p<0.001$) and complications ($p<0.001$).

Inter-rater agreement for use of STICC elements in individual patient discussions was 90 % and kappa 0.8. STICC element utilization and STICC scores are shown in Table 3. Most teams discussed the "situation" in all of their patient discussions, but few discussed issues of concern or calibration. We stratified teams into low and high groups. For use of each STICC element, low was defined as 0–2 and high as 3–4. For STICC scores, low was less than 15 and high was15 or higher. The associations between use of STICC and patient outcomes are shown in Table 5. Higher STICC scores, reflecting more frequent use of all STICC elements in patient discussions, were significantly associated with lower LOS, ULOS, and complications. Discussion of situation was associated with lower LOS and complication rates. Discussion of task and intent were associated with lower LOS, ULOS, and complications.

Because we could not adjust for patient characteristics and team workload, we compared these between "low" and "high" sensemaking teams. The five teams with high STICC scores were the same teams that purposefully rounded. Four of these five teams used a "high" number of rounding orders and rounded in a patient-centered order. Therefore, we compared both of these groupings of high sensemaking teams versus other teams. The only consistent significant difference between low and high teams was age-adjusted Charlson score (6.4 vs. 5.5, $p<0.001$). High and low teams did not differ with regard to workload (average daily census or number of admissions).

DISCUSSION

Inpatient teams differ in how they make sense of their rounding activities and individual patients' illnesses. Higher sensemaking measures were associated with lower LOS, ULOS, and complication rates. We used rounding practices to assess how teams made sense of their rounding tasks. Explicit discussion of rounding order was infrequent. When it occurred, it usually reflected team needs (such as an intern having afternoon clinic) rather than patient needs. Despite this physician focus, purposeful rounding was associated with improved outcomes, suggesting that assessing team daily tasks and rounding accordingly might be an important marker of general team function or workload / time management. There was high overlap between teams who used a higher number of rounding orders, rounded purposefully, and rounded in a patient-driven order, suggesting these activities reflect a similar sensemaking capability.

Teams also differed in their sensemaking of individual patients as assessed using STICC, and teams using more elements had improved outcomes. Interestingly, teams with higher rounding sensemaking measures also utilized more

Table 5 Comparison of Length of Stay, Unnecessary Length of Stay, and Complication Rates by Results of the STICC Framework (Situation, Task, Intent, Concern, Calibrate) Results Using the Kruskal-Wallis Test

Outcome Mean (std dev) Median	Situation		Task		Intent		Concern		Calibrate		Total STICC	
	Low	High	Low	High	Low	High	Low	High	Low	High	Low	High
LOS (in days)	6.4 (4.7) 4	4.9* (4.1) 3	5.7 (4.7) 4	4.6* (3.7) 3	5.7 (4.6) 4	4.3* (3.5) 3	5.1 (4.3)4	5.0 (3.9) 4	5.2 (4.4) 3	4.6 (3.5) 4	5.7 (4.7) 4	4.6* (3.7) 4
ULOS (in days)**	0.6 (1.7)	0.4 (1.5)	0.3 (1.2)	0.6 (1.7)	0.6 (1.8)	0.2 (0.8)	0.5 (1.5)	0.3 (1.5)	0.5 (1.6)	0.2 (0.7)	0.6 (1.7)	0.3* (1.2)
Compli-cations (per pt per day)**	0.6 (1.1)	0.3* (0.7)	0.5 (0.9)	0.2* (0.6)	0.5 (0.9)	0.2* (0.6)	0.3 (0.7)	0.3 (0.7)	0.4 (0.8)	0.3 (0.8)	0.5 (0.9)	0.2* (0.6)

**Significant at p<0.05*
*** Median ULOS and complication rates were 0 for all groups*
LOS Length of stay, ULOS Unnecessary length of stay

STICC elements, raising the possibility that we are assessing two aspects of a "sensemaking capacity" on physician teams.

Sensemaking provides one lens to view clinical reasoning not just as an individual endeavor, but rather as a shared, distributed, or social experience. Similar to distributed cognition,[13,26] sensemaking recognizes that it is the interactions among the physicians, other providers and allied health care workers, the patient, and the environment (e.g., time constraints or an intern thinking about afternoon clinic) that lead to the emergence of understanding, meaning and action. With its organizational origins, sensemaking expands the traditional focus of clinical reasoning from individual patients to the management of all patients on a team.

Using a sensemaking lens rather than only a clinical reasoning perspective makes several important aspects of patient care explicit. First, teams have to not only take care of individual patients, but also actively manage their census of patients, prioritizing tasks, deploying resources, and avoiding harm. Second, sensemaking makes "systems" aspects of caring for patients more explicit, touching on issues of how we best care for patients in a particular context, at a specific time, all within the constraints of a specific healthcare system. While we did not assess sensemaking with regard to other providers taking care of the patient—notably nursing—we believe that the ways that all of these providers make sense of the patient are important, representing an area for future research.

The social nature of sensemaking and distributed cognition highlight the creation of a shared mental model among providers,[27] which may also reflect tacit knowledge among the team.[28,29] Our findings reinforce the importance of this shared meaning, as higher discussion of "intent," the "why we are doing this," was associated with lower complication rates and LOS. While concern and calibrate were not in themselves associated with outcomes, we do not believe that this reflects a lack of importance. Instead, it may reflect the infrequency of their use. The association of total STICC score and outcomes may speak to their importance. More explicit inclusion of intent, concern, and calibrate as part of assessment and plan discussions may be a strategy for enriching patient discussions and further improving patient outcomes.

The recognition that teams of providers must manage groups of patients within a specific organizational context highlights the importance of sensemaking to overall system function and high reliability. By explicitly recognizing sensemaking as a social activity among providers and patients, we can expand our social view of clinical reasoning beyond its traditional focus on diagnosis and treatment, to include safety, efficiency, provider interdependencies, and harm prevention. Additionally, we extend our understanding of the types of things to which providers must attend when taking care of hospitalized patients: communication, contextualization, and complex task management. The shared mental models fostered by effective sensemaking enable the consistently effective performance found in high reliability systems.

Our study is limited by its being conducted in two hospitals and utilizing a small number of teams. While we listened to only four discussions per team, we were able to identify similar patients across teams, allowing us to make comparisons that would not otherwise be possible. It is possible that our focus on specific types of discussions and complications limits the generalizability of these results to other clinical contexts or presentations. This small sample did allow us to obtain a richness that afforded an in-depth assessment of how teams round and discuss patients that can serve as platform for future work. It is possible that we missed some discussion of rounding order, but by dichotomizing teams based on ever observing a discussion, we believe we have assessed the team's general approach to how they would round. Finally, teams with higher sensemaking metrics cared for patients with lower Charlson scores, but given the small degree of difference, we do not believe it was clinically impactful. Examining sensemaking across institutions, across an expanded scope of patients and providers, and developing interventions that incorporate aspects of sensemaking into patient care activities, would be directions for future work that would build on and enrich this initial understanding.

Acknowledgements: *We would like to thank the members of the teams involved in this study.*
The research reported here was supported by the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development Service (CDA 07-022). Investigator salary support is provided through this funding, and through the South Texas Veterans Health Care System. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs.
A subset of this analysis was presented at the 2014 Society of General Internal Medicine annual meeting in San Diego, California.

Conflict of Interest: *The authors declare that they do not have a conflict of interest.*

Corresponding Author: *Luci K. Leykum, M.D., M.B.A., M.Sc.; South Texas Veterans Health Care System, 7400 Merton Minter, San Antonio, TX 78229, USA (e-mail: Leykum@uthscsa.edu).*

REFERENCES

1. **James JT.** A new, evidence-based estimate of patient harms associated with hospital care. J Patient Saf. 2013;9(3):122–8.
2. **Williams SC, Schmaltz SP, Morton DJ, Koss RG, Loeb JM.** Quality of care in U.S. hospitals as reflected by standardized measures, 2002–2004. NEJM. 2005;353:255–64.
3. **Landrigan CP, Parry GJ, Bones CB, Hackbarth AD, Goldmann DA, Sharek PJ.** Temporal trends in rates of patient harm resulting from medical care. N Engl J Med. 2010;323(22):2124–35.
4. **Perrow, C**. Normal Accidents: Living with High-Risk Technologies. Princeton University Press; 1999.
5. http://archive.ahrq.gov/professionals/quality-patient-safety/quality-resources/tools/hroadvice/hroadvice.pdf. Accessed 4-16-15.
6. **Roberts KH.** Managing high reliability organizations. Calif Manag Rev. 1990;32(4):101–14.
7. **Roberts KH.** Some characteristics of one type of high reliability organization. Organ Sci. 1990;1(2):160–76.
8. **Weick KE, Sutcliffe KM, Obstfeld D.** Organizing for high reliability: processes of collective mindfulness. In: Staw BM, Cummings LL, eds. Research in organizational behavior, vol. 21. Greenwich, CT: JAI Press; 1999:81–123.

9. **Weick KE.** The collapse of sense making in organizations: the Mann Gulch disaster. Adm Sci Q. 1993;12:628–52.
10. **Blatt R, Christianson MK, Sutcliffe KM, Rosenthal MM.** A sensemaking lens on reliability. J Organ Behav. 2006;27:897–917.
11. **Weick KE, Sutcliffe KM**. Managing the unexpected: assuring high performance in an age of complexity. University of Michigan Business School Series; 2001.
12. **Weick KE, Sutcliffe KM, Obstfeld D.** Organizing and the process of sensemaking. Organ Sci. 2005;16(4):409–21.
13. **Holmboe ES, Durning SJ.** Assessing clinical reasoning: moving from in vitro to in vivo. Diagnosis. 2014;1(1):111–7.
14. **Ghaferi AA, Birkmeyer JD, Dimick JB.** Complications, failure to rescue, and mortality with major inpatient surgery in medicare patients. Ann Surg. 2009;250(6):1029–34.
15. **Stickrath C, Noble M, Prochazka A, Anderson M, Griffiths M, Manheim J, Sillau S, Aagaard E.** Attending rounds in the current era. JAMA Intern Med. 2013;173(12):1084–9.
16. **McAllister C, Leykum LK, Lanham HJ, Reisinger HS, Kohn JL, Palmer RF, Pezzia C, Agar M, Parchman ML, Pugh JA, McDaniel RR.** Relationships within inpatient physician housestaff teams and their association with hospitalized patient outcomes. J Hosp Med. 2014. doi:10.1002/jhm.2274.
17. **Patton MQ**. Qualitative research. Encyclopedia of statistics in behavioral science. 2005.
18. **Pope C, van Royen P, Baker R.** Qualitative methods in research on healthcare quality. Qual Saf Health Care. 2002;11:148–52.
19. **Deyo RA, Cherkin DC, Ciol MA.** Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. J Clin Epidemiol. 1992;45(6):613–9.
20. **Sutcliffe KM, Lewton E, Rosenthan MM.** Communication failures: an insidious contributor to medical mistakes. Acad Med. 2005;79:186–95.
21. **Tukey JW.** Exploratory data analysis. Reading, MA: Addison-Wesley; 1977.
22. **Zar JH.** Biostatistical analysis 4th Ed. Upper Saddle River, NJ: Pearson Prentice-Hall; 2010.
23. **Dunn OJ.** Multiple contrasts using rank sums. Technometrics. 1964;6:241–52.
24. SAS Inc. SAS/STAT software version 9.1. Cary, NC: SAS Institute Inc; 2003.
25. STATA Version 13. College Station, TX
26. **Durning SJ, Artino AR.** Situativity theory: a perspective on how participants and the environment can interact: AMEE guide no. 52. Med Teach. 2011;33:188–99.
27. **Huber GP, Lewis K.** Cross-understanding: implications for group cognition and performance. Acad Manage Rev. 2010;35(1):6–26.
28. **Nonaka I.** A dynamic theory of organizational knowledge creation. Organ Sci. 1994;5(1):14–37.
29. **Polanyi M**. The tacit dimension. London, UK: Routledge & Kegan; 1966.