Objective Measures of Communication Behavior Predict Clinical Performance

Yuhao Peng, BS,* Nicholas E. Anton, MS,† Jackie Cha, MS,* Tomoko Mizota, MD,† Julie M. Hennings, MSN, RN, CCRN,† Ryan Stambro‡, Megan A. Rendina, BS,† Katie J. Stanton-Maxey, MD,† Dimitrios Stefanidis, MD, PhD,† and Denny Yu, PhD*

*School of Industrial Engineering, Purdue University, West Lafayette, Indiana; †Department of Surgery, Indiana University School of Medicine, Indianapolis, Indiana; and ‡Simulation Center at Fairbanks Hall, Indiana University Health, Indianapolis, Indiana

OBJECTIVE: Effective teamwork and communication are critical to patient outcomes, and subjective assessment tools have been studied in predicting team performances. However, inherent biases remain while using subjective assessment tools. This study hypothesizes that objective communication features can assess and predict clinical performance.

DESIGN: Forty 3rd-year medical students participated in the Acute Care Trauma Simulation as the role of doctor, teaming up with a nurse confederate and a simulated patient. Participants conducted postoperative patient management, patient care diagnoses, and treatment. Audio from all team members were recorded, speech variables (e.g., speech duration, number of conversations, etc.) were extracted, and statistical analyses were performed to associate communication with clinical performance.

SETTING: This study was performed at the simulation center located at Fairbanks Hall, Indiana University School of Medicine.

PARTICIPANTS: Data from forty 3rd-year medical students were collected and analyzed.

RESULTS: Majority (67%) of the communications were initiated by student. Speech ratio, intensity, and frequency of communications differed when students communicate with nurse than with patient (e.g., student communication to patient had higher intensity than nurse). Increasing frequency of check-backs between student and nurse ($p < 0.05$) and speech duration from student to patient ($p = 0.001$) positively associated with student's clinical performance score.

CONCLUSION: Objective communication features can predict medical trainee's clinical performance and provide an objective approach for simulation-based trauma care training. (J Surg Ed 76:1337–1347. © 2019 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: team communication, communication analysis, patient safety

COMPETENCIES: Patient Care, Professionalism, Interpersonal and Communication Skills

INTRODUCTION

Effective teamwork in healthcare has been a known contributor to positive patient outcomes.[1-5] Communication contributes to building cohesive healthcare teams and also influences patient care.[6-8] Studies on healthcare teams show that teams with frequent communication among team members achieve higher productivity than teams with less frequent communication.[9] According to the Joint Commission, communication failures are one of the most frequently identified root causes (30%) of sentinel events.[10,11] In 6 Danish hospitals, 52% of 84 root cause analysis reports cite verbal communication errors (e.g., misinterpreted information between physicians, especially during handoffs), which provides further evidence that effective communication is a major factor contributing to patient safety.[12] Furthermore, when deconstructing communication failures over 22 hours of audio recorded during 6 high-acuity surgical procedures, researchers found that communication failures occurred once every 8 minutes.[13] Studies have found that communication failures also varied depending on the phase of the surgery, with 54% of errors

Correspondence: Inquiries to Denny Yu, PhD, Purdue University, School of Industrial Engineering, 315 N. Grant Street, West Lafayette, IN 47907-2023; e-mail: dennyyu@purdue.edu

occurring between incision and closure. Moreover, the importance of communication in safety is not just limited to healthcare, but has been frequently shown important in other safety critical domains, such as aviation, military, etc.[14-17]

Communication is a well-established area of study. Generally defined as the behavior of reducing uncertainty from one individual to another, it occurs whenever there is a need for information processing.[18-20] In hospital settings like the emergency department (ED) or clinical ward, communication can appear in many different forms and modalities. Some examples include verbal (e.g., face-to-face, telephone, radio), nonverbal (e.g., gestures), and literal (e.g., writing and reading).[21,22] In a study classifying information-sharing activities among 38 clinicians during 159 patient encounters across 7 sites, verbal communication (i.e., with patient, staff, or colleague) contributed to 60% of patient encounter time, the remaining time included writing, reading, examination, and other.[23] Communication patterns also differ whether the target audience is a healthcare professional or patient. Healthcare professionals often use both medical language and everyday language among themselves and patients; however, everyday language is the most frequently used form when communicating between healthcare professionals and patients.[24] In addition to varying audiences, communication patterns may also change according to circumstances at the moment, including time, safety and sterility, resources, roles, and situation.[25]

Understanding and monitoring communication can offer insight into healthcare providers' abilities to deliver safe and effective patient care. Several tools and measurement constructs have been proposed to quantify communication, identify patterns that impact outcomes, and guide team training in healthcare. For example, the Team Strategies and Tools to Enhance Performance and Patient Safety (TeamSTEPPS) is one tool that has been widely used to improve team performance and has demonstrated positive effects on patient outcomes.[26-28] TeamSTEPPS is a comprehensive multiphase teamwork intervention that focuses on leadership, situation monitoring, mutual support, communication, and communication is one of the TeamSTEPPS competencies.[29] Another tool is the Communication Assessment Tool-Team (CAT-T), designed to assess team communications; however, the tool similarly relies on subjective evaluation (i.e., 5-point response scale, from 1 = poor to 5 = excellent).[21,30,31] Communication assessment is also emphasized in the area of interprofessional nontechnical skills.[32-34] Similar to the previous tools, nontechnical skills assessment tools assess communication with behavioral anchors used by expert raters as reference points. However, like other tools, a key limitation is the dependency on subjective Likert scales.[34,35]

Subjective rating tools are limited by several inherent biases. In medical education, systemic rater errors have been observed and include biased ratings due to halo effects, severity, central tendency, leniency, logical error, inattention, restriction of the range, etc.[36] For example, subjective evaluation by raters tended to result in a positive/leniency bias, meaning higher scores were given to persons even they performed poorly; low severity on evaluations from expert to novice or vice versa; and other serious shortcomings that were usually ignored.[37] In addition, common in all the aforementioned tools for assessing communication, there are limitations in scalability due to the need for trained experts to perform the evaluations.[38] Due to the recognized impact of communication on patient care and limitations of the currently available subjective assessment tools, there is a need to identify objective and scalable approaches to assess and evaluate healthcare providers' communication.

Sensing-based approaches may overcome some limitations of current assessment tools by providing objective and potentially automated communication assessments. Several studies have used sensors to measure communication in team-based interpersonal interactions.[39] Audio sensors have been used to infer communication patterns, nurses' personality, and team-workload distribution.[40-42] Previous work has shown that continuous audio streams can estimate social interaction patterns through wearable devices,[39] and techniques have been proposed (e.g., Latent Semantic Analysis) to assess team communication by transforming raw speech into textual input.[20,43] One of the most common uses of audio sensors in healthcare is for content-analysis to complement field observations.[44] These studies capture audio to measure the usage of content-based communication events, distinguish between verbal or nonverbal communication, and quantify frequency and duration of communication events. However, many of these approaches still rely on an analyst to code the data, and few studies have investigated the ability of audio metrics that can be automatically processed from audio recordings (e.g., speech intensity, duration, rate, etc.) to help access clinical performance.

Our research team is therefore working on identifying more objective and continuous assessments of nontechnical performance in the clinical environment using audio sensors that can ultimately facilitate targeted performance-enhancing interventions. Toward this goal, this study aimed to (1) identify communication features and patterns that can be derived from audio data-streams and represent healthcare provider's communication behavior and (2) assess whether objectively obtained communication features are associated with clinical performance.

METHODS

Study Participants

This research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board at Indiana University (IRB: #1611105172). Informed consent was obtained from each participant. Participants were recruited from one academic medical institution, and each participant provided informed consent to allow for audio and video recording of scenario-based simulations. These Acute Care Team Simulations (ACTS) served as a summative assessment of medical student performance after they had completed their general surgery clerkship rotation.

Study Procedure

Each ACTS scenario took place in a state-of-the-art simulation center located at Fairbanks Hall, Indiana University School of Medicine. This simulation center was designed to educate multidisciplinary healthcare providers through the replication of multiple facets of the patient care environment including a fully functional operating room environment and intensive care unit rooms that can be controlled entirely by a simulation technologist using a computer system. This design allowed healthcare providers to engage in immersive simulated patient care scenarios and learn safe practices for actual patient care.

During ACTS, the manikin features pulse and breathing sounds among other physiological features, which can be manipulated according to the details of the simulation scenario. Each ACTS session involved a team consisting of one medical student (i.e., all students were on their surgery clerkship rotation during their 3rd year of medical school), a nurse confederate, and a simulation technologist controlling the manikin (SimMan 3G, Laerdal Medical, Wappingers Falls, NY). The technologist acted as the participant and his voice was heard through the manikin. Students were randomized into 1 of 6 scenarios: (1) Motor vehicle accident shock, (2) Pneumothorax, (3) Hyponatremia, (4) Leg compartment syndrome, (5) Pulmonary embolism, and (6) Heparin-Induced Thrombocytopenia.

These 6 scenarios captured the range of care management from patient arrival to the ED, postoperative patient management, patient care diagnoses, and treatment. Each scenario averaged approximately 10 minutes in length. Typical tasks performed by the medical student in every scenario included: completing a patient assessment, determining a diagnosis, and identifying an appropriate treatment. Medical students were in charge of the patient's care and they were expected to communicate with both the simulated patient and nurse to deliver correct patient care safely and effectively. Based on the medical student's actions, a simulation technologist manipulated the patient's health (i.e., improve, remain the same, or worsen) under the guidance of an experienced faculty observer. Each student performed 1 scenario. At the end of each scenario, the medical student's clinical performance was rated by the same nurse confederate, who had extensive experience with ACTS assessments.

Data Collection

Audio from each healthcare team member and patient was recorded throughout the scenario. Voice recorders (Zoom H1, Zoom, Inc, Hauppauge, NY) were placed in the participant's pocket, and a lapel microphone (RØDE smartLav+ Microphone, RØDE Microphones, Silver Water, NSW, Australia) was attached to the scrub or jacket collar of the student, nurse, and patient. Although more intrusive than audio recordings from a video recorder, this approach allowed better localization of audio source, noise cancelling, and <1-minute setup time. Video recordings with 3 room views (i.e., patient view, overhead view, and a view of patient vitals) were collected using the cameras (Panasonic WV-CS574, Panasonic, Kadoma, Osaka Prefecture, Japan) built into the simulation center (B-Line Medical, Washington, DC). These were used as needed to verify and interpret audio observations.

Data Analysis

Communication Variables

Communication patterns were analyzed by a trained study researcher listening and annotating the recordings. The assessments focused on "check-backs," part of the closed-loop communication (CLC) strategy according to Härgestam et al. that consists of 3 elements[45]: (1) call-out, (2) check-back, and (3) closed-loop. First, the conversation initiator transmits a message as either with a question or statement (call-out). Then, the receiver acknowledges the message with a response to the initiator (check-back). Lastly, the response from the receiver should be verbally verified by the initiator to complete CLC. In this study, closed-loop communication was excluded because the principles of CLC were not part of the standard medical curriculum, hence they were not expected to utilize this form of specialized communication. Accordingly, only call-out and check-back were analyzed, and any verbal response is considered as a check-back. For example, if person A transmitted a message to person B, and person B provided a response to A with a message, then a check-back was recorded. In another example, person A transmitted a message and person B responded with a question as follows:

Person A: "Can I have . . .?"
Person B: "Do you mean . . .?"

In this example, there was a verbal response from person B to A, and at the same time, person B transmitted a new message. Our study considered person B's response a check-back, and a call-out from person B to A. In summary, conversations in the present study were classified using the following categories:

(1) Communication initiated with a question followed by a response,
(2) Communication initiated with a statement (i.e., non-question) followed by a response,
(3) Communication initiated with a question followed by no response, and
(4) Communication initiated with a statement followed by no response.

In the ACTS scenarios, communication could originate from 3 potential sources (i.e., student, nurse, and patient). This resulted in a total of 6 combinations (i.e., only two-way communications were measured in this study): student to nurse, student to patient, nurse to student, nurse to patient, patient to student, and patient to nurse. Although rare, when there was three-way communication, for example, both nurse and patient responding to student's message, it was classified by the content of whom the message initiator was speaking to. The number of check-backs, also known as the frequency of communications, was calculated from all 6 two-way combinations at the 4 different categories listed above. This analysis was completed by a research team member using custom Microsoft Excel software to annotate while listening to the audio recordings.

Audio Processing

In addition to decomposing the full audio files into segments as mentioned previously, audio features from the scenarios were also extracted using Praat software.[46] Since ACTS scenarios were focused on assessing medical students' clinical performance, this study concentrated on the student's audio features and communication frequency with the patient and nurse. For this study, a communication event (or segment) was defined by each change in medical students' communication target (nurse or patient). Figure 1 demonstrates an example of a series of 3 communication events from the student to the nurse and patient; the first communication event during the first 34 seconds was between the student and the nurse. The background noise of patient/nurse audio have been filtered. In each of the three-communication events, the following variables of medical students' audio were calculated: (1) speech duration: the total time the student spoke to nurse or patient, (2) speech ratio: the percentage of time medical students spoke during a communication event, (3) speech intensity: loudness, and (4) speech rate: speech speed or pace.

Clinical Performance

Medical students' ACTS performance was evaluated by the nurse confederate who participated in all the scenarios and sessions. Clinical performance was rated using a 100-point Visual Analogue Scale (VAS) from 1 = unacceptable for their level to 100 = outstanding for their level. Validity evidence for this type of performance assessment in medical education is presented elsewhere.[47]

Statistical Analysis

All statistical analyses were performed on Minitab.[48] *t* Tests were performed to identify differences in audio variables between student to nurse and student to patient. Pearson's correlation was used to determine associations between individual audio variables and performance. Regression analysis with forward stepwise variable selection was performed to determine statistically (alpha = 0.05) significant audio predictors for clinical performance. Furthermore, two-way random

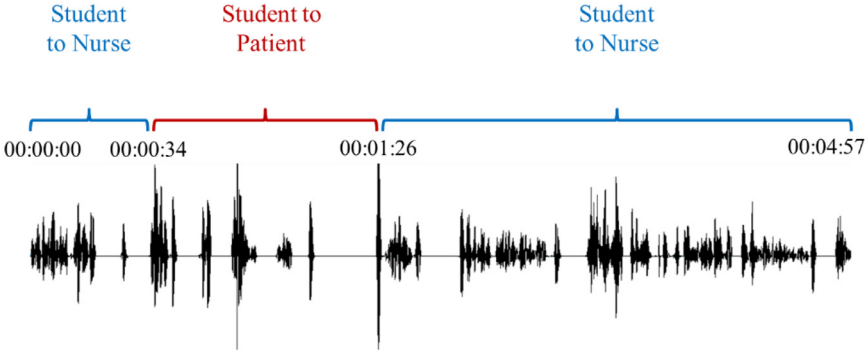

FIGURE 1. An example of student's audio and separated communication events.

TABLE 1. Average Duration and Clinical Score of Each Scenario

Scenario	Participated Students	Clinical Score Average	Clinical Score Standard Deviation
Motor vehicle accident shock	8	62.8	28
Pneumothorax	7	65.1	25
Hyponatremia	7	76.4	13.1
Leg compartment syndrome	6	71.3	26.2
Pulmonary embolism	6	72	17.8
Heparin-Induced Thrombocytopenia	6	45	25.9

intraclass correlation (ICC) coefficients were calculated for the video analysis for agreement and consistency. Five random cases were tested for reliability, where the raters labeled question/nonquestion communications following with/without check-backs. ICC scores within the range of 0.21 to 0.40 were interpreted as fair, 0.41 to 0.60 as moderate, 0.61 to 0.80 as substantial, and 0.81 to 1.00 as almost perfect agreement.[49]

RESULTS

All 40 participants completed the study. The frequency of each scenario and clinical performance score is shown in Table 1. Frequency across scenarios was not uniform as participation in the study was voluntary and randomly assigned.

Average frequency of communication segments per scenario were as follows: 39 segments initiated by medical students, 11 segments initiated by nurse, and 5 segments initiated by patient. Figure 2 summarizes the conversations with (Fig. 2a) and without (Fig. 2b) response from initiator to receiver. The nodes represent each individual role in the simulation, and the thickness of the arrows represents the proportion of communications over all scenarios. The direction of the arrow head represents initiator to receiver relationship. The majority of conversations with response (Fig. 2a) were initiated from the student (67%). Conversations initiated by the patient were the least frequent. Five percent of communication from the medical students did not receive a response from either the nurse or the patient (Fig. 2b). Only 1% of the communication resulted in medical students not responding to the nurse or patient (Fig. 2b).

Communications were further analyzed by type (question or statement), completion of a check-back, and roles, for example, student to nurse, student to patient, nurse to patient, etc. (Fig. 3). The ICC analysis for labeling type and check-back resulted in 0.859 and 0.879 for agreement and consistency, respectively. These coefficients fell in into the range of 0.81 to 1.00, which indicate significant reliability.

The most frequent type of communication was student to nurse questions with check-back (Fig. 3a). Out of an average 16 student-to-nurse questions per scenario (Fig. 3a and c), 99% received a check-back from the nurse. About 2.6% of questions from nurse or patient did not receive a student's check-back.

For statements, medical students received check-backs from the nurse 91% of the time. However, student statements to the patient received check-backs least frequently at 75% per scenario. Comparing across roles, communication between nurse and patient were the least frequently observed regardless of conversation type (question vs. statement) or occurrence of check-backs.

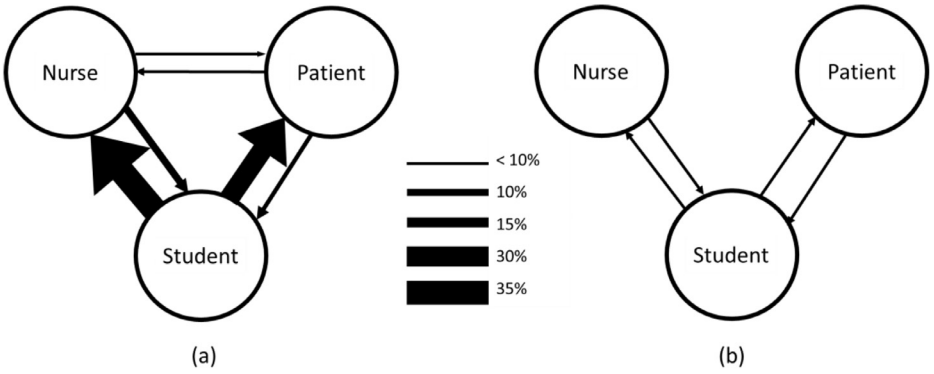

FIGURE 2. Link analyses of proportion of interpersonal communications observed between the team members (link widths are scaled to frequency): (a) with check-back (response), (b) without check-back (response).

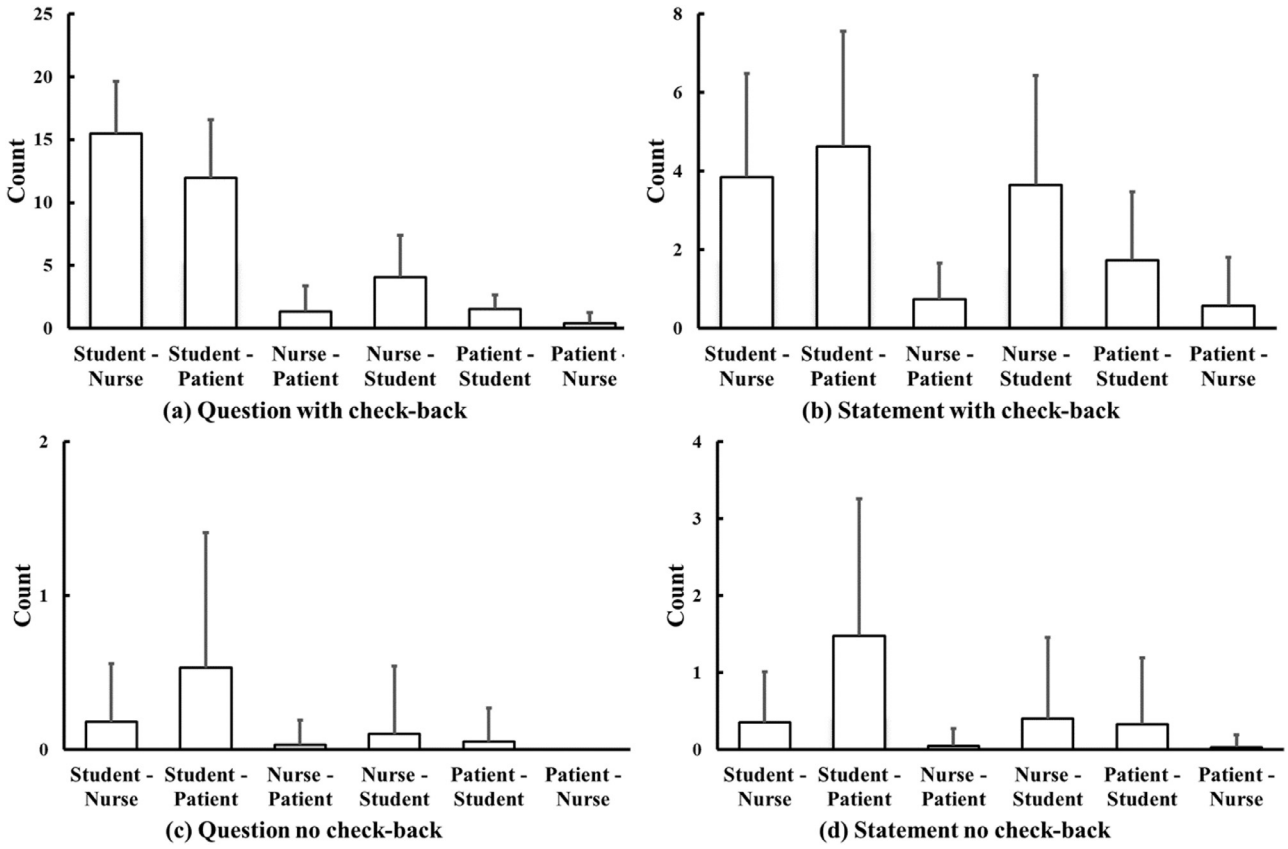

FIGURE 3. Interpersonal communications among student, nurse, and patient of all cases. (a) Initiated questions with a check-back (response). (b) Initiated statements with a check-back (response). (c) Initiated questions without a check-back (response). (d) Initiated statements without a check-back (response).

Findings from the audio processing are shown in Table 2. Comparison of speech features between student to nurse and student to patient showed that student audio features differed depending on whether s/he was speaking to the patient or nurse. On average, communication to patient was 3 dB louder ($p < 0.05$) than nurse, speech ratio to patient was 5% more than nurse ($p < 0.05$), and more questions and statements with check-back were directed to nurse ($p < 0.05$). Speech duration and speech rate were not statistically different whether the student was communicating with the nurse or patient.

Features derived from the audio recordings showed significant correlation to performance scores. Out of all the audio features, the strongest relationship with respect to performance score was the positive association (Pearson's $\rho = 0.493$, $p = 0.001$) between frequency communication initiated by the student (regardless of statement or question or to which team member) as shown in Figure 4. Frequency of communication between student and nurse (with check-back) were also significantly correlated with performance ($\rho = 0.456$, $p = 0.003$), while speech duration between student to nurse approached but did not reach significance ($\rho = 0.294$, $p = 0.066$).

TABLE 2. Comparison Between Student to Nurse and Student to Patient With Significant Differences Indicated

Speech Measurables	To Nurse		To Patient		*t* Value	p Value
	Mean	SD	Mean	SD		
Speech duration (s)	65.77	20.67	61.92	22.85	0.86	0.40
Speech ratio	0.21	0.06	0.26	0.07	−5.01	**<0.001**
Speech intensity: average (dB)	59.6	3.66	62.56	2.54	−6.7	**<0.001**
Speech rate (syllable/second)	2.48	0.6	2.47	0.58	0.14	0.89
Question and statement with check-backs	19.33	4.5	16.58	5.03	2.62	**0.01**

Bolded p-values indicated statistically significant variable ($p < 0.05$).

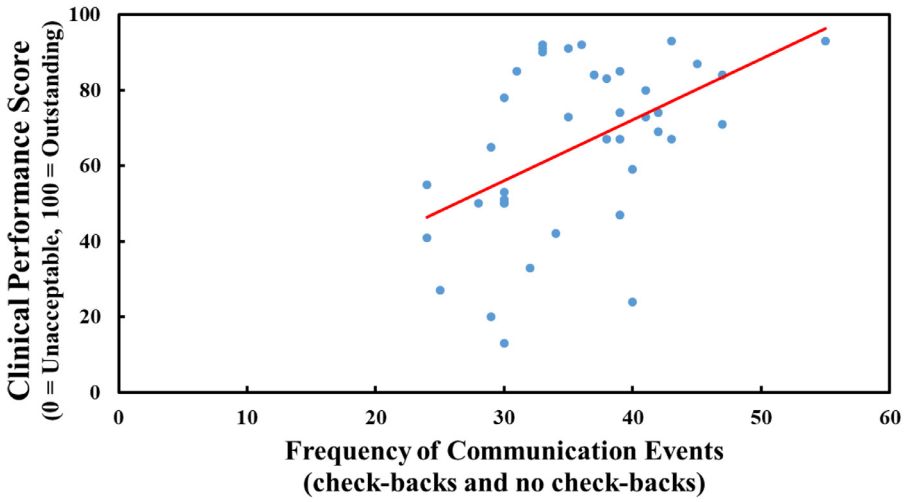

FIGURE 4. Correlation between frequency of communications and clinical performance.

Leveraging all communication metrics to determine the relationship between communication and performance, a forward stepwise regression approach identified significant predictors of performance. The stepwise selection removed variables that may be collinear, and stepwise selection result can be seen in Table 3. The final model resulted with an R-squared value of 0.61. With increasing frequency of *statements* from student to nurse that received check-backs resulted in improved performance ($p < 0.05$). Similarly, increased frequency of *questions* from student to nurse that received check-backs improved performance ($p < 0.001$). Variables describing conversation without check-backs (i.e., communication that received no response from the other individual) were not predictive of performance. For speech features, duration of speech directed toward the patient also had positive association ($p < 0.001$) with performance. Although not statistically significant, increasing speech intensity to patient ($p = 0.07$) and speech ratio ($p = 0.10$) were also associated with improved performance.

TABLE 3. Resulting Model From Stepwise Regression of Speech Features to Predict Clinical Performance

Terms	Coef	p Value
Constant	77.7	0.266
Student to nurse with check-back (Q)	3.268	**<0.001**
Nurse to student with check-back (Q)	1.23	0.154
Student to nurse with check-back (NQ)	2.52	**0.037**
Nurse to student with check-back (NQ)	−1.39	0.203
Patient to nurse with check-back (NQ)	4.49	0.059
Speech ratio overall	−112.8	0.1
Student to nurse speech Intensity	1.72	0.115
Student to patient speech Duration	0.586	**0.001**
Student to patient speech Intensity	−3.02	0.065

Q, communication started with question; NQ, communication started with statement.
Bolded p-values indicated statistically significant variable ($p < 0.05$).

DISCUSSION

Communication is widely known to impact clinical performance, and the findings of this study explored quantitative communication metrics that predict clinical performance in a team-based healthcare simulation. Several speech features were significantly associated with clinical performance, thus, the results of the study support our first hypothesis that communication features can be derived from audio data to further understand healthcare providers' communication behavior. The first approach distinguished communication by roles that initiated the dialogue, intent of the communication (i.e., statement or question), and by whether check-backs were observed was based on previous work in the clinical literature.[50] Our results suggest that medical students tended to direct more questions to the nurse rather than the patient, included questions such as asking the nurse for the patient's current status and requesting equipment. However, the main communications from student to patient were statements, which primarily focused on stating the causes of the patient's symptoms and communicating the care plan. In addition to interprofessional communication, patient communication is another area of major focus for medical education, training, and practice. Studies have shown that it is important for patients to receive unambiguous statements from healthcare providers during the diagnostic process, appropriate treatments can then be conveyed precisely.[51] According to closed-loop communication, this ensures a clear understanding regarding clinical information, and this information is shared among healthcare

providers before conducting appropriate patient care.[52] Results showed conversations with no response from patient to student were commonly detected, which was likely the result of the variety of symptoms and conditions presented in the different scenarios. For example, repeated examination by the student may be causing severe pain to the patient, and thus resulted in no response from the patient.

In seeking additional insight into communication, our results from medical students' speech features showed that several audio metrics yielded a significant outcome in comparison between students to nurse and students to patient. Although the audio recordings revealed that the students spent less time talking to the patient than to the nurse, the difference in speech duration was not significant statistically. Speech ratio was calculated based on the speech duration of the student during each conversation segment, where higher speech ratio from student to patient could indicate less action and thinking time, and a smaller speech ratio from student to nurse could indicate that there were more hands-on examinations of the patient and time spent making clinical decisions. Another observation was that medical students spoke with a higher volume to the patient than to nurse. In a study of patient satisfaction, a greater satisfaction was reported when physician was speaking in a higher volume to the patients.[53] It is unclear why this phenomenon was observed; however, we hypothesize that medical students were trying to communicate clearly with a patient who may be in physiological or psychological distress, or who they are not familiar with vs. a team member who they have worked with throughout their training. It is also possible that the lack of facial expressions of the manikin may have also influenced student behavior.

Lastly, statistical methods were performed to identify the relationships between measured variables and performance scores. A statistically significant variables had a positive relationship with their scores. The frequency of communication with nurse and patient yielded a high correlation with performance score (Fig. 4) compared with other communication variables: for example, speech duration, communication from student to nurse with check-backs. Students with a higher communication frequency had higher performance scores and may suggest that medical students asked questions to gather more information about the patient, to further identify the patient's symptoms with learned knowledge. Accordingly, statement-initiated conversations were later made by medical students to conduct corresponding treatments (e.g., think-aloud), which therefore resulted in a positive patient outcome. In addition to correlation, a predictive model that related these variables with performance was developed. The result from the stepwise regression model showed that there were 9 factors, resulting an R-squared value of 0.61. Communications from student to nurse with check-backs was especially significant in the regression model, which also suggests a strong correlation between frequency communication and performance score. From the regression model, the number of both question and statement events between student and nurse was significant, but not the frequency of conversation with patients. Rather, the duration speaking to patients was more predictive of performance.

Though clinical performance is known to be dependent on personal knowledge, communication is clearly a significant factor that contributes to clinical performance. Indeed, medical students' performance differed according to measured audio variables. Thus, the objective audio variables have the potential to augment clinical training and assessment. These variables may be useful in evaluating healthcare providers' communication in team-based simulations and the relationship of communication variables and healthcare providers' performance in the clinical setting.

This study contains limitations which are important to note. First, although the sample size of the study (40 participants) was adequate, this study was not designed to test participant characteristics such as gender, age, or ethnicity. Given that these demographic factors may affect speech behaviors, future study participant recruitment needs to focus on gender, age, and ethnicity to assess their impact on communication features. In this simulation, limited communication was observed between nurse and patient, which could be due to the simulation setting that medical students were the main evaluation subject, being tested by the nurse confederate and the manikin patient controlled by a technologist. Although the participants were simulating hospital roles, the diagnostic process, environment, and equipment were identical to real hospital settings. For the frequency of communications, check-back is only one part of close-loop communication where the receiver accepts the message and acknowledges its receipt; however in this study, any verbal response was accepted as a check-back. Future work should explore the more rigorous definition of check-back.

Six different scenarios were randomized across participants, and each diagnostic process was different across all scenarios. Furthermore, each scenario was different and thus resulted in different diagnostic process, medical examinations, and treatments. Some scenarios may require more hands-on examinations, causing a longer silence in the segment which resulted in a smaller speech ratio. Despite these differences, several speech metrics were significant suggestion that the approach using communication features for assessment can be generalizable

across a variety of clinical case scenarios, communication across a variety of clinical case scenarios.

Finally, the statistical models were generated through stepwise regression, but due to the richness of continuous audio recordings and features, future research with more robust statistical approaches may provide additional insights on the relationship between audio metrics and performance.

CONCLUSION

Effective communication in healthcare settings is important in enhancing team performance and providing positive patient outcomes. By analyzing interpersonal communication among medical student, nurse, and patient during a simulated clinical scenario using audio recordings this study demonstrated the feasibility of quantifying communication features and patterns among participants. Importantly, the results showed that the frequency of student communication with the nurse and patient predicted their clinical performance during the scenarios. Our future work will focus on discovering improved communication metrics, and an improved statistical modeling of performance.

ACKNOWLEDGMENTS

This research was made possible by Indiana University Health and the Indiana Clinical and Translational Sciences Institute, funded in part by grant #ULI TR002529 from the National Institutes of Health, National Center for Advancing Translational Sciences, Clinical and Translational Science Award and The Advances in Medicine (AIM) grant from Cook Medical. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or Cook Medical.

REFERENCES

1. Manser T. Teamwork and patient safety in dynamic domains of healthcare: a review of the literature. *Acta Anaesthesiol Scand*. 2009. https://doi.org/10.1111/j.1399-6576.2008.01717.x.
2. Baker DP, Gustafson S, Beaubien J, Salas E. Medical teamwork and patient safety: the evidence-based relation. Literature review. *Agency Healthc Res Qual*. 2005: 1–64. Publ No 05-0053.
3. Martin JS, Ummenhofer W, Manser T, Spirig R. Interprofessional collaboration among nurses and physicians: making a difference in patient outcome. *Swiss Med Wkly*. 2010;140(17-18):1–12.
4. Kohn LT, Corrigan JM, Donaldson MS. To Err Is Human: Building a Safer Health System, 6. Washington DC: The National Academies Press; 1999. https://doi.org/10.17226/9728.
5. Tiel Groenestege-Kreb D, Van Maarseveen O, Leenen L. Trauma team. *Br J Anaesth*. 2014. https://doi.org/10.1093/bja/aeu236.
6. Schmutz J, Manser T. Do team processes really have an effect on clinical performance? A systematic literature review. *Br J Anaesth*. 2013. https://doi.org/10.1093/bja/aes513.
7. Ellingson LL. Communication, collaboration, and teamwork among health care professionals. *Commun Res Trends*. 2002.
8. Grumbach K, Bodenheimer T. Can health care teams improve primary care practice? *JAMA*. 2004. https://doi.org/10.1001/jama.291.10.1246.
9. Reagans R, Zuckerman EW. Networks, diversity, and productivity: the social capital of corporate R&D teams. *Organ Sci*. 2001. https://doi.org/10.1287/orsc.12.4.502.10637.
10. Commission J. Sentinel event statistics released through second quarter 2015. *Jt Comm Online*. 2015.
11. CRICO Strategies. Malpractice risks in communication failures 2015 annual benchmarking report. CRICO.
12. Rabøl LI, Andersen ML, Østergaard D, Bjørn B, Lilja B, Mogensen T. Republished error management: descriptions of verbal communication errors between staff. An analysis of 84 root cause analysis-reports from Danish hospitals. *Postgrad Med J*. 2011. https://doi.org/10.1136/pgmj.2010.040238rep.
13. Hu YY, Arriaga AF, Peyre SE, Corso KA, Roth EM, Greenberg CC. Deconstructing intraoperative communication failures. *J Surg Res*. 2012. https://doi.org/10.1016/j.jss.2012.04.029.
14. Helmreich RL, Foushee HC. Why CRM? Empirical and theoretical bases of human factors training. Crew Resource Management; 2010. https://doi.org/10.1016/B978-0-12-374946-8.10001-9.
15. Sexton JB, Helmreich RL. Analyzing cockpit communications: the links between language, performance, error, and workload. *J Hum Perform Extrem Environ*. 2000. https://doi.org/10.7771/2327-2937.1007.
16. Anca J, Helmreich RL, Kanki BG. Crew Resource Management. 2010. https://doi.org/10.1055/s-0029-1241139.

17. Leonard M, Graham S, Bonacum D. The human factor: the critical importance of effective teamwork and communication in providing safe care. *Qual Saf Heal Care*. 2004;13(suppl 1). https://doi.org/10.1136/qshc.2004.010033.

18. Buck R, VanLear CA. Verbal and nonverbal communication: distinguishing symbolic, spontaneous, and pseudo-spontaneous nonverbal behavior. *J Commun*. 2002. https://doi.org/10.1093/joc/52.3.522.

19. Hackman JR. The design of work teams. *Handb Organ Behav*. 1987. https://doi.org/10.1016/0090-2616(78)90031-1.

20. Cooke NJ, Gorman JC, Kiekel PA. Communication as team-level cognitive processing. Macrocognition in Teams: Theories and Methodologies; 51–64. https://content.taylorfrancis.com/books/e/download?dac=C2015-0-90514-8&isbn=9781315593166&doi=10.1201/9781315593166-4&format=pdf.

21. Eisenberg EM, Murphy AG, Sutcliffe K, et al. Communication in emergency medicine: implications for patient safety. *Commun Monogr*. 2005;72:390–413. https://doi.org/10.1080/03637750500322602.

22. Coiera E, Tombs V. Communication behaviors in a hospital setting: an observational study. *BMJ*. 1998. https://doi.org/10.1136/bmj.38638.497477.68.

23. Tang PC, Jaworski MA, Fellencer CA, Kreider N, LaRosa MP, Marquardt WC. Clinician information activities in diverse ambulatory care practices. *Proc AMIA Annu Fall Symp*. 1996.

24. Bourhis RY, Roth S, MacQueen G. Communication in the hospital setting: a survey of medical and everyday language use amongst patients, nurses and doctors. *Soc Sci Med*. 1989. https://doi.org/10.1016/0277-9536(89)90035-X.

25. Lingard L, Reznick R, Espin S, Regehr G, DeVito I. Team communications in the operating room: talk patterns, sites of tension, and implications for novices. *Acad Med*. 2002. https://doi.org/10.1097/00001888-200203000-00013.

26. Weaver SJ, Rosen MA, Diazgranados D, et al. Does teamwork improve performance in the operating room? A multilevel evaluation. *Jt Comm J Qual Patient Saf*. 2010. https://doi.org/10.1016/S1553-7250(10)36022-3.

27. King HB, Battles JB, Baker DP, et al. TeamSTEPPS: team strategies and tools to enhance performance and patient safety advances. *Patient Saf New Dir Altern Approaches*. 2008. doi:NBK43686 [bookaccession].

28. Mayer CM, Cluff L, Lin W-T, et al. Evaluating efforts to optimize TeamSTEPPS implementation in surgical and pediatric intensive care units. *Jt Comm J Qual Patient Saf*. 2011. https://doi.org/10.1016/S1553-7250(11)37047-X.

29. Clapper TC, Kong M. TeamSTEPPS®: the patient safety tool that needs to be implemented. *Clin Simul Nurs*. 2012. https://doi.org/10.1016/j.ecns.2011.03.002.

30. Mercer LM, Tanabe P, Pang PS, et al. Patient perspectives on communication with the medical team: pilot study using the communication assessment tool-team (CAT-T). *Patient Educ Couns*. 2008. https://doi.org/10.1016/j.pec.2008.07.003.

31. Fairbanks RJ, Bisantz AM, Sunm M. Emergency department communication links and patterns. *Ann Emerg Med*. 2007. https://doi.org/10.1016/j.annemergmed.2007.03.005.

32. Flin R, O'Connor P, Crichton M. Safety at the sharp end. Safety at the Sharp End; doi:39/1/1 [pii].

33. Hull L, Arora S, Kassab E, Kneebone R, Sevdalis N. Observational teamwork assessment for surgery: content validation and tool refinement. *J Am Coll Surg*. 2011. https://doi.org/10.1016/j.jamcollsurg.2010.11.001.

34. Yule S, Flin R, Paterson-Brown S, Maran N, Rowley D. Development of a rating system for surgeons' non-technical skills. *Med Educ*. 2006. https://doi.org/10.1111/j.1365-2929.2006.02610.x.

35. Hull L, Arora S, Aggarwal R, Darzi A, Vincent C, Sevdalis N. The impact of nontechnical skills on technical performance in surgery: a systematic review. *J Am Coll Surg*. 2012. https://doi.org/10.1016/j.jamcollsurg.2011.10.016.

36. Downing SM. Threats to the validity of clinical teaching assessments: what about rater error? *Med Educ*. 2005. https://doi.org/10.1111/j.1365-2929.2005.02138.x.

37. Albanese MA. Challenges in using rater judgements in medical education. *J Eval Clin Pract*. 2000. https://doi.org/10.1046/j.1365-2753.2000.00253.x.

38. Albrecht G. Using subjective health assessments in practice and policy-making. Vol 4.; 1996. https://link.springer.com/content/pdf/10.1007%2FBF02249315.pdf. Accessed September 12, 2018.

39. Onnela JP, Waber BN, Pentland A, Schnorf S, Lazer D. Using sociometers to quantify social interaction

patterns. *Sci Rep*. 2014. https://doi.org/10.1038/srep05604.

40. Yu D, Blocker RC, Sir MY, et al. Intelligent emergency department: validation of sociometers to study workload. *J Med Syst*. 2016. https://doi.org/10.1007/s10916-015-0405-1.

41. Yu D, Blocker RC, Hallbeck S, Patel AM, Pasupathy KS. Wearable sociometers: calibration in chaotic simulated environments. *J Med Device*. 2015;9:020946. https://doi.org/10.1115/1.4030150.

42. Olguin O, Gloor PA, Pentland A, Olguin DO. Wearable sensors for pervasive healthcare management. 2009. doi:10.4108/ICST.PERVASIVEHEALTH2009.6033.

43. Gorman JC, Foltz PW, Kiekel PA, Martin MJ, Cooke NJ. Evaluation of latent semantic analysis-based measures of team communications content. *Proc Hum Factors Ergon Soc Annu Meet*. 2003. https://doi.org/10.1177/154193120304700337.

44. Xiao Y, Seagull FJ, Mackenzie C, Ziegert J, Klein KJ. Team communication patterns as measures of team processes: exploring the effects of task urgency and shared team experience. *Proc Hum Factors Ergon Soc Annu Meet*. 2003. https://doi.org/10.1177/154193120304701228.

45. Härgestam M, Lindkvist M, Brulin C, Jacobsson M, Hultin M. Communication in interdisciplinary teams: exploring closed-loop communication during in situ trauma team training. *BMJ Open*. 2013;3. https://doi.org/10.1136/bmjopen-2013-003525.

46. Boersma, P, Weenink D. Praat: doing phonetics by computer [Computer program]. 2018. http://www.praat.org/.

47. Cha J, Anton N, Mizota T, et al. Use of non-technical skills can predict medical student performance in acute care simulated scenarios. *Am J Surg*. 2019;217:323–328.

48. Minitab I. MINITAB statistical software. *Minitab Release*. 2000.

49. McHugh ML. Interrater reliability: the Kappa statistic. *Biochem Medica*. 2012. https://doi.org/10.11613/BM.2012.031.

50. Capella J, Smith S, Philp A, et al. Teamwork training improves the clinical care of trauma patients. *J Surg Educ*. 2010. https://doi.org/10.1016/j.jsurg.2010.06.006.

51. Kripalani S, Weiss BD. Teaching about health literacy and clear communication. *J Gen Intern Med*. 2006;21:888–890. https://doi.org/10.1111/j.1525-1497.2006.00543.x.

52. Miller K, Riley W, Davis S. Identifying key nursing and team behaviours to achieve high reliability. *J Nurs Manag*. 2009. https://doi.org/10.1111/j.1365-2834.2009.00978.x.

53. Mast MS, Hall JA, Köckner C, Choi E. Physician Gender Affects How Physician Nonverbal Behavior Is Related to Patient Satisfaction. *Medical Care*. 2008;46(12):1212–1218. https://doi.org/10.1097/mlr.0b013e31817e1877.