Correlation of attending and patient assessment of resident communication skills in the emergency department

Jason J. Lewis MD¹ | Lakshman Balaji MPH² | Anne V. Grossestreuer PhD² | Edward Ullman MD¹ | Carlo Rosen MD¹ | Nicole M. Dubosh MD¹

¹Department of Emergency Medicine, Beth Israel Deaconess Medical Center/ Harvard Affiliated Emergency Medicine Residency, Boston, Massachusetts, USA
²Department of Emergency Medicine, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA

Correspondence
Jason J. Lewis, MD, Department of Emergency Medicine, Beth Israel Deaconess Medical Center, Rosenberg Building 2, Boston, MA 02215, USA. Email: jlewis@bidmc.harvard.edu

Funding information
Funded by a Society for Academic Emergency Medicine Education Research Grant, 2018, EF2018-002.

Supervising Editor: Wendy C. Coates MD

Background: Communication and interpersonal skills are one of the Accreditation Council for Graduate Medical Education’s six core competencies. Validated methods for assessing these among trainees are lacking. Educators have developed various communication assessment tools from both the supervising attending and the patient perspectives. How these different assessment methods and tools compare with each other remains unknown. The goal of this study was to determine the degree of agreement between attending and patient assessment of resident communication skills.

Methods: This was a retrospective study of emergency medicine (EM) residents at an academic medical center. From July 2017 to June 2018, residents were assessed on communication skills during their emergency department shifts by both their supervising attending physicians and their patients. The attendings rated residents’ communication skills with patients, colleagues, and nursing/ancillary staff using a 1 to 5 Likert scale. Patients completed the modified Communication Assessment Tool (CAT), a 14-item questionnaire based on a 1 to 5 Likert scale. Mean attending and patient CAT scores were calculated for each resident. Means were divided into tertiles due to nonparametric distribution of scores. Agreement between attending and patient ratings of residents were measured using Cohen’s kappa for each attending evaluation question. Scores were weighted to assign adjacent tertiles partial agreement.

Results: During the study period, 1,097 attending evaluations and 952 patient evaluations were completed for 26 residents. Attending scores and CAT scores of the residents showed slight to fair agreement in the following three domains: patient communication (κ = 0.21), communication with colleagues (κ = 0.21), and communication with nursing/ancillary staff (κ = 0.26).

Conclusions: Attending and patient ratings of EM residents’ communication skills show slight to fair agreement. The use of different types of raters may be beneficial in fully assessing trainees’ communication skills.
INTRODUCTION

Interpersonal and communication skills are essential to the clinical practice of emergency medicine (EM). Effective communication benefits both the patient and the physician and is an integral aspect of the medical encounter. Clear and empathetic physician communication improves clinical outcomes, patient satisfaction, and adherence to treatment.1–3 Additionally, effective and patient-centered communication decreases the likelihood of a physician being named in a malpractice claim.4 Developing these skills is, therefore, an essential aspect of a residency training program and the Accreditation Council for Graduate Medical Education (ACGME) has deemed communication and interpersonal skills as one of the six core competencies.5 In EM, these skills are evaluated by residency programs and reported semiannually as part of the EM Milestones Project.6,7 Nevertheless, validated tools to assess interpersonal and communication skills in trainees are currently lacking and the most effective methods to assess these competencies remain unknown.8

Medical educators have developed a variety of modalities for assessing resident physicians’ communication and interpersonal skills. These include direct faculty observation, patient assessment of the medical encounter, simulation, and standardized patient encounters.9,10 Traditionally, residents are evaluated using supervising attending feedback based on clinical encounters.11 One patient perspective assessment instrument is the Communication Assessment Tool (CAT), a 15-item questionnaire that assesses practitioner communication skills on a 1 to 5 Likert scale.12,13 The CAT score has previously been used to successfully educate and evaluate EM, surgery, and family medicine residents.14–19 It has also been shown to be effective in assessing hospitalists’ communication skills.20

Graduate medical education is shifting its focus toward 360-degree assessment of trainees.21,22 This involves incorporating multisource feedback from supervising faculty as well as from patients, nurses, and others involved in resident training. How these various assessors compare on the domains of interpersonal and communication skills in trainees remains unknown. The goal of this study was to determine the correlation between two different assessment modalities of resident communication: supervising attending physician evaluation and patient assessment via the CAT score.

METHODS

Study design, setting, and population

This was a retrospective study of first- and second-year EM residents at an academic medical center with a 3-year EM residency program from July 2017 to June 2018. Third-year residents were excluded from the analysis due to the unique supervisory role specific to our institution’s emergency department (ED) as opposed to primary ownership of most patients. This study was determined exempt from further review by the institutional review board of Beth Israel Deaconess Medical Center. This study was funded in part from a 2018 Society for Academic Emergency Medicine Education Research Grant.

Study protocol and outcome measure

During the study period, residents were assessed on communication skills during their ED shifts by both the supervising attending physician and a convenience sample of discharged patients for whom they cared. Supervising attendings were all academic faculty in EM at our institution. Following completion of a clinical shift, the supervising attending was asked to complete an electronic summative evaluation of the residents with whom he or she worked. The evaluation form was developed by the residency program leadership based on the ACGME common requirements and the EM Milestones.5,7 Domains are rated using a 1 to 5 Likert scale. Three of the assessment domains on this electronic evaluation form pertain to residents’ communication skills with patients, physician colleagues, and consultants as well as nursing and ancillary staff (Figure 1) and apply to the initial EM Milestones of Interpersonal and Communication Skills (ICS) 1: Patient-Centered Communication and ICS 2: Team Communication.7 The attending evaluation form is part of our standard general resident evaluation process and has been in place since 2013.

In addition, we implemented a system to collect patient assessments of the residents’ communication skills. Trained research assistants (RAs) received an electronic page whenever a first- or second-year EM resident signed up for a patient through our ED’s online tracking system between the hours of 8 a.m. to 11 p.m.

| Interpersonal and Communication Skills | Outstanding (5) | Exceeds Expected Level (4) | At Expected Level (3) | Below Expected Level (2) | Unsatisfactory (1) | N/A |
|---------------------------------------|-----------------|---------------------------|----------------------|-------------------------|-------------------|-----|
| Demonstrates effective patient centered communication |                 |                           |                      |                         |                   |     |
| Communication with other physician colleagues and consultants |                 |                           |                      |                         |                   |     |
| Communication with nursing and ancillary staff |                 |                           |                      |                         |                   |     |

**FIGURE 1** Attending evaluation form for resident communication
Monday through Friday. Patients were included if they could identify the resident who cared for them by photo; did not require interpreter services; and were at baseline alert and oriented to person, place, and time. Additionally, CAT surveys were administered to discharged patients only in accordance with our institution’s patient survey policy. If eligible, the RA invited the patient to complete the CAT to provide valuable feedback on the resident physician’s communication skills. The CAT, as described above, includes 15 questions based on a 1 to 5 Likert scale with 14 physician-specific questions for a total out of 70 possible points (Figure 2). Our modified CAT included these 14 questions to obtain information specific to the individual resident and not the entire health care team. By completing the survey, patients gave consent for use of deidentified data for research purposes. Both attending physicians and patients filling out their respective assessments were made aware that their responses remain anonymous to the resident. Mean attending ratings for each resident across all three domains and mean CAT scores using the 14 physician-specific questions were calculated for each resident and compared for agreement.

Data analysis

Mean attending ratings and CAT scores were divided into tertiles due to nonparametric distribution of scores. Agreement between attending ratings and CAT scores of residents were measured using Cohen’s kappa for each attending evaluation question. We used weighted scores to assign credit to partially concordant tertiles. Tertiles that were in complete agreement were assigned a weight of 1, tertiles that was in partial agreement were assigned a weight of 0.5, and tertiles that were completely discordant were assigned a weight of 0. We a priori decided that a kappa less than 0 indicated poor agreement, a kappa between 0 and 0.2 represented slight agreement, a kappa between 0.2 and 0.4 represented fair agreement, a kappa between 0.4 and 0.6 represented moderate agreement, a kappa between 0.6 and 0.8 represented substantial agreement, and a kappa between 0.8 and 1 represented almost perfect agreement. Given that the scale does have overlap between the categories, we chose the more conservative approach of choosing the lower agreement category.

RESULTS

Twenty-six residents were evaluated during the study period. Thirteen of them were postgraduate year (PGY)-1 residents and 13 were PGY-2 residents. Seven of the 26 residents were female. A total of 1,097 supervising attending evaluations with a median number of 44 evaluations per resident (interquartile range [IQR] = 33–50, minimum = 22) and 952 CAT questionnaires with a median number of 37 evaluations per resident (IQR = 36–37, minimum = 33) were completed for the 26 residents. Two responses

![Figure 2](https://example.com/f2.png)

**Figure 2.** Communication Assessment Tool. From “Measuring patient views of physician communication skills: development and testing of the Communication Assessment Tool,” by G. Makoul, E. Krupat and CH. Chang, Patient Education and Counseling, 2007, 67(3), p. 341. Copyright 2007 by Gregory Makoul, PhD. Reprinted with permission from Gregory Makoul, PhD. The 14 doctor specific questions were used in analysis. Question 15 was not included.
**TABLE 1** Resident CAT scores by question and PGY

| CAT question                                                                 | All residents (n = 26) | PGY-1 (n = 13) | PGY-2 (n = 13) | p-value |
|-----------------------------------------------------------------------------|------------------------|----------------|----------------|---------|
| 1. Greeted me in a way that made me feel comfortable                        | 4.8 (±0.1)             | 4.8 (±0.1)     | 4.8 (±0.1)     | 0.642   |
| 2. Treated me with respect                                                  | 4.9 (±0.1)             | 4.9 (±0.1)     | 4.9 (±0.1)     | 0.819   |
| 3. Showed interest in my ideas about my health                             | 4.8 (±0.1)             | 4.8 (±0.1)     | 4.8 (±0.2)     | 0.407   |
| 4. Understood my main health concerns                                       | 4.8 (±0.1)             | 4.8 (±0.1)     | 4.8 (±0.1)     | 0.518   |
| 5. Paid attention to me                                                     | 4.8 (±0.1)             | 4.9 (±0.1)     | 4.8 (±0.2)     | 0.470   |
| 6. Let me talk without interruptions                                         | 4.9 (±0.1)             | 4.9 (±0.1)     | 4.9 (±0.1)     | 0.504   |
| 7. Gave me as much information as I wanted                                  | 4.8 (±0.2)             | 4.8 (±0.1)     | 4.7 (±0.2)     | 0.366   |
| 8. Talked in terms I could understand                                       | 4.9 (±0.1)             | 4.9 (±0.1)     | 4.9 (±0.1)     | 0.727   |
| 9. Checked to be sure I understood everything                               | 4.8 (±0.1)             | 4.8 (±0.1)     | 4.8 (±0.2)     | 0.953   |
| 10. Encouraged me to ask questions                                           | 4.6 (±0.2)             | 4.7 (±0.2)     | 4.5 (±0.3)     | 0.160   |
| 11. Involved me in decisions as much as I wanted                            | 4.7 (±0.2)             | 4.7 (±0.2)     | 4.7 (±0.3)     | 0.767   |
| 12. Discussed next steps, including any follow-up                           | 4.7 (±0.2)             | 4.8 (±0.1)     | 4.7 (±0.3)     | 0.458   |
| 13. Showed care and concern                                                 | 4.8 (±0.2)             | 4.8 (±0.1)     | 4.8 (±0.2)     | 0.621   |
| 14. Spent the right amount of time with me                                  | 4.7 (±0.2)             | 4.8 (±0.2)     | 4.7 (±0.2)     | 0.620   |

Note: Data are reported as mean (±SD).
Abbreviations: CAT, Communication Assessment Tool; PGY, postgraduate year.
TABLE 2  Tertile ranges for mean attending and CAT scores

| Range of responses by tertile                                      | Tertile 1         | Tertile 2         | Tertile 3         |
|-------------------------------------------------------------------|-------------------|-------------------|-------------------|
| Overall CAT                                                      | 60.8–67.2         | 67.3–68.4         | 68.5–69.2         |
| Demonstrates effective patient centered communication            | 3.9–4.1           | 4.2–4.4           | 4.4–4.6           |
| Communication with other physician colleagues and consultants    | 3.5–4.1           | 4.2–4.4           | 4.4–4.7           |
| Communication with nursing and ancillary staff                   | 3.6–4.1           | 4.1–4.3           | 4.4–4.6           |

Range of mean CAT scores by tertile of attending response

| Mean CAT score associated with                                       | Tertile 1         | Tertile 2         | Tertile 3         |
|---------------------------------------------------------------------|-------------------|-------------------|-------------------|
| demonstrates effective patient centered communication              | 64.9–68.3         | 65.5–69.2         | 60.8–68.9         |
| communication with other physician colleagues and consultants      | 64.9–67.7         | 65.6–69.2         | 60.8–68.9         |
| communication with nursing and ancillary staff                     | 64.9–68.3         | 60.8–69.2         | 67.2–69.0         |

Abbreviation: CAT, Communication Assessment Tool.

were missing for the domain of "communication with other physician colleagues and consultants" and one response was missing for "communication with nursing and ancillary staff" from the attending evaluations. Mean evaluation scores for all residents are provided in Figure 3. There was no difference between mean attending scores for PGY-1 and PGY-2. Mean scores for the individual CAT questions are displayed in Table 1. Patients rated the residents most highly on the CAT question of "treated me with respect," "let me talk without interruptions," and "talked in terms I could understand." They rated them the lowest on "encouraged me to ask questions." There was no difference between PGY-1 and PGY-2 across all CAT questions (Table 1). Tertile mean ranges for each scoring system are presented in Table 2. The range of mean CAT score associated with each tertile of attending evaluation question are also shown in Table 2. Attending scores and CAT scores of the residents showed slight to fair agreement in the following three domains: patient communication (κ = 0.21, expected agreement = 56%, actual = 65%), communication with colleagues (κ = 0.21, expected agreement = 56%, actual = 65%), and communication with nursing/ancillary staff (κ = 0.26, expected agreement = 55%, actual = 67%).

DISCUSSION

The results of our study demonstrate statistically fair agreement between supervising attending physician and patient assessment of several different communication domains in EM residents. Given that the kappas of are just above the statistical threshold for fair, the clinical significance of this is closer to slight to fair agreement. To the best of our knowledge, this is the first study to investigate degree of concordance in these two different assessor groups in EM residents and our results highlight the utility of using a multisource feedback approach to assessing resident communication and interpersonal skills. Even among faculty raters, prior work has shown marked variability and even poor agreement of assessing resident's skills.24,25 Thus, the slight to fair agreement we found invites further research for educators and residency programs seeking to include patient feedback in resident evaluations in that it may provide additional critical information that may be missed by utilizing only attending evaluation methods.

The faculty attending and patient perspectives on residents' abilities are inherently different and most certainly account for lack of "strong" agreement. For one, these two groups typically spend different amounts of time in the actual bedside interaction. While attending level direct observation has been noted to be a valid and well-received method of assessment,9,11,26 the reality of clinical practice is that the time a supervising attending physician can devote to the direct observation of residents communicating with patients at the bedside is often limited. The patient, moreover, is involved in the entire interaction being evaluated. Previous work has shown that overall direct observation time in ED was less than 4% of total shift time.27 Interestingly, the highest concordance was found between attending evaluation of communication with nursing staff and the CAT score. This may have been influenced by direct feedback provided by nurses to the attending physician rather than first-hand observation by the attending physician. Alternatively, it is possible that residents communicate differently with individuals not formally evaluating them, which may provide a more genuine representation of their true communication skills. This finding would be worth investigating further in a future study. Nevertheless, while supervising attending physician evaluation has historically been the primary mode of assessing resident interpersonal and communication skills, given the
limitations of direct observation time, utilizing real-time patient assessors may be beneficial in fully evaluating trainees' skills.

In contrast to the attending, the patient has the opportunity to assess the resident’s communication through a different lens: as the direct recipient rather than as an observer. Not only may this lead to a more encompassing period of observation compared to a single direct observation session by a supervisor but it also allows for the patient's own emotional experience of the health care interaction to enhance or otherwise influence their assessment of the resident. Prior work has demonstrated that patient feedback collected through a validated and credible method can have a positive impact on medical performance, although feedback needs to be provided to residents in a manner that facilitates discussion and encourages an actionable behavior change if needed. There is increasing data supporting the fact that multisource feedback is essential in assessing all resident skills and can improve professional practice.

Our findings highlight the potential benefit of using a multimodal approach to evaluating resident interpersonal and communication skills as relying solely on one method may miss certain deficiencies that could subsequently be addressed. We feel that program directors should use the CAT or other similar patient assessment tools combined with direct feedback from supervising attendings on any potential communication or interpersonal skills concerns as it may be beneficial for resident education.

Prior studies have sought to correlate different raters for trainees in various clinical settings. While there has been some positive correlation of standardized patients and faculty evaluations in radiology and physical medicine and rehabilitation residents, other studies in different medical specialties have shown inconsistent correlation between different raters for interpersonal skills. While the majority of these studies have used standardized patients, our study used assessments from actual patients in a real-time clinical setting. In doing so, we sought to eliminate factors that may inherently skew evaluations in a simulation lab, objective structured clinical evaluation, or other nonclinical setting while including the natural interruptions that arise during real encounters in the ED. Further studies are needed to determine how real patient feedback compares to standardized patients and other raters, as well as the validity of existing tools. However, based on our findings of only slight to fair agreement between attending and patient evaluation methods, we suggest utilizing a multimodal approach incorporating patient evaluation in residency assessment.

**LIMITATIONS**

This study has several limitations. We used two different assessment tools with different questions, which may limit comparison, particularly since the CAT score does not specifically ask about “communication with physician colleagues and consultants” or “nursing and ancillary staff.” This was done because of the existing faculty evaluation forms that were already in place based on the need for EM Milestone-based assessment and integration of a patient-centered assessment tool. However, given the large sample size and comparison of means, we can certainly appreciate trends in our analysis. Additionally, given institutional restrictions on surveying hospitalized patients, only patients who were discharged completed CAT questionnaires. This skews our CAT results to lower acuity, and therefore we cannot draw conclusions about resident communication in the higher-acuity patient population. The gender imbalance of our study population may bias the results; however, previous studies have shown no difference in patient perspective of communication skills between male and female residents of the same year. Furthermore, there are no anchors for the attending evaluation Likert scale and it is possible that individual attendings interpret each individual score differently. Additionally, while it is recommended that attendings complete resident evaluations after the shift, it is unknown how long after a shift an evaluation is completed and therefore may be subject to recall bias. Both methods are also subjective assessments and thus prone to inherent bias, but given the large number of completed evaluations from multiple different faculty and patients, these effects should be limited. Finally, the kappa score has overlap on ranges and may be open to interpretation. We chose the more conservative approach of assigning the lower agreement category and further supported the concordance with expected and actual agreement as previously suggested to help clarify inter-rater reliability.

**CONCLUSIONS**

We found slight to fair agreement between supervising attending and patient ratings of emergency medicine residents’ communication and interpersonal skills. The use of a multimodal approach using different types of raters, including both attending evaluation and patient evaluation, may be beneficial in assessing trainees’ communication and interpersonal skills. Based on our findings, we recommend incorporating the Communication Assessment Tool score in resident evaluation. Further studies are needed to investigate the validity of attending physician and patient evaluation modalities.

**CONFLICT OF INTEREST**

JL, LB, AG, EU, CR, and ND report grant money to Beth Israel Deaconess Medical Center to conduct research conceived and written by Nicole Dubosh from Society for Academic Emergency Medicine Education Research Grant, 2018.

**AUTHOR CONTRIBUTION**

Jason J. Lewis contributed to study concept and design, acquisition of the data, drafting of the manuscript, critical revision of the manuscript for important intellectual content, and study supervision. Lakshman Balaji contributed to analysis and interpretation of the data, critical revision of the manuscript for important intellectual content, and statistical expertise. Anne V. Grossestreuer contributed...
to analysis and interpretation of the data, critical revision of the manuscript for important intellectual content, and statistical expertise. Edward Ullman contributed to study concept and design and critical revision of the manuscript for important intellectual content. Carlo Rosen contributed to study concept and design, and critical revision of the manuscript for important intellectual content. Nicole M. Dubosh contributed to study concept and design, acquisition of the data, drafting of the manuscript, critical revision of the manuscript for important intellectual content, acquisition of funding, and study supervision.

ORCID
Jason J. Lewis  https://orcid.org/0000-0001-5421-407X
Nicole M. Dubosh  https://orcid.org/0000-0001-8674-5334

REFERENCES
1. Hojat M, Louis DZ, Markham FW, et al. Physicians’ empathy and clinical outcomes for diabetic patients. Acad Med. 2011;86(3):359-364.
2. Boissy A, Windover AK, Bokar D, et al. Communication skills training for physicians improves patient satisfaction. J Gen Intern Med. 2016;31(7):755-761.
3. Anhang Price R, Elliott MN, Zaslavsky AM, et al. Examining the role of patient experience surveys in measuring health care quality. Med Care Res Rev. 2014;71(5):522-554.
4. Levinson W, Roter DL, Mullooly JP, Dull VT, Frankel RM. Physician–patient communication: the relationship with malpractice claims among primary care physicians and surgeons. JAMA. 1997;277(7):553-559.
5. ACGME Program Requirements for Graduate Medical Education in Emergency Medicine. Accreditation Council for Graduate Medical Education website. Effective July 1, 2019. https://www.acgme.org (Accessed March 9, 2020).
6. Beeson MS, Carter WA, Christopher TA, et al. The development of the emergency medicine milestones. Acad Emerg Med. 2013;20(7):724-729.
7. Beeson MS, Carter WA, Christopher TA, et al. Emergency medicine milestones. J Grad Med Educ. 2013;5(1):5-13.
8. Chan TM, Wallner C, Swoboda TK, Leone KA, Kessler C. Assessing interpersonal and communication skills in emergency medicine. Acad Emerg Med. 2012;19(12):1390-1402.
9. Hobgood CD, Riviello RJ, Jouriles N, Hamilton G. Assessment of communication and interpersonal skills competencies. Acad Emerg Med. 2002;9(11):1257-1269.
10. Dittom-Phare P, Sandhu H, Kelly B, Kissane D, Loughland C. Pilot evaluation of a communication skills training program for psychiatry residents using standardized patient assessment. Acad Psychiatry. 2016;40(5):768-775.
11. Craig S. Direct observation of clinical practice in emergency medicine education. Acad Emerg Med. 2011;18(1):60-67.
12. Makoul G, Krupat E, Chang CH. Measuring patient views of physician communication skills: development and testing of the Communication Assessment Tool. Patient Educ Couns. 2007;67(3):333-342.
13. 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;73(2):220-223.
14. Dubosh NM, Hall MM, Novack V, Shafat T, Shapiro NI, Ullman EA. A multimodal curriculum with patient feedback to improve medical student communication: pilot study. West J Emerg Med. 2020;21(1):115.
34. Nadeem N, Zafar AM, Ahmad MN, Zuberi RW. Faculty and patient evaluations of radiology residents’ communication and interpersonal skills. *J Pak Med Assoc*. 2012;62(9):915.

35. Udawatta M, Alkhalid Y, Nguyen T, et al. Patient satisfaction ratings of male and female residents across subspecialties. *Neurosurgery*. 2020;86(5):697-704.

36. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276-282.

How to cite this article: Lewis JJ, Balaji L, Grossestreuer AV, Ullman E, Rosen C, Dubosh NM. Correlation of attending and patient assessment of resident communication skills in the emergency department. *AEM Educ Train*. 2021;5:e10629. https://doi.org/10.1002/aet2.10629