

Development of the Objective, Structured Communication Assessment of Residents (OSCAR) Tool for Measuring Communication Skills With Patients

ALEECE CARON, PhD
ADAM PERZYNSKI, PhD
CHARLES THOMAS, MS
JIMMY Y. SAADE, MD
MICHAEL MCFARLANE, MD
JEFFERY BECKER, MD

Abstract

Background Although interpersonal and communication skills are essential to physician practice, there is a dearth of effective tools to meaningfully teach and assess communication skills.

Objective The purpose of our study was to create a standardized tool for evaluation of communication skills for residents across specialties.

Methods We designed an Objective, Structured Communication Assessment of Residents (OSCAR) tool, consisting of 4 clinical stations, to assess intern communication skills with relationship development, their establishment of case goals, and their organization and time management skills. Interns from 11 training programs completed the stations, with senior residents trained to function as standardized patients. The 4 stations' scenarios were a disruptive patient, handling a phone call for a narcotics refill, disclosing a medical mistake, and delivering bad news.

Results Eighty-three interns completed OSCAR during orientation. The assessment took interns about 40 minutes to complete, and participants were given immediate feedback by the standardized patients. The total possible score for each station was 50. Resident performance was highest for disclosing a medical error (94%, 47 of 50), followed by handling a disruptive patient (90%, 45 of 50), disclosing bad news (86%, 43 of 50), and handling the phone call for the narcotics refill (62%, 31 of 50). Multivariate analysis of variance results indicated differences between residents from US and international medical schools, but there were no significant differences across specialties. Interrater reliability was excellent for each station (> 0.80).

Conclusions OSCAR is a practical tool for assessing interns' communication skills to provide timely results to program directors.

All authors are with the MetroHealth Medical Center, Case Western Reserve University, Cleveland. **Aleece Caron, PhD**, is Assistant Professor of Medicine and Senior Medical Educator; **Adam Perzynski, PhD**, is Senior Instructor of Medicine in the Center for Healthcare Research and Policy; **Charles Thomas, MS**, is Biostatistician in the Center for Healthcare Research and Policy; **Jimmy Y. Saade, MD**, is Radiology Resident; **Michael McFarlane, MD**, is Professor of Medicine, Vice Chair of the Department of Medicine, and Program Director of the Division of Internal Medicine; and **Jeffery Becker, MD**, is Assistant Professor of Medicine and Senior Associate Program Director of the Division of Internal Medicine.

Funding: The authors report no external funding source for this study.

The authors would like to acknowledge the following manuscript reviewers: Duncan Neuhauser, PhD, the Charles Elton Blanchard, MD Professor of Health Management, Professor of Medicine, Professor of Family Medicine, Professor of Organizational Behavior, and Co-Director, Health Systems Management Center, Case Western Reserve University; Neal V. Dawson, MD, Professor of Medicine, the MetroHealth System, Case Western Reserve University School of Medicine; and Ashwini Sehgal, MD, Professor of Medicine, the MetroHealth System, Director, Center for Reducing Healthcare Disparities, and the Duncan Neuhauser Professor of Community Health Improvement, Case Western Reserve University School of Medicine.

The authors would also like to acknowledge Marcie Becker, MBA, Director of International Affairs and Graduate Medical Education (GME) at the MetroHealth System, and Christine Redovan, GME Consultant Partners in Medical Education Inc, for their contributions to this project.

Corresponding author: Aleece Caron, PhD, 2500 MetroHealth Drive R-256, Cleveland, OH 44109, 216.778.5834, acaron@metrohealth.org

Editor's Note: The online version of this article contains reliability statistics, the evaluation form, and an example of the 4 stations used in this study.

Introduction

Interpersonal and communication skills are essential to physician practice. Designing tools to meaningfully teach and assess communication skills is, however, difficult and effective tools are largely lacking. Residents come from a variety of cultural and educational backgrounds that shape their ability to communicate with patients and colleagues. Providing residents with an opportunity to practice their interpersonal and communication skills in a safe environment, along with immediate constructive feedback, would assist in closing this gap in their training.¹

Received August 10, 2012; revisions received January 9, 2013, and February 22, 2013; accepted February 23, 2013.

DOI: <http://dx.doi.org/10.4300/JGME-D-12-00230.1>

Direct observation of residents is the most reliable and valid means of skills assessment, but has limitations, including variation in the patients and circumstances residents are exposed to, and in the observers (ie, attending physicians or senior residents).^{2–12} Several articles describe appropriate protocols for an objective, structured clinical examination (OSCE) and for standardized patient (SP) evaluations, yet the financial and opportunity costs of those modalities appear to be major limiting factors.^{13–15} Although program directors agree that communication competency is important for residents to master, there has been little agreement in the literature about how those communication skills should be evaluated.¹⁶

The purpose of our study was to create a practical, useful, and expedient evaluation of communication skills for residents, across specialties; to determine whether additional skill development was necessary; and to provide timely feedback to the program directors regarding these skills. We designed a program that used SPs to provide meaningful data to program directors and immediate feedback to the residents.¹⁷

Methods

Setting and Participants

The Objective, Structured Communication Assessment of Residents (OSCAR) assessment took place at the MetroHealth System, a 731-bed facility that also serves as the county's safety-net hospital and is affiliated with Case Western Reserve School of Medicine. Eight residents from a variety of disciplines were trained as SPs and evaluators by an experienced preceptor. Interns from 11 residency programs participated in OSCAR.

Intervention

An interdisciplinary committee designed the OSCAR clinical stations and evaluation tools in 3 phases: Phase I, station development (January–March 2010); Phase II, a feasibility study (end of March 2010); and Phase III, implementation and evaluation by interns during the residency orientation in July 2010. Stations were designed to simulate experiences that incoming residents in most disciplines would encounter during their training^{18–22} and were patient-safety priorities.^{23–28}

MetroHealth Medical Center's Institutional Review Board determined this study was exempt.

Phase I resulted in the development of OSCAR, which consisted of 4 stations: (1) interacting with a disruptive patient who wants to leave the hospital against medical advice, (2) returning a patient's phone call for a narcotics refill, (3) disclosing that a patient received the wrong medication, and (4) informing a patient about an abnormal

What was known

Interpersonal and communication skills are essential, yet there is a dearth of effective tools to teach and assess communication skills.

What is new

A tool providing new interns with an objective, structured assessment of their communication skills, using senior residents as standardized patients.

Limitations

Small sample size limits generalizability. The evaluation tool requires further validation.

Bottom line

The use of senior residents as standardized patients allows for immediate feedback to new interns about their communication skills and provides meaningful feedback to program directors.

mammogram and discussing further assessment. For each station, the resident was provided with tasks to complete, patient demographics, reason for the encounter, and a description of the situation. We adapted the format of OSCE and its use of SPs to focus on interpersonal and communication skills with patients. Clinical content was included in the scenarios, but residents were not evaluated on their medical knowledge.

Residents were assigned to stations so they could become proficient in 1 station and were not moved between stations. Training took approximately 2 hours and included an explanation of the project, review of the stations, demonstration by the preceptor, opportunity for role playing, and evaluation of their SP performance by the preceptor. It also included instruction, practice, and evaluation on how to give feedback. Faculty preceptors were available for questions and informally observed the process. Interns had 2 minutes to read the case and objectives, 5 minutes to interview the patient, 2 minutes to receive immediate feedback from the SP, and 1 minute to change stations.

In Phase II, we conducted a feasibility study with 18 internal medicine residents and designed an evaluation form for the SPs to complete on each resident they observed (described above). The evaluation form was piloted using 18 internal medicine residents and 8 SPs. It took approximately half a day to run the feasibility study. Two administrators ran OSCAR and were trained on how to instruct the residents to rotate through the stations, with one acting as timekeeper. We used structured feedback from residents and SPs in the pilot study to refine the scenarios. Changes included using actual clinical examination rooms and changing the response choices on the evaluation form. This pilot showed that we could conduct OSCAR efficiently and that the SPs could complete the evaluations in a timely manner. The committee trained

TABLE 1 INTERN DEMOGRAPHICS

Demographics	Respondents, No. (%)
Sex	
Female	41 (49)
Male	42 (51)
Total	83 (100)
US Born	
Yes	46 (55)
No	37 (45)
Total	83 (100)
International medical school graduate	
Yes	35 (42)
No	48 (58)
Total	83 (100)
Program	
Internal medicine	30 (36)
Emergency medicine	13 (16)
Obstetrics-gynecology	7 (8)
Family Practice	6 (7)
Medicine-pediatrics	6 (7)
Pediatrics	5 (6)
Dentistry	5 (6)
Psychiatry	4 (5)
Physical medicine and rehabilitation	3 (4)
Pathology	2 (2)
Radiology	2 (2)
Total	83 (100)
Ethnic origin	
Non-Hispanic White	30 (36)
Indian subcontinent	24 (29)
Asian/Pacific Islander	8 (10)
Other	21 (25)
Total	83 (100)

senior residents to act as SPs. The committee felt that allowing residents to give feedback to interns would be a valuable learning experience, in addition to developing a sustainable cadre of experienced SPs. Feedback from the residents suggested that the experience was positive overall.

Outcome Measures

To evaluate resident performance, the committee agreed on and developed 3 domains critical to patient care and effective communication for the stations described above: relationship development, case goals, and organization and time management. Based on our experience in the pilot, we anticipated high performances overall, and we selected response categories optimized to distinguish scenario performance on a 4-point scale (outstanding, advanced, proficient, and unsatisfactory) with a total possible score of 50.

Residents were reevaluated at the end of their postgraduate year (PGY)-1, using the same stations and the same SPs. There was no additional formal training for the OSCAR scenarios during the year because we wanted to see if resident performance improved after a year of training. Interns completed the stations in a random order each time.

Analysis

Contrast analysis was used to examine general differences in resident performance across stations. To examine differences according to the sex of the resident, program (primary care versus specialty programs), and medical school type, we used multivariate analysis of variance (MANOVA). To examine changes from the beginning to the end of PGY-1, we used general linear modeling to conduct a repeated measures (RM)-MANOVA.²⁹ The Bonferroni procedure was used to account for the multiple comparisons. Contrasts and pairwise comparisons were then examined. We used multiple imputations to handle missing values. Five imputed data sets were created using the fully conditional Markov chain Monte Carlo procedure. Results of pooled estimates across the 5 imputations are presented. Most variables had complete or near-complete data. Examination of residuals indicated multivariate normality. Data analysis was conducted using SPSS version 19.0 (IBM Corp, Armonk, NY).

Interrater reliability was assessed by having the SP and an observer faculty rater independently assess each intern. The SPs and the faculty rater were instructed not to speak to each other about that intern's performance.

Results

Eighty-three of 91 interns (91%) from 11 training programs (TABLE 1) completed the OSCAR process. Forty-one (49%) of the interns were women, 48 (58%) were US medical school graduates, and 46 (55%) were born in the United States. Of those born in the United States, 6 (13%) attended medical school outside the United States. Thirty-six percent (30 of 83) of the interns were non-Hispanic White, 29% (24 of 83) were from the Indian subcontinent, and 35% (29 of 83) had other ethnic origins. Twenty-seven of 46 interns (59%) born in the United States were

TABLE 2 OVERALL RESULTS (n = 83)

Scenario	Mean ^a , %	Mean (SD)	Range	Total Items	Total Possible
Medical error	95	47.6 (3.2)	30.0–50.0	13	50.0
Relationship development	96	21.2 (1.4)	14.0–22.0	6	22.0
Case goals	93	14.9 (1.7)	9.0–16.0	4	16.0
Organization and time management	94	11.3 (1.2)	7.0–12.0	3	12.0
Disruptive patient	90	44.8 (6.1)	24.0–50.0	13	50.0
Relationship development	90	19.8 (2.1)	12.0–22.0	6	22.0
Case goals	88	14.0 (2.8)	6.0–16.0	4	16.0
Organization and time management	91	10.9 (1.8)	5.0–12.0	3	12.0
Bad news	86	42.2 (6.9)	18.0–49.0	13	50.0
Relationship development	87	19.1 (2.5)	10.0–21.0	6	22.0
Case goals	84	13.4 (3.0)	5.0–16.0	4	16.0
Organization and time management	88	10.5 (2.0)	3.0–12.0	3	12.0
Narcotics refill	62	31.2 (4.9)	21.0–40.0	13	50.0
Relationship development	63	13.9 (2.1)	10.0–20.0	6	22.0
Case goals	60	9.6 (2.2)	4.0–14.0	4	16.0
Organization and time management	63	7.6 (1.4)	5.0–10.0	3	12.0

^a Means as a percentage of total possible score.

non-Hispanic White, 10 of 46 (22%) were of Indian descent, and the remainder (9 of 46; 20%) were from other racial/ethnic backgrounds.

It took each resident approximately 40 minutes to complete all 4 stations, and it took 3 one-half day sessions for all interns to complete the OSCAR process. Interrater reliability for each station was excellent (Cohen $\kappa > 0.80$) and was also excellent for each domain within the stations (reliability statistics and examples of stations are provided as online supplemental material). Interitem reliability (Cronbach α) ranged from acceptable (0.66) to excellent (0.94).²⁸

The primary outcome measure was overall station performance. The percentage of total possible score, overall mean, and standard deviation are illustrated in TABLE 2. Resident performance was highest (percentage of total possible score; mean; SD) on disclosing medical error (95%; 47.6; 3.2), followed by handling a disruptive patient (90%; 44.8; 6.1), delivering bad news (84%; 42.2; 6.9), and handling a phone call for a narcotics refill (62%; 31.2; 4.9). Across the 4 stations, average performance was slightly lower (~3% across stations) on meeting case goals.

There was a significant contrast in performance across the stations ($F_{4,50} = 2734.68$, $P < .001$), suggesting residents' skills differed from station to station. Among the 83 residents, the MANOVA results indicated a significant

between-subjects effect of medical-school type ($F_{4,72} = 6.079$, $P < .001$) but not of program type ($F_{4,72} = 1.889$, $P = .12$) or sex ($F_{4,72} = 1.379$, $P = .25$). The multivariate test for within-subjects differences from beginning to the end of PGY-1 was highly significant ($F_{4,72} = 26.753$, $P < .001$), suggesting a general improvement in resident performance.

Results of Bonferroni-adjusted pairwise comparisons between estimated marginal means are shown in TABLE 3. Those comparisons showed that US medical school graduates ($n = 48$) performed better than international medical graduates did ($n = 35$) on all 4 stations; women performed slightly better on the disruptive patient station; and residents in primary care programs performed better on the disclosing bad news station.

Of the 83 residents in the initial application of OSCAR, 68 (82%) were evaluated at the end of PGY-1. Comparing performance at the end of PGY-1 with the beginning of the year, Bonferroni-adjusted pairwise comparisons found significant within-subjects improvement in 2 of 4 stations (narcotics refill and disruptive patient).

Discussion

We show the practicality of an SP-based assessment of interns' communication skills and demonstrate that we could

TABLE 3 COMPARISON OF RESIDENT PERFORMANCE ON THE OBJECTIVE, STRUCTURED COMMUNICATION ASSESSMENT OF RESIDENTS (OSCAR) ASSESSMENT ^a								
OSCAR Station	Program Type, ^b %		Sex, ^b %		Medical School Type, ^b %		Time Point, ^c %, n = 68	
	Primary (n = 54)	Specialty (n = 29)	Male (n = 42)	Female (n = 41)	United States (n = 48)	International (n = 35)	Begin PGY-1	End PGY-1
Medication error	97	95	97	94	99	92	94	98
Bad news	89	83	86	86	91	81	84	88
Narcotics refill	71	71	71	71	73	68	62	79
Disruptive patient	93	90	89	94	94	89	89	93

Abbreviation: PGY, postgraduate year.
^a Higher scores indicate better performance. Statistically significant differences ($P < .05$) after Bonferroni adjustment are shown in bold.
^b Estimates produced by multivariate analysis of variance (MANOVA) contrast analysis.
^c Estimates produced by repeated measures MANOVA contrast analysis.

implement and evaluate communication skills using a practical tool for all incoming interns, regardless of specialty. We also report that this did not require an unreasonable time commitment from senior residents and faculty.

The interns' mean performance on the stations ranged from an average score of 62% to 95% of the total possible points. Our study was not designed to determine a threshold level for competence, and there was variation among stations for all participants. The narcotics refill results showed a mean performance across all groups of 62% at baseline and 79% at follow-up, suggesting that despite improvement, additional training for this aspect of clinical and communication practice may be necessary.

Our results show differences in performance for US medical school graduates and international medical graduates. The US medical school graduates performed better overall, which we do not attribute to problems with their prior education but to cultural issues. More studies need to be conducted to determine whether this trend extends beyond our institution and whether it is a training problem or an experience problem.

We opted to train residents as SPs for several reasons: (1) hiring SPs is costly, (2) we wanted to provide senior residents with a unique teaching and learning opportunity, and (3) we wanted to determine whether using senior residents as SPs was something we would want to do in the future to assess resident performance.

Although the OSCAR approach is similar to an OSCE, there are differences. The OSCAR tool was specifically designed to evaluate skills in communicating with patients, not medical knowledge or clinical skills. The SPs were trained to not correct issues of medical knowledge, and the residents were told that the focus was solely on communication skills. In addition, because of its limited focus, we were able to tailor the training for the SPs to allow them to

complete the evaluation; give appropriate, meaningful feedback; and gain teaching experience from this process.

The OSCAR tool showed a similar level of reliability as other techniques used to evaluate communication skills, such as the 360-degree evaluations, with some added advantages. Ratings on 360-degree evaluations are often unidimensional in that raters give an overall impression rather than differentiate areas of specific competence.^{30,31} With 360-degree evaluations, there is often a substantial burden in collecting, managing, and reporting data collected from the comprehensive set of persons interacting with residents.³⁰ Individuals evaluating residents using 360-degree surveys may not have much time to actually observe the resident's communication skills in challenging situations.³⁰ Additionally, feedback to residents and program directors from 360-degree evaluations is often delayed, and our approach to assessing communication skills provided for real-time feedback.

Once we have accumulated a larger sample, we plan to further evaluate the construct validity of the OSCAR stations and to create a tool for SPs to reflect on their experience in evaluating the interns. We would also like to evaluate the SPs on their ability to give constructive feedback.

Our study has some limitations. The small sample size limits generalizability. Another limiting factor could be the design of the instrument because ratings were not anchored to specific behaviors. Finally, the baseline scores were high, suggesting a possible ceiling effect and that the stations were not challenging.

Conclusion

The OSCAR tool provides program directors with an inexpensive and feasible method of assessing trainees' skill in communicating with patients at the beginning of training. OSCAR had high interrater reliability when used

with interns from a variety of specialties. Residents reported that this was an excellent learning experience, and the program was able to consistently identify residents who might have communication issues, to intervene quickly, and to implement an improvement plan.

References

- 1 Makoul G. Essential elements of communication in medical encounters: the Kalamazoo consensus statement. *Acad Med*. 2001;76(4):390–393.
- 2 Silverman J, Kurtz S, Draper J. *Skills for Communicating With Patients*. Oxon, England: Radcliffe Medical Press; 1998.
- 3 Maguire P, Pitceathly C. Key communication skills and how to acquire them. *BMJ*. 2002;325(7366):697–700.
- 4 Duffy FD, Gordon GH, Whelan G, Cole-Kelly K, Frankel R, Buffone N, et al. Participants in the American Academy on Physician and Patient's Conference on Education and Evaluation of Competence in Communication and Interpersonal Skills. Assessing competence in communication and interpersonal skills: the Kalamazoo II report. *Acad Med*. 2004;79(6):495–507.
- 5 Holmboe E, Hawkins RE, Huot SJ. Effects of training in direct observation of medical residents' clinical competence: a randomized trial. *Ann Intern Med*. 2004;140(11):874–881.
- 6 Epstein RM. Assessment in medical education. *N Engl J Med*. 2007;356(4):387–396.
- 7 Schuh LA, London Z, Neel R, Brock C, Kissela BM, Schultz L, et al. Education research: bias and poor interrater reliability in evaluating the neurology clinical skills examination. *Neurology*. 2009;73(11):904–908.
- 8 Hallock JA, Seeling SS, Norcini JJ. The international medical graduate pipeline. *Health Aff (Millwood)*. 2003;22(4):94–96.
- 9 Joshi R, Ling FW, Jaeger J. Assessment of a 360-degree instrument to evaluate residents' competency in interpersonal and communication skills. *Acad Med*. 2004;79(5):458–463.
- 10 Dyche L. Interpersonal skill in medicine: the essential partner of verbal communication. *J Gen Intern Med*. 2007;22(7):1035–1039.
- 11 Weigelt JA, Brasel KJ, Bragg D, Simpson D. The 360-degree evaluation: increased work with little return? *Curr Surg*. 2004;61(6):616–626.
- 12 Wood J, Collins J, Burnside ES, Albanese MA, Propeck PA, Kelcz F, et al. Patient, faculty, and self-assessment of radiology resident performance: a 360-degree method of measuring professionalism and interpersonal/communication skills. *Acad Radiol*. 2004;11(8):931–939.
- 13 Abraham J, Wade DM, O'Connell KA, Desharnais S, Jacoby R. The use of simulation training in teaching health care quality and safety: an annotated bibliography. *Am J Med Qual*. 2011;26(3):229–238.
- 14 Taylor DK, Buterakos J, Campe J. Doing it well: demonstrating general competencies for resident education utilising the ACGME Toolbox of Assessment Methods as a guide for implementation of an evaluation plan. *Med Educ*. 2002;36(11):1102–1103.
- 15 Kelly M, Murphy A. An evaluation of the cost of designing, delivering and assessing an undergraduate communication skills module. *Med Teach*. 2004;26(7):610–614.
- 16 Fallowfield L, Lipkin M, Hall A. Teaching senior oncologists communication skills: results from phase I of a comprehensive longitudinal program in the United Kingdom. *J Clin Oncol*. 1998;16(5):1961–1968.
- 17 Skillings JL, Porcerelli JH, Markova T. Contextualizing SEGUE: evaluating residents' communication skills within the framework of a structured medical interview. *J Grad Med Educ*. 2010;2(1):102–107.
- 18 Donnelly MB, Sloan D, Plymale M, Schwartz R. Assessment of residents' interpersonal skills by faculty proctors and standardized patients: a psychometric analysis. *Acad Med*. 2000;75(suppl 10):93–95.
- 19 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.
- 20 Buckman R, Kason Y. *How to Break Bad News: A Guide for Health Care Professionals*. Baltimore: Johns Hopkins University Press; 1992.
- 21 Kaplan SH, Greenfield S, Ware JE Jr. Assessing the effects of physician-patient interactions on the outcomes of chronic disease. *Med Care*. 1989;27(suppl 3):110–127.
- 22 Bootman JL, Harrison DL, Cox E. The health care cost of drug-related morbidity and mortality in nursing facilities. *Arch Intern Med*. 1997;157(18):2089–2096.
- 23 Holmboe ES, Hawkins RE. Methods for evaluating the clinical competence of residents in internal medicine: a review. *Ann Intern Med*. 1998;129(1):42–48.
- 24 The Joint Commission website. www.jointcommission.org. Accessed November 10, 2012.
- 25 Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington DC: Institute of Medicine of the National Academies; 2001.
- 26 Griffen FD, Stephens LS, Alexander JB, Bailey HR, Maizel SE, Sutton BH, et al. Violations of behavioral practices revealed in closed claims reviews. *Ann Surg*. 2008;248(3):468–474.
- 27 Hwang SW, Li J, Gupta R, Chien V, Martin RE. What happens to patients who leave hospital against medical advice? *CMAJ*. 2003;168(4):417–420.
- 28 Jeremiah J, O'Sullivan P, Stein MD. Who leaves against medical advice? *J Gen Intern Med*. 1995;10(7):403–405.
- 29 O'Brien RG, Kaiser MK. MANOVA method for analyzing repeated measures designs: an extensive primer. *Psychol Bull*. 1985;97(2):316–333.
- 30 Sorg JC, Wilson RD, Perzynski AT, Tran D, Vargo MM. Simplifying the 360-degree peer evaluation in a physical medicine and rehabilitation residency program. *Am J Phys Med Rehabil*. 2012;91(9):797–803.
- 31 Risucci DA, Tortolani AJ, Ward RJ. Ratings of surgical residents by self, supervisors and peers. *Surg Gynecol Obstet*. 1989;169(6):519–526.