The American
Journal of Surgery®

Comparison of simulation-based assessments and faculty ratings for general surgery resident milestone evaluation: Are they telling the same story?

CrossMark

Aimee K. Gardner, Ph.D.[a,*], Kareem AbdelFattah, M.D.[b]

[a]*Department of Surgery, School of Allied Health Sciences, Baylor College of Medicine, Houston, TX, USA; and [b]Department of Surgery, UT Southwestern Medical Center, Dallas, TX, USA*

Abstract

BACKGROUND: The goal of this article is to examine the role of simulation in documenting resident nontechnical competencies outlined by the general surgery milestones, and compare those assessments to faculty ratings.

METHODS: Trainees completed a multiphasic simulation scenario to assess ability to obtain informed consent, lead a preoperative time out, crisis management, communication, and delivering bad news. Assessments from this scenario were compared with ratings of these same competencies collected from clinical rotations.

RESULTS: Twenty-six PGY1 trainees participated in the training program. Results revealed no significant correlations between simulation performance and faculty rotation ratings for any of the 5 competencies. With the exception of communication in the operating room, faculty reported an overall inability to reliably observe these competencies 26% to 41% of the time.

CONCLUSIONS: This work suggests that traditional end-of-rotation evaluations may not be the most feasible and valid method to evaluate resident nontechnical skills for milestone assessment. Simulation-based assessments should be considered to fill this gap.

© 2016 Elsevier Inc. All rights reserved.

Program directors are tasked with ensuring that their residency program adequately prepares the next generation of trainees to be knowledgeable and skilled surgeons. Unfortunately, this undertaking is becoming increasingly difficult as the areas of competency continue to adapt and grow and the clinical and administrative demands placed on residents and faculty increase. To support this effort, the Accreditation Council for Graduate Medical Education, has implemented the Next Accreditation System, which defines specific milestones regarding resident knowledge, skills, and other competencies along a continuum.[1] The specific and granular nature of these milestones requires that residency programs implement new assessment tools so that faculty can make informed decisions about resident progression.

Simulation has been proposed to play a major role in satisfying these requirements, as it provides an avenue to

evaluate resident competence in a cost-efficient, objective, and standardized manner that may be too impractical or burdensome through traditional assessment systems.[2–4] For example, the milestones require that residents be systematically evaluated twice a year on leadership and communication when unexpected events occur in the operating room (OR) (interpersonal and communication skills [ICS] 3) and delivering bad news (DBN) to patients and families sensitively and effectively (ICS1). Obviously, these clinical events occur with high variability in terms of frequency, resident role, and characteristics of the case, making it difficult to assess all residents in a systematic fashion. Simulation, then, may provide an ideal avenue for "leveling the playing field" and ensuring that all trainees have an opportunity to practice and demonstrate their skills, faculty have a standard backdrop to evaluate skills, and program directors can comprehensively examine any large scale deficits for program improvement. However, without specific comparison of these simulation-based evaluations to those obtained from clinical faculty, the value of creating milestone assessments based in simulation remains unknown. Given the time, logistics, and resources needed to create and implement such a robust curricula, more work is needed to demonstrate the value of simulation-based milestone assessment.

To examine these concepts further, our team implemented a robust simulation program to complement the milestones project with 3 overarching aims. First, we wanted to examine the feasibility of measuring milestone 5 competencies in a simulated environment, including the ability to deliver bad news (ICS1; level 3), communication with team members (ICS2; level 2), performing an informed consent (ICS3; level 2), leading a preoperative "time out" (ICS3; level 2), and demonstrating leadership when unexpected events occur in the OR (ICS3; level 4). Second, we sought to investigate how these evaluations compared with those in these areas provided by faculty during clinical rotations. Finally, we wanted to examine performance across an entire cohort to ascertain if any programmatic gaps emerged to inform future curricular development.

Methods

Participants consisted of PGY1 trainees in the UT Southwestern General Surgery Residency program. The IRB deemed this project exempt. All training sessions took place in the last 2 months of Internship. Participants received pre training materials 1 week before attending the simulation session. Pretraining materials consisted of the SPIKES Delivering Bad News protocol,[5] a surgical safety checklist from the World Health Organization (WHO),[6] the Society of American Gastroenterologists and Endoscopic Surgeons laparoscopic troubleshooting guide,[7] and an article discussing preparation of first year General Surgery residents for obtaining informed consent.[8] However, no quizzes or surveys of completion regarding these materials were used.

On the day of training, participants were provided with an overview of the goals of the training session. Participants then began the 1-hour training session in which the first phase was to consent a standardized patient (SP) for a routine laparoscopic cholecystectomy. Trainees were provided an overview of the patient's medical history and diagnosis along with the hospital's consent form, which they were to complete with the SP. Phase II of the training session took place in a high-fidelity OR suite. Team members consisted of a confederate anesthesiologist, scrub tech, and circulating nurse.

On entering the OR, the confederate nurse instructed the surgeon that the patient was ready and prompted the surgeon to lead the time out using the WHO safety checklist displayed on the wall. The scenario then progressed as a modified version of Laparoscopic Troubleshooting Module included in the ACS/APDS National Skills Curriculum,[9] in which trainees must lead the team in systematic troubleshooting of laparoscopic equipment and identify and treat a physiologic disturbance. Specifically, trainees had to troubleshoot the laparoscopic tower to remedy loss of visualization before the procedure started and also lead the team to perform Advanced Cardiac Life Support (ACLS) when the patient became bradycardic during insufflation. After 3 rounds of Advanced Cardiac Life Support, regardless of trainee actions, the participant was informed that the patient was unable to be resuscitated. The trainee then left the OR and was provided information on specifics of the physiological problem (a CO2 embolism) such that all trainees had a similar understanding regarding why the patient died. The trainee was also instructed that they needed to discuss the death with the sister of the family member (also an SP). Trainees were allowed as much time as they needed to gather their thoughts and prepare for the conversation. When ready, the trainee was shown to the room where the sister of the patient was awaiting updates from the surgeon. The SP followed a scripted set of behaviors and responses so that all trainees had to respond to similar prompts during the conversation.

At the completion of the training, all trainees were debriefed by a simulation faculty member using the debrief with good judgment approach.[10] After the debriefing, trainees were provided a copy of their video-recorded performance for reflection and review. Trainees were specifically instructed to rate their communication in the OR using the Operating Room Communication Assessment (ORCA) tool[11] and to provide a written critique of their DBN performance. The ORCA ratings were done on-site immediately after the scenario, whereas the DBN critique was emailed to the simulation staff within the following week.

All phases of the training session were video recorded for performance assessment. Performance during the consent was evaluated using 12 checklist items derived from the consent form. Time out performance was assessed by 9 checklist items from the ACS/APDS Curriculum.[11] Communication was assessed using the 16-item ORCA