Simulation training in critical care: Does practice make perfect?

Rachael Springer, MD,a,b John Mah, MD,a,b Ian Shusdock,b Robert Brautigam, MD,a,b Steve Donahue,c and Karyn Butler, MD,a,b Hartford, CT

Background. Few data exist regarding the effectiveness of simulation in resident education in critical care. The purpose of this study was to determine whether multiple-simulation exposure (MSE) or single-simulation exposure (SSE) improved residents’ recognition of shock and initial management of the critically ill simulated surgical patient.

Methods. Data were collected at a level 1 trauma center. Surgery, anesthesiology, and emergency medicine residents were given a multiple-choice question (MCQ) pretest before a tutorial on the recognition and management of shock followed by high-fidelity simulation/debriefing and MCQ post-test. MSE residents had 1.5 hours of simulation per resident over 3 days, and SSE residents had 1.5 hours of simulation as a group in 1 day. Pre- and posttest comparisons overall and subgroup analysis for MSE versus SSE were performed.

Results. Data was available for 45 MSE residents and 15 SSE residents. Overall posttest percent correct was greater than pretest percent correct (81% ± 9% vs 75% ± 13%, post- versus pre-, P = .01). Subgroup analysis demonstrated significantly improved post- versus pretest performance for MSE residents only. There were no differences in pre- or posttest performance for MSE residents during the first 4 months of the academic year versus the last 4 months. Pretest performance over 12 months of observation for MSE residents showed no significant differences.

Conclusion. Repeated simulation exposure was more effective than single simulation exposure at improving MCQ performance designed to measure the recognition and management of shock in the critically ill simulated surgical patient. Duration of training did not impact MCQ performance. (Surgery 2013;154:345-50.)

From the Department of Surgery,a University of Connecticut, Division of Surgical Critical Care,b Hartford Hospital, and Center for Education, Simulation, Innovation,c Hartford Hospital, Hartford, CT

JOHN DEWEY modeled early concepts of education by recognizing the impact of experience, democracy, continuity, and interaction, and these building blocks formed the foundation of successful educational curricula. In 1956, Bloom’s taxonomy of educational objectives focused curricula on the three imperatives of learning: cognitive, psychomotor, and affective. Although surgical critical care training requires all three of these imperatives, there are few formal curricula that incorporate them into postgraduate educational programs. Construct validity is defined (Cronback and Meehl) as the extent to which an instrument measures the construct that it is supposed to measure. “Successful” education is inherently dependent on the components of construct validity (content validity, response process, internal structure, consequential validity, and relationship to other variables) to determine whether the measurement of learning is reflective of true learning. It is surprising that this paradigm has not been widely integrated in modern critical care training.

Throughout medical training there is a concept of “see one, do one, teach one.” In theory this concept may seem applicable to all aspects of medicine, but in practice, especially in regard to surgical critical care, the concept proves challenging and does not take into consideration the need to reduce the patient’s risk of harm. In fact, Landrigan et al reported that additional efforts were needed to introduce safety measures into routine healthcare to lower the rate of preventable harm. Because of the complex and dynamic nature of the critically ill patient, it may be difficult for the inexperienced resident to understand concepts of physiologic compromise while simultaneously managing the patient. In fact, in the midst of critical care crises, the least-experienced provider generally

Accepted for publication April 19, 2013.
Reprint requests: Karyn Butler, MD, Department of Surgery, Hartford Hospital, 80 Seymour Street, Hartford, CT 06102.
E-mail: kbutler@harthosp.org.
0039-6060/ - see front matter © 2013 Mosby, Inc. All rights reserved.
http://dx.doi.org/10.1016/j.surg.2013.04.038
functions in an observatory role and major decisions are made by those with the most training. This paradigm contributes to an educational gap in managing the critically ill patient that is further compounded by restricted resident work hours.\textsuperscript{5}

Simulation as an educational tool enhances practice and learning, and it can be applied to many different disciplines and trainees, including medicine, aviation, and the military.\textsuperscript{6-8} Industry has led the way in educational training for high acuity scenarios (airline pilot training) and improving consumer safety (crash testing in the automotive industry). Surgeons have used simulation for centuries to teach dissection skills with the use of cadavers and animal models. These disciplines have capitalized on the expansion of simulation to include standardized patients, computers and virtual reality. In the medical field the most important attribute of simulation is that it permits exploration and repeated practice without exposing patients to risk of harm.\textsuperscript{9,10}

Numerous studies investigating the role of simulation in medical education identify benefits such as improved overall performance, reduced response time, improved team-interaction skills under crisis, and less deviation from the standard of practice.\textsuperscript{11} Despite these benefits, there are few studies examining the role of simulation in surgical critical care training. The purpose of this observational study was to determine whether the simulation training of postgraduate residents improved the recognition and initial management of shock in the critically ill simulated surgical patient. As many programs struggle with incorporating educational sessions in the wake of duty-hour restrictions, we also sought to determine whether single exposure to simulation was as effective as multiple exposures.

**MATERIALS AND METHODS**

**Design.** All simulation exposures were performed in the Center for Education, Simulation & Innovation, a 20,000-square-foot facility at Hartford Hospital, the main teaching hospital for the University of Connecticut ACGME training programs. The Center for Education, Simulation & Innovation is equipped with state-of-the-art high-fidelity mannequins with computer-supported physiologic responses. Separate classrooms are available for viewing of scenarios and for interactive debriefing. Data collected included resident performance on multiple-choice question (MCQ) tests designed to measure knowledge of the recognition and immediate management of patients in shock.

<table>
<thead>
<tr>
<th>Table I. Goals and objectives of shock simulation curriculum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recognize the etiology of shock</td>
</tr>
<tr>
<td>- Hypovolemic</td>
</tr>
<tr>
<td>- Cardiogenic</td>
</tr>
<tr>
<td>- Distributive</td>
</tr>
<tr>
<td>- Septic</td>
</tr>
<tr>
<td>- Neurogenic</td>
</tr>
<tr>
<td>- Obstructive</td>
</tr>
</tbody>
</table>

Initiate management to stabilize patient
- Establish diagnosis
- Implement therapy
- Communicate to primary team and intensive care unit supervisor

**Simulation participants.** All PGY-1 surgery residents participate in the shock curriculum while in boot camp during the first month of internship. This group is referred to as the single-simulation exposure (SSE) group as the entire hands-on simulation or immersion is completed as a group in a single 1.5-hour setting. The multiple-simulation exposure (MSE) group comprised PGY-1 surgery, PGY-2 or -3 anesthesiology residents, and PGY-2 emergency medicine residents who completed the shock curriculum as part of their required month-long rotation through the surgical intensive care unit (ICU). The MSE residents were immersed in simulation for a total of 1.5 hours per resident over 3 separate days.

**Simulation curriculum.** The curriculum is based on a set of objectives designed to simulate the manifestation of the four types of shock: hypovolemic, cardiogenic, obstructive, or distributive (Table I). Figure 1 outlines the details of the curriculum. The residents are first given a 14 MCQ pretest, followed by a 30-minute didactic presentation on the recognition and management of the different types of shock. After the didactic presentation, the goals of the module are provided before introduction of the scenario. The MSE residents are divided into pairs; the first pair is immersed in simulation whereas the second pair observes in the classroom. The roles are switched after 30 minutes of simulation. Learners are instructed to work together and articulate what they believe is the evidence supporting the type of shock and to then initiate immediate management to stabilize the “patient.” The simulation instructor provides positive or constructive feedback during the simulation module to help modify responses as needed. The simulation scenario covers at least two types of shock depending on time available. All residents are exposed to hypovolemic shock during their immersion.
At the completion of the simulation module, the instructor performs an interactive debriefing that allows the observers to provide peer-to-peer comments. The next day, the resident pairs are again immersed into simulation. This second exposure provides an opportunity for repetitive practice and escalating level of difficulty and by the third simulation immersion 5–10 days later, residents are exposed to an even greater degree of difficulty. The total simulation immersion time for the MSE residents is 1.25-1.5 hours per pair distributed over three sessions, where each resident serves the role of the team leader at least once. After the third debriefing, a 16 MCQ posttest is given, which completes the curriculum. The SSE residents undergo the same pretest, the same 30 minutes of didactic presentation and introduction of the goals and objectives of the module as the MSE residents. The residents are immersed in simulation in pairs, but this occurs only once for each pair of residents for a total of 30 minutes of simulation immersion per resident pair distributed over one session. The balance of the time is spent as a group observing other resident pairs and participating in the debriefing. The curriculum is concluded with completion of the 16 MCQ posttest.

**Construct validity of MCQ test.** Known-groups validation was performed to provide evidence to support the construct validity of the MCQ tests. The tests were administered to clinicians at three levels of performance: inexperienced, experienced, and expert. Inexperienced and expert clinicians had no previous exposure to the curriculum. The inexperienced clinicians were third-year medical students, experienced clinicians were critical care or cardiology fellows, and the experts were surgical intensivists/traumatologists. Performance on the pre- and post-MCQ was recorded as percent correct of the total questions.

---

*Fig 1. Algorithm outlining the simulation curriculum on the recognition and management of shock. Residents were conveniently assigned to MSE or SSE based on rotation requirements determined by training program. MSE, Multiple simulation exposure; SSE, single simulation exposure.*
**Statistical analysis.** Data are presented as mean ± SD. Pre- and post-test comparisons overall and subgroup analysis for MSE versus SSE was performed via the use of independent *t* test and repeated-measures analysis of variance.

**RESULTS**

**Study group.** The observation period for data collection was from August 2011 through July 2012. Residents in the SSE group were observed only in July 2012 because this was the only time the curriculum was offered as part of the PGY-1 surgery boot camp. Different groups of residents (anesthesiology, surgery, emergency medicine) rotated monthly in the surgical ICU and were assigned to the MSE group. A total of 60 residents completed the shock simulation curriculum during the period of data collection.

**Construct validity of MCQ test.** Pretest results were available for eight learners at the inexperienced level, seven at the experienced level, and six at the expert level. For the post-test, results were available from four learners at the inexperienced level, seven at the experienced level and six at the expert level. Performance results are shown in Table II. There was a difference in pretest percent correct between the expert group and the experienced and inexperienced groups. These results support the validity of the MCQ test as an instrument to evaluate knowledge gained from the simulation.

**MCQ test performance.** Pretest scores were available for 60 residents (MSE; *n* = 45, SSE; *n* = 15) and posttest scores for 45 residents (MSE; *n* = 28, SSE; *n* = 17). Overall (MSE + SSE) posttest percent correct was greater than pre-test percent correct (81% ± 9% vs 75% ± 13%, post-versus pre-, *P* = .01). Subgroup analysis demonstrated significantly improved post-versus pretest performance for MSE residents only. There were no differences in mean pretest scores among the SSE or the MSE residents (Fig 2).

Comparison of pretest performance over time was analyzed for the MSE residents. During the 12-month observation period, data was available for 10 of the 12 months and showed no differences in MCQ performance over time. There were no differences in pre- or posttest performance for residents during the first 4 months of the academic year versus the last 4 months (Fig 3).

**DISCUSSION**

The results from our observational study showed that simulation training of residents on recognizing and initiating immediate management of shock increased knowledge as measured by MCQ test performance. Moreover, we found that repeated simulation exposure was more effective than single simulation exposure at improving MCQ performance. Importantly, the beneficial effects of simulation training were independent of cumulative experiential training or specialty.

**Construct validity of MCQ test.** Kim et al.\(^{13}\) validated the Ottawa Global Rating Score (GRS) as an instrument to assess resident performance in the resuscitation of critically ill simulated patients. Similar to our study, residents from multiple disciplines underwent three half-day sessions of high-fidelity simulation preceded by a tutorial. Subsequent simulation immersions were conducted 2 days and 2–3 weeks after the tutorial. The GRS scored performance on a scale of 1 (novice) to 7 (expert), and comparisons were made between PGY-1 and PGY-3 residents. The authors showed that the GRS significantly

<table>
<thead>
<tr>
<th>Table II. Known-groups validation of MCQ test</th>
<th>Pre-test* (% correct)</th>
<th>Post-test (% correct)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inexperienced</td>
<td>64 ± 13</td>
<td>79 ± 5</td>
</tr>
<tr>
<td>Experienced</td>
<td>78 ± 13</td>
<td>85 ± 12</td>
</tr>
<tr>
<td>Expert</td>
<td>91 ± 10</td>
<td>92 ± 5</td>
</tr>
</tbody>
</table>

*\(P = .002\): comparisons between known-groups for pretest performance. MCQ, Multiple choice questions.
differentiated PGY-1 performance from PGY-3 performance with high inter-rater reliability. They concluded that there is evidence of the construct validity of the GRS but that further revisions were needed before it could be used as a summative evaluation tool.

Several important differences exist with our study and the work by Kim et al. First, although the simulation curricula were similar, we evaluated acquisition of knowledge based on MCQ performance. Second, we used one simulation scenario (shock) to reduce variability between evaluations. Third, we evaluated cognitive performance over time as well as in response to two different types of simulation immersion (multiple versus single exposure). Despite these differences, our results are in alignment with Kim et al in supporting construct validity; this is an essential component of successful educational curricula and one that must not be overlooked as simulation models are developed.

**Simulation in resident education.** Young et al evaluated cognitive performance of surgery and emergency medicine residents in critical care simulation. The study group consisted of residents who had greater than 24 months of postgraduate training and those who had less than 24 months. These researchers also looked at resident experience dichotomized at 10 weeks of ICU exposure and assessed residents decision making skill by counting the number of proactive versus reactive tasks performed by each subject. They found that the amount of ICU experience rather than the duration of residency training had the greater impact on performance. They further noted that residents with more ICU experience but less clinical training outperformed more senior residents.

Similarly, in our study, the MSE group outperformed the SSE group. Our results are in alignment with Young et al’s observation that more experience results in better performance. The lack of difference in pretest scores between the MSE and SSE group further supports the observation that test performance was not dependent on the length of training of the residents or their specialty. We did observe a difference between pretest and posttest performance at the end of the study period; however, because of the small sample size per month, we were not able to determine monthly improvement in pre-post test performance. Further investigation with the use of paired resident performance is needed to address this issue. It is important to note that we could not control for independent study by residents in either group and thus cannot exclude the impact this may have on MCQ test performance.

A limitation of the study was that we did not evaluate whether patient care was improved as a result of simulation experience and this is clearly an important potential benefit of simulation. Nishisaki et al demonstrated that pediatric critical care fellows who underwent high-fidelity simulation with repetitive practice, defined as repetition of a task until proficiency was achieved, had improved clinical performance when managing resuscitation, sepsis, trauma and traumatic brain injury. The ability to transfer cognitive performance from the simulation laboratory to the bedside is an imperative that all curricula should strive to achieve. A second limitation of the study is that we did not have paired resident performance for analysis. This was primarily attributed to absences that were beyond control such as approved vacation or absence due to duty
hour restrictions. This was evident in the unequal numbers of participants for both the MSE and SSE groups.

CONCLUSION

Our results show that simulation in surgical critical care improves cognitive performance. Our curriculum emphasized three important principles: (1) Construction of a predetermined curriculum with specific goals and objectives; (2) the use of known-group validated measuring instruments for construct validity; and (3) the focus on an established educational domain (knowledge). We believe that construct validity is an important first step in evidence-based education. Moreover, the method of curriculum delivery is an essential component to an effective educational program in simulation. We found that repeated simulation immersion, combined with passive and active learning, resulted in improved post-test performance. It is imperative that simulation in surgical critical care training be “as accessible as a book on a library shelf” so that the risk of patient harm is minimized during resident training.

The authors acknowledge Karyl Burns, PhD, for her help with review of the manuscript.

REFERENCES