The Effect of Simulation-based Training on Medical Students' Perceptions, Knowledge, and Skill at Baseline and 6-month Follow-up

Cameron G. Shultz, PhD, MSW (Corresponding author)
St. Mary Mercy Livonia Hospital
Graduate Medical Education
36475 Five Mile Road, Livonia, MI, 48154, USA
Tel: 1-734-655-2303 E-mail: cshultz@umich.edu

James M. Cooke, MD
Department of Family Medicine and Learning Health Sciences,
University of Michigan
1150 W. Medical Center Drive, M7300 Med Sci I
Ann Arbor, MI 48109, USA
Tel: 1-734-998-7120 E-mail: cookej@med.umich.edu

Keri L. Denay, MD
Department of Family Medicine, University of Michigan
1150 W. Medical Center Drive, M7300 Med Sci I
Ann Arbor, MI 48109, USA
Tel: 1-734-998-7120 E-mail: kschwide@med.umich.edu

Mikel Llanes, MD
Department of Family Medicine
University of Michigan
1150 W. Medical Center Drive, M7300 Med Sci I
Ann Arbor, MI 48109, USA
Tel: 1-734-998-7120 E-mail: mllanes@med.umich.edu

Philip Zazove, MD, MM
Department of Family Medicine
University of Michigan
1150 W. Medical Center Drive, M7300 Med Sci I
Ann Arbor, MI 48109, USA
Tel: 1-734-998-7120 E-mail: pzaz@med.umich.edu

Scott A. Kelley, MD
Department of Family Medicine
University of Michigan
1150 W. Medical Center Drive, M7300 Med Sci I
Ann Arbor, MI 48109, USA
Tel: 1-734-998-7120 E-mail: scottak@med.umich.edu

Abstract

The principal goal of most simulation-based learning is enhanced clinical skill, with the desired outcome being skill retention and improved performance over time. While evidence supports simulation-based training at the clerkship, graduate, and post-graduate level, the evidence supporting its long-term benefit at the pre-clerkship level is less well established. Using quantitative methods, this research assessed the effect of skill-based clinical simulation training on all second-year medical students' enrolled in two simulation courses at the University of Michigan during the 2013-2014 academic year (n=39). Pre-course, post-course, and 6-month follow-up questionnaires were administered, and analysis of variance was used to assess change in students' knowledge, self-reported confidence, and competency. Students in Course 1 (n=12) completed a post-course and 6-month follow-up skills test, and results were compared to faculty controls. Differences between groups were assessed using t-tests. Odds ratios were calculated. Knowledge improved from pre-course to post-course; the gain was retained at follow-up only for students in Course 1. Students perceived the courses as improving knowledge, confidence, and skill, though scores were significantly lower at follow-up. For the subset of students participating in the skills test, the time to complete central line and thoracentesis at follow-up did not significantly differ from post-course; however, an increase was observed for lumbar puncture. Compared to faculty, students took significantly more time to complete the central line procedure. Outcomes for needle redirects and correct sequencing were mixed, with most procedures showing no significant difference between measurement periods. Assessing students' skill—and the maintenance of that skill over time—using objective and empirically derived measures can be challenging. The evaluation strategy described herein could be adapted to many procedures commonly practiced within primary care and other medical specialties. Brief exposure to clinical simulation early in medical training and prior to clerkships can have a positive and lasting effect on medical students' knowledge, confidence, and skill for selected procedures.

Keywords: Clinical simulation, education, medical student, knowledge, skill

1. Introduction

Creating new opportunities for medical students to develop and practice clinical procedures remains an important priority within undergraduate medical education programs. To address this need, simulation mannequins and other simulation-based teaching methods are growing in popularity due to their low cost, safety, and capacity for repetitive use.(Bradley, 2006; Issenberg, McGaghie, Petrusa, Gordon, & Scalese, 2005; McGaghie, Issenberg, Cohen, Barsuk, & Wayne, 2011; McGaghie, Issenberg, Petrusa, & Scalese, 2010; Passiment, Sacks, & Huang, 2011) Simulation-based instructional programs permit learners to practice and receive focused feedback on clinical skills within a controlled setting, thus preparing learners for experiences with real patients. The principle goal of most simulation-based learning is the enhancement of clinical skill, with the desired outcome being skill retention and improved performance in the clinical setting. While evidence supports the use of simulation-based training at the clerkship level (ie, 3rd and 4th years of undergraduate training) and within graduate and post-graduate training programs,(McGaghie, Issenberg, Barsuk, & Wayne, 2014; McGaghie, et al., 2011; McGaghie, et al., 2010) the evidence supporting its use at the pre-clerkship level (ie, 1st and 2nd year undergraduate medical students) is less well established.(Chou, Abdelshehid, Clayman, & McDougall, 2006; Consoli et al., 2012; Fraser et al., 2009; Fraser et al., 2011; Goulart et al., 2012; Halm, Lee, & Franke, 2010; Heitz, Brown, Johnson, & Fitch, 2009; Karabilgin, Vatansever, Caliskan, & Durak, 2012; Kelly et al., 2013; Lee et al., 2009; Mittal, Morris, & Kelz, 2011; Schubart et al., 2012; van Zuilen, Kaiser, & Mintzer, 2012) A limited but growing number of studies show that learners can also retain simulation-acquired skills over time,(Ahya et al., 2012; Barsuk, Cohen, McGaghie, & Wayne, 2010; Crofts et al., 2007; Didwania et al., 2011; Wayne, Siddall, et al., 2006) with several studies assessing retention among students during pre-clerkship years.(Chou, et al., 2006; Consoli, et al., 2012; Fraser, et al., 2009; Schubart, et al., 2012; van Zuilen, et al., 2012) However, we know of no previous study that systematically assesses retention among pre-clerkship medical students at an interval extending beyond several months.

The objective of this study was to evaluate the effect of skill-based clinical simulation training on second-year medical students' knowledge, confidence, and skill at baseline, and retention at 6-month follow-up.

1.1 Simulation-based instruction

As described by Bradley (2006), simulation is "[t]he technique of imitating the behaviour of some situation or process...by means of a suitably analogous situation or apparatus, especially for the purpose of study or personnel training." Simulation-based medical training can take many different forms, including life-like mannequins, medical team training, software-based multimedia simulation, and standardized patients (among others). Many procedures, treatments, and clinical scenarios can be modeled using simulation, ranging from the highly specialized (eg, laparoscopic surgery) to the relatively straightforward (eg, hand-washing).

A number of factors contribute to fewer opportunities for pre-clerkship medical students to actively participate in procedures performed on real patients, including changing supervision standards for students and residents, the need for residents to perform procedures at the expense of student participation, and concerns over safety and outcomes. For such reasons, alternative learning experiences—eg, medical simulation—that incorporate rigorous evaluation methods are needed.

2. Methods

Prior to implementation, this study was evaluated and classified as exempt from ongoing review by the University of Michigan (UM) Institutional Review Board.

2.1 Overview

As described by Kneebone (2005), 4 essential constructs exemplify effective simulation based training—knowledge, skill, professional practice, and clinical context. Corresponding to each construct, Kneebone proposed specific criteria for consideration when evaluating training programs:

(1) Simulations should allow for sustained, deliberate practice within a safe environment, ensuring that recently acquired skills are consolidated within a defined curriculum which assures regular reinforcement;

(2) Simulations should provide access to expert tutors when appropriate, ensuring that such support fades when no longer needed;

(3) Simulations should map onto real-life clinical experience, ensuring that learning supports the experience gained within communities of actual practice; and

(4) simulation-based learning environments should provide a supportive, motivational, and learner-centered milieu which is conducive to learning.

These criteria informed the development and implementation of the skill-based clinical simulation courses targeting second-year medical students at the UM.

Given the wide variety of inpatient and outpatient procedures within the practice of family medicine, family physicians are well positioned to provide targeted instruction to preclinical medical students on diagnostic and therapeutic procedures regularly encountered during medical and surgical clerkships. Building off the success of a pilot course using clinical simulation to provide pre-clerkship medical students with hands-on experience in family medicine-related procedures,(Kelley & Cooke, 2013) two simulation-based courses were offered to 2nd year medical students at the UM: Course 1 focused on central line placement, lumbar puncture, and thoracentesis; Course 2 focused on splinting and casting. Course 1 was conducted over three 2-hour sessions, one for each procedure; Course 2 was limited to a single 2-hour session. Both courses were taught by clinical faculty from the UM Department of Family Medicine. The primary goal for each course was to increase students' knowledge, confidence, and skill.

2.1.1 Simulation Procedures for Course 1

Course 1 utilized several simulation instruments commonly used within academic medical education settings: the Blue Phantom central line mannequin, Kyoto Kagaku lumbar puncture simulator, and Blue Phantom midscapular thoracentesis model. The Blue Phantom central line mannequin offers realistic anatomy of the upper thorax and neck. The model has an internal jugular vein and additional landmarks such as the trachea, manubrium, and clavicle. Veins are compressible, and ultrasound can be utilized to visualize anatomy and needle placement (Blue Phantom, Sarasota, FL). The Kyoto Kagaku lumbar puncture simulator includes life-like skin and tissue resistance to the spinal needle, enabling learners to gain skill in performing epidural and lumbar puncture procedures. Cerebral spinal fluid can be collected, and fluid pressure measured (Kyoto Kagaku, Torrance, CA). The Blue Phantom midscapular thoracentesis model helps learners develop essential skills, such as ultrasound-guided catheter insertions for those with pleural effusions. It contains realistic anatomical features (eg, chest wall, ribs, intercostal spaces, pleural cavity), offers positive fluid flow, and can provide instant feedback when pleural effusion fluid is accessed (Blue Phantom, Sarasota, FL).

With an experienced family medicine faculty instructor for each group, students practiced procedures in small groups of three to four students per simulator.

2.1.2 Simulation Procedures for Course 2

Course 2 focused on applying a volar wrist splint and short arm cast using fiberglass casting material. Students were paired, and each took turns being either the patient or the provider. The family medicine faculty instructor floated from pair to pair.

2.1.3 Course 1 and 2

For both courses, sessions consisted of a presentation (eg, anatomic landmarks, indications, contraindications, technique, etc) and detailed step-by-step demonstration on how to conduct the procedure. This was followed by hands-on practice with one-on-one coaching and in-depth discussion of the finer details of the procedure. Discussion between faculty and students included qualitative feedback on both performance and technique.

2.1.4 Incentive

To compensate students for their time, a small financial incentive (up to $25 total) was offered for participation in the pre-course, post-course, and 6-month follow-up assessment activities.

2.2 Measures and Statistical Analyses

2.2.1 Questionnaire

Using quantitative research methods, an objective measure of students' knowledge and a subjective, self-reported measure of knowledge, confidence, and skill were assessed at three time periods: immediately prior to the course (pre-course), immediately after the course (post-course), and 6 months after completing the course (follow-up) using a paper-based (pre- and post-course) or web-based (follow-up) questionnaire. In the absence of a standardized tool assessing knowledge and confidence for the two simulation courses, study investigators employed a multi-step process to develop instruments addressing each course's specific content. Steps included explicit delineation of the measured constructs (eg, knowledge and confidence), an assessment of each question to confirm its face and content validity, piloting the instrument with study team members, and repeating the process over several revision cycles. Using this collaborative and iterative process, each question was reviewed and approved via consensus by at least 3 clinical faculty and a PhD methodologist within the UM Department of Family Medicine.

2.2.2 Knowledge

Students in Course 1 were asked a total of 9 knowledge-based questions at each measurement period. With 3 questions for each procedure, possible scores ranged from 0 to 9. Students in Course 2 were asked 3 knowledge-based questions (all for the single procedure) at each measurement period. Possible scores ranged from 0 to 3. Questions for both courses were designed to resemble those encountered in Step 2 of the US Medical Licensing Examination (USMLE).

One-way repeated measures ANOVA and pairwise comparisons (paired-samples t-test with Bonferroni adjustment) were employed for each course to assess for differences in students' knowledge between pre-course, post-course, and follow-up measurement periods; p-values and effects sizes are reported.

2.2.3 Self-reported Knowledge, Confidence, and Skill

Student's self-reported knowledge, confidence, and skill were assessed post-course and at follow-up using a course-specific Likert-type question tailored to each domain (knowledge, confidence, and skill). Students' agreement with the statement was measured using a 10-point scale, ranging from strongly disagree to strongly agree. All students from both courses were asked the same questions; hence, results were pooled.

One way repeated measures ANOVA were performed to compare self-reported scores between measurement periods; p-values and effect sizes are reported.

2.2.4 Objective Skill Assessment

For students in Course 1, skill was assessed post-course and at follow-up using a set of objectively derived performance measures including the time to complete the procedures, the number of needle redirects, and whether or not the student followed the correct sequence of ordered steps for each procedure. Because time to complete the procedures in most cases was not normally distributed, time (in seconds) was transformed to its log.

Paired samples t-tests were employed to compare students' time for each procedure post-course and at follow-up; paired t-tests were also used for comparing student re-directs post-course and at follow-up. Effect sizes and confidence intervals are reported. One-sample t-tests—using faculty mean time as the reference—were used to test for differences between student times and faculty times; effect sizes and confidence intervals are reported. Odds ratios were used to measure the association between correctly following the sequence of ordered steps for each procedure post-course and at follow-up; p-values and confidence intervals are reported.

3. Results

3.1 Sample

Twelve students participated in Course 1, and 27 in Course 2. There were no statistically significant differences in the age, gender, or race/ethnicity of the students in each course. Mean age was 25.77 years (standard deviation [SD], 2.03); 19 (48%) students were female.

3.2 Knowledge Test

Knowledge test means and SDs for both courses and at each measurement period are indicated in Table 1.

Table 1. Means and standard deviations in knowledge test scores for Courses 1 and 2

Measurement period	N	Mean score	Standard Deviation
COURSE 1*			
Pre-course	12	6.67	1.44
Post-course	11	7.73	1.19
Follow-up	12	7.83	1.19
COURSE 2†			
Pre-course	27	1.93	0.55
Post-course	27	2.78	0.42
Follow-up	24	2.25	0.61

* The highest possible score for Course 1 is 9.

† The highest possible score for Course 2 is 3.

3.2.1 Course 1

Eleven students from Course 1 (92%) completed the knowledge test at all 3 measurement periods. One-way repeated measures ANOVA revealed a significant difference in scores, Wilks’ Lambda=0.42, F(2, 9)=6.34, p=0.02, effect size (partial eta squared)=0.59. Pairwise comparisons with Bonferroni adjustment show post-course and follow-up scores were significantly higher than pre-course scores (Table 2).

3.2.2 Course 2

Twenty-four students from Course 2 (89%) completed the knowledge test at all 3 measurement periods. One-way repeated measures ANOVA revealed a significant difference in scores, Wilks’ Lambda=0.43, F(2, 22)=14.667, p<0.0005, effect size (partial eta squared)=0.57. Pairwise comparisons with Bonferroni adjustment show post-course scores were significantly higher than pre-course and follow-up scores (Table 2).

Table 2. Change in knowledge test scores for Courses 1 and 2

	Mean Difference	p (two-tailed)*	Effect size (eta squared)†,‡
COURSE 1			
Pre-course -vs- Post-course	-1.18	0.022	0.53
Pre-course -vs- Post-course	-1.17	0.011	0.58
Pre-course -vs- Post-course	-0.18	1.000	-
COURSE 2			
Pre-course -vs- Post-course	-0.85	<0.0005	0.56

Pre-course -vs- Post-course	-0.33	0.172	-
Pre-course -vs- Post-course	0.50	0.008	0.33

* Adjustment for multiple comparisons: Bonferroni

[Ɨ] Reported only for pairs reaching statistical significance

[‡] Cohen's (1988) guidelines suggest values of 0.01 indicate a small effect, 0.06 a moderate effect, and 0.14 a large effect.

3.3 Self-reported Knowledge, Confidence, and Skill

Thirty-five students (90%) completed the self-reported knowledge, confidence, and skill questions both post-course and at follow-up. As shown in Table 3, post-course and follow-up scores were universally high (>7), though follow-up scores were significantly lower for all three domains (knowledge, confidence, and skill). Self-reported skill had the largest decline, with a mean score of 8.23 (SD=1.50) post-course and 7.03 (SD=1.60) at follow-up.

Table 3. Change in self-reported knowledge, confidence, and skill*

	Mean (SD)	Wilks' Lamda	F	p-value	Effect size (partial eta squared)[†]
Knowledge	Post-course: 8.29 (1.32) Follow-up: 7.69 (1.62)	0.85	$5.92_{1,34}$	0.02	0.15
Confidence	Post-course: 8.09 (1.58) Follow-up: 7.31 (1.62)	0.79	$8.83_{1,34}$	0.005	0.21
Skill	Post-course: 8.23 (1.50) Follow-up: 7.03 (1.60)	0.59	$23.28_{1,34}$	<0.0005	0.41

SD = standard deviation

* Measured using 10-point scale, ranging from strongly disagree (1) to strongly agree (10)

[†] Cohen's (Cohen, 1988) guidelines suggest values of 0.01 indicate a small effect, 0.06 a moderate effect, and 0.14 a large effect.

3.4 Objective Skill Assessment

Twelve students (100%) from Course 1 completed the post course skill assessment for lumbar puncture and thoracentesis, and 11 students (92%) completed the post-course skill assessment for central line placement. Nine students (75%) from Course 1 completed the follow-up skill assessment for each of the 3 procedures.

3.4.1 Central Line

There was no significant difference between students' post course time and follow-up time for central line placement; likewise, the change in the number of needle redirects post course and at follow-up did not significantly differ (Table 4).

Students' mean post-course time for central line placement was significantly longer than faculty's. The magnitude of the difference was large. Four of 11 students (36%) had a post-course time exceeding 2 SDs beyond that of faculty's (Table 5).

Students' mean follow-up time for central line placement was significantly longer than faculty's. The magnitude of the difference was large. Seven of 9 students (78%) had a follow-up time exceeding 2 SDs beyond faculty's (Table 5).

While all 11 students (100%) correctly followed the sequence of ordered steps for central line post course, only 3 of 9 (33%) did so at follow up. The odds of correctly following the sequence of ordered steps at follow-up were 98% less than post-course (OR=0.02, 95% CI=0.001 to 0.53, z=2.362, p=0.018).

3.4.2 Lumbar Puncture

Students' post course time was significantly shorter than follow-up for lumbar puncture; likewise, the change in the number of needle redirects significantly increased from post course to follow-up (Table 4).

Students' post-course time for lumbar puncture did not significantly differ from faculty's (Table 5). One of 12 students (8%) had a post-course time exceeding 2 SDs beyond faculty's.

Students' follow-up time for lumbar puncture was significantly longer than faculty's. Five of 9 students (56%) had a follow-up time exceeding 2 SDs beyond faculty's (Table 5).

All 12 students (100%) correctly followed the sequence of ordered steps for lumbar puncture post-course, and 7 of 9 (78%) did so at follow-up. The odds of correctly following the sequence of ordered steps for lumbar puncture at follow-up were not statistically different than post-course (OR=0.12, 95% CI=0.01 to 2.85, z=1.31, p=0.190).

3.4.3 Thoracentesis

Students' post course time did not significantly differ from follow-up time for thoracentesis; likewise, the number of needle redirects post course did not significantly differ from follow-up (Table 4).

Students' post-course time for thoracentesis did not significantly differ from faculty's (Table 5). One of 12 students (8%) had a post-course time exceeding 2 SDs beyond faculty's.

Students' follow-up time for thoracentesis did not significantly differ from faculty's. One of 9 students (11%) had a time exceeding 2 SDs beyond faculty's.

Seven of 12 students (58%) correctly followed the sequence of ordered steps for thoracentesis post-course, and 3 of 9 (33%) did so at follow-up. The odds of correctly following the sequence of ordered steps at follow-up were not significantly different than at post-course (OR=0.36, 95% CI=0.06 to 2.16, z=1.22, p=0.262).

Table 4. Comparison of student post-course and follow-up time to complete procedures and the number of needle redirects*

	Mean post-course (SD)	Mean follow-up (standard deviation)	t (df)	p (two-tailed)	Effect size (eta squared)[†,‡]
CENTRAL LINE					
Time	4.11 (0.59)	4.25 (0.53)	-0.49 (8)	0.64	-
Needle redirects	1.11 (1.27)	1.22 (1.39)	-0.22 (8)	0.83	-
LUMBAR PUNCTURE					
Time	2.86 (0.74)	4.46 (1.04)	-3.60 (8)	0.007	0.62
Needle redirects	0.25 (0.62)	3.00 (2.69)	-3.67 (8)	0.006	0.63
THORACENTESIS					
Time	2.60 (0.91)	2.73 (0.53)	-0.71 (8)	0.50	-
Needle redirects	1.44 (2.79)	0.11 (0.33)	1.41 (8)	0.20	-

NOTE: Because time to complete the procedures in most cases was not normally distributed, time (in seconds) was transformed to its log.

* Missing cases excluded pairwise

[†] Reported only for pairs reaching statistical significance

[‡] Cohen's (1988) guidelines suggest values of 0.01 indicate a small effect, 0.06 a moderate effect, and 0.14 a large effect.

Table 5. Comparison of student time to complete procedures against faculty reference time*

	Mean student time (SD)	Mean faculty time (SD)	t (df)	p (two-tailed)	Effect size (Cohen's d) [§,**]	Number of student outliers (%)[††]
CENTRAL LINE						
Post-course	4.07 (0.54)	3.38 (0.41)[†]	4.24 (10)	0.002	1.28	4 (36)
Follow-up	4.25 (0.53)		4.93 (8)	0.001	1.64	7 (78)
LUMBAR PUNCTURE						

Post-course	2.74 (0.71)	2.84 (0.63)†	-0.46 (11)	0.65	-	1 (8)
Follow-up	4.46 (1.04)		4.68 (8)	0.002	1.56	5 (56)
THORACENTESIS						
Post-course	2.44 (0.85)	2.55 (0.50)‡	-0.44 (11)	0.67	-	1 (8)
Follow-up	2.73 (0.53)		1.04 (8)	0.33	-	1 (11)

NOTE: Because time to complete the procedures in most cases was not normally distributed, time (in seconds) was transformed to its log.

* Missing cases excluded pairwise

† Faculty sample size = 3

‡ Faculty sample size = 4

§ Reported only for pairs reaching statistical significance

** Cohen's (1988) guidelines suggest values of 0.2 indicate a small effect, 0.5 a moderate effect, and 0.8 a large effect.

†† An outlier is defined as a student time that is ≥2 standard deviations from the faculty reference time.

4.1 Discussion

The objective of this study was to measure the effect of skill-based clinical simulation training on second-year medical students' knowledge, confidence, and skill between baseline and 6-month follow-up. We found that students' scores on knowledge tests improved from pre-course to post-course for Course 1 and Course 2. Test scores were also higher at follow-up when compared to pre-course, though the gain fell short of statistical significance for Course 2 ($p=0.172$). Students self-reported knowledge, confidence, and skill was universally high both post-course and at follow-up, though scores were significantly lower for all three measures at follow-up. The largest decline was observed for self-reported skill. One explanation for this finding may stem from the hands-on and observable nature of skill-based performance. Whereas knowledge and confidence reflects a person's internal state of mind, skill is an externally observable trait that requires purposeful interaction with one's environment. Because skill can be readily witnessed by others, the criteria for self-assessing skill may be qualitatively different. Notably, previous research suggests self-perceived confidence and skill—ie, self-assessment—may not be an accurate predictor clinical skill, thus requiring the need for objective skill-based measures.(Barsuk, et al., 2010; Davis et al., 2006; Wayne, Butter, et al., 2006)

For both central line and thoracentesis simulations, students' time to complete the procedure and number of needle redirects did not differ between post-course and follow-up. This finding is consistent with previous research assessing learners in graduate and post-graduate medical education settings.(Crofts, et al., 2007; Wayne, Siddall, et al., 2006) In contrast, students' time and needle redirects for lumbar puncture worsened at follow-up. It is not clear why skills degraded for this single simulation only, though it may be related to the procedure's relative difficulty.

Students were significantly less likely to correctly follow the sequence of ordered steps at follow-up for the central line simulation only. The observed decline may be a function of the multiple steps inherent in this procedure—whereas both lumbar puncture and thoracentesis had 5 steps assessed, 8 steps were assessed for central line.

Compared to faculty, students' time to complete the simulations took longer for only half of the measurements: central line (post-course and follow-up) and lumbar puncture (follow-up). This finding is quite remarkable given that these 2nd year medical students have had little to no previous experience with the procedures or with using simulation mannequins.

An important next step for this line of research would be to measure the effect of the training on student performance in years 3 and 4 when working with real patients. Future research would also benefit from having a larger sample, not only to increase power but to enable use of more sophisticated multivariate techniques that control for potential confounders.

Several limitations to this study warrant discussion. First, the number of participating students was small. This was especially true for Course 1, where only 75% participated in the 6-month follow-up skill assessment. Despite the small sample, findings were quite robust: subsequent analyses using non-parametric statistical methods (not shown) resulted in essentially identical findings. Second, while the knowledge test was developed by seasoned clinical faculty using an iterative and consensus-based process, the instruments were course specific and thus may not be generalizable to all simulation-based teaching environments. Third, findings may have been inflated due to testing or maturation bias. To minimize this potential threat, correct responses were never disclosed to students, and students were not given access to the instruments between measurement periods. Fourth, while efforts were made to secure 100% participation—including a small financial incentive (up to $25 total)—a number of students declined to participate at follow-up. The reasons for dropping out were not known; however, several students indicated conflicts stemming from exams, travel, and studying for Step 1 of the USMLE. And fifth, pooling results for some analyses may have introduced error, as both the instructors and content varied between courses. To minimize this risk, each course was taught by faculty from the same institution and department, the courses were structured similarly, the courses adhered to the 4 essential constructs described by Kneebone (2005), and instructors were purposeful in keeping the hands-on and skill-based training the course's primary focus.

Conducting medical procedures is an integral component to the practice of many medical specialties; however, assessing students' skill—and the maintenance of that skill over time—using objective and empirically derived measures can be challenging. The family medicine clinical simulation courses for second-year medical students' at the UM employed simple yet robust methods to assess students' knowledge, confidence, and skill over time. Given their relative simplicity, these methods could be adapted to many procedures commonly practiced within primary care and other medical specialties.

Findings showed that brief exposure to clinical simulation early in one's medical training and prior to clerkships can have a positive and lasting effect on knowledge, confidence, and skill. We anticipate this early exposure will serve as a primer, helping medical students maximize their learning during clerkships years.

References

Ahya, S. N., Barsuk, J. H., Cohen, E. R., Tuazon, J., McGaghie, W. C., & Wayne, D. B. (2012). Clinical performance and skill retention after simulation-based education for nephrology fellows. Semin Dial, 25(4), 470-473. doi: 10.1111/j.1525-139X.2011.01018.x

Barsuk, J. H., Cohen, E. R., McGaghie, W. C., & Wayne, D. B. (2010). Long-term retention of central venous catheter insertion skills after simulation-based mastery learning. Acad Med, 85(10 Suppl), S9-12. doi: 10.1097/ACM.0b013e3181ed436c

Bradley, P. (2006). The history of simulation in medical education and possible future directions. Med Educ, 40(3), 254-262. doi: 10.1111/j.1365-2929.2006.02394.x

Chou, D. S., Abdelshehid, C., Clayman, R. V., & McDougall, E. M. (2006). Comparison of results of virtual-reality simulator and training model for basic ureteroscopy training. [Comparative Study Randomized Controlled Trial]. J Endourol, 20(4), 266-271.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: L. Erlbaum Associates.

Consoli, A., Fraser, K., Ma, I., Sobczak, M., Wright, B., & McLaughlin, K. (2012). Diagnostic performance 1 h after simulation training predicts learning. Adv Health Sci Educ Theory Pract, 18(5), 893-800. doi: 10.1007/s10459-012-9431-7

Crofts, J. F., Bartlett, C., Ellis, D., Hunt, L. P., Fox, R., & Draycott, T. J. (2007). Management of shoulder dystocia: skill retention 6 and 12 months after training. Obstet Gynecol, 110(5), 1069-1074. doi: 10.1097/01.aog.0000286779.41037.38

Davis, D. A., Mazmanian, P. E., Fordis, M., Van Harrison, R. R., Thorpe, K. E., & Perrier, L. (2006). Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. JAMA, 296(9), 1094-1102. doi: 10.1001/jama.296.9.1094

Didwania, A., McGaghie, W. C., Cohen, E. R., Butter, J., Barsuk, J. H., Wade, L. D., . . . Wayne, D. B. (2011). Progress toward improving the quality of cardiac arrest medical team responses at an academic teaching hospital. J Grad Med Educ, 3(2), 211-216. doi: 10.4300/jgme-d-10-00144.1

Fraser, K., Peets, A., Walker, I., Tworek, J., Paget, M., Wright, B., & McLaughlin, K. (2009). The effect of simulator training on clinical skills acquisition, retention and transfer. Med Educ, 43(8), 784-789. doi: 10.1111/j.1365-2923.2009.03412.x

Fraser, K., Wright, B., Girard, L., Tworek, J., Paget, M., Welikovich, L., & McLaughlin, K. (2011). Simulation training improves diagnostic performance on a real patient with similar clinical findings. Chest, 139(2), 376-381. doi: 10.1378/chest.10-1107

Goulart, J. M., Dusza, S., Pillsbury, A., Soriano, R. P., Halpern, A. C., & Marghoob, A. A. (2012). Recognition of melanoma: A dermatologic clinical competency in medical student education. J Am Acad Dermatol, 67(4), 606-611. doi: http://dx.doi.org/10.1016/j.jaad.2011.12.007

Halm, B. M., Lee, M. T., & Franke, A. A. (2010). Improving medical student toxicology knowledge and self-confidence using mannequin simulation. Hawaii Med J, 69(1), 4-7.

Heitz, C., Brown, A., Johnson, J. E., & Fitch, M. T. (2009). Large group high-fidelity simulation enhances medical student learning. [Research Support, N.I.H., Extramural

Research Support, Non-U.S. Gov't]. Med Teach, 31(5), e206-210.

Issenberg, S. B., McGaghie, W. C., Petrusa, E. R., Gordon, D. L., & Scalese, R. J. (2005). Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. [Research Support, Non-U.S. Gov't

Research Support, U.S. Gov't, Non-P.H.S.

Review]. Medical Teacher, 27(1), 10-28.

Karabilgin, O. S., Vatansever, K., Caliskan, S. A., & Durak, H. I. (2012). Assessing medical student competency in communication in the pre-clinical phase: objective structured video exam and SP exam. Patient Educ Couns, 87(3), 293-299.

Kelley, S. A., & Cooke, J. M. (Society of Teachers of Family Medicine, 2013). Simulated procedures in family medicine in a pre-clinical medical student curriculum Retrieved 19 August, 2014, from http://www.fmdrl.org/index.cfm?event=c.beginBrowseD&clearSelections=1&criteria=kelley#4384

Kelly, L. P., Garza, P. S., Bruce, B. B., Graubart, E. B., Newman, N. J., & Biousse, V. (2013). Teaching ophthalmoscopy to medical students (the TOTeMS study). [Randomized Controlled Trial

Research Support, N.I.H., Extramural]. Am J Ophthalmol, 156(5), 1056-1061. doi: http://dx.doi.org/10.1016/j.ajo.2013.06.022

Kneebone, R. (2005). Evaluating clinical simulations for learning procedural skills: a theory-based approach. Acad Med, 80(6), 549-553.

Lee, J. T., Qiu, M., Teshome, M., Raghavan, S. S., Tedesco, M. M., & Dalman, R. L. (2009). The utility of endovascular simulation to improve technical performance and stimulate continued interest of preclinical medical students in vascular surgery. [Comparative Study

Research Support, Non-U.S. Gov't]. J Surg Educ, 66(6), 367-373. doi: http://dx.doi.org/10.1016/j.jsurg.2009.06.002

McGaghie, W. C., Issenberg, S. B., Barsuk, J. H., & Wayne, D. B. (2014). A critical review of simulation-based mastery learning with translational outcomes. Med Educ, 48(4), 375-385. doi: 10.1111/medu.12391

McGaghie, W. C., Issenberg, S. B., Cohen, E. R., Barsuk, J. H., & Wayne, D. B. (2011). Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. Acad Med, 86(6), 706-711. doi: 10.1097/ACM.0b013e318217e119

McGaghie, W. C., Issenberg, S. B., Petrusa, E. R., & Scalese, R. J. (2010). A critical review of simulation-based medical education research: 2003-2009. Med Educ, 44(1), 50-63. doi: 10.1111/j.1365-2923.2009.03547.x

Mittal, M. K., Morris, J. B., & Kelz, R. R. (2011). Germ simulation: a novel approach for raising medical students awareness toward asepsis. [Comparative Study

Research Support, Non-U.S. Gov't]. Simul Healthc, 6(2), 65-70.

Passiment, M., Sacks, H., & Huang, G. (2011). Medical simulation in medical education: results of an AAMC Survey Retrieved 14 March, 2014, from https://www.aamc.org/download/259760/data/

Schubart, J. R., Erdahl, L., Smith Jr, J. S., Purichia, H., Kauffman, G. L., & Kass, R. B. (2012). Use of

Breast Simulators Compared with Standardized Patients in Teaching the Clinical Breast Examination to Medical Students. J Surg Educ, 69(3), 416-422. doi: http://dx.doi.org/10.1016/j.jsurg.2011.10.005

van Zuilen, M. H., Kaiser, R. M., & Mintzer, M. J. (2012). A competency-based medical student curriculum: taking the medication history in older adults. [Comparative Study
Research Support, Non-U.S. Gov't]. J Am Geriatr Soc, 60(4), 781-785.

Wayne, D. B., Butter, J., Siddall, V. J., Fudala, M. J., Wade, L. D., Feinglass, J., & McGaghie, W. C. (2006). Graduating internal medicine residents' self-assessment and performance of advanced cardiac life support skills. Med Teach, 28(4), 365-369. doi: 10.1080/01421590600627821

Wayne, D. B., Siddall, V. J., Butter, J., Fudala, M. J., Wade, L. D., Feinglass, J., & McGaghie, W. C. (2006). A longitudinal study of internal medicine residents' retention of advanced cardiac life support skills. Acad Med, 81(10 Suppl), S9-S12.

Copyright Disclaimer