Practice improves performance on a coronary anastomosis simulator, attending surgeon supervision does not

Daniel H. Enter, MD,[a] Xiaoying Lou, BS,[a] Dawn S. Hui, MD,[b] Adin-Cristian Andrei, PhD,[a] Hendrick B. Barner, MD,[b] Luke Sheen, BA,[a] and Richard Lee, MD, MBA[b]

Objectives: Enthusiasm for simulation early in cardiothoracic surgery training is growing, yet evidence demonstrating its utility is limited. We examined the effect of supervised and unsupervised training on coronary anastomosis performance in a randomized trial among medical students.

Methods: Forty-five medical students were recruited for this single-blinded, randomized controlled trial using a low-fidelity simulator. After viewing an instructional video, all participants attempted an anastomosis. Subsequently, the participants were randomized to 1 of 3 groups: control (n = 15), unsupervised training (n = 15), or supervised training with a cardiothoracic surgeon or fellow (n = 15). Both the supervised and unsupervised groups practiced for 1 hour per week. After 4 weeks, the participants repeated the anastomosis. All pre- and post-training performances were videotaped and rated independently by 3 cardiothoracic surgeons blinded to the randomization. All raters scored 13 assessment items on a 1 to 5 (low-high) scale along with an overall pass/fail rating.

Results: After the training period, all 3 groups showed significant improvements in composite scores (control: +0.52 ± 0.69 [$P = .014$], unsupervised: +1.05 ± 0.48 [$P < .001$], and supervised: +1.10 ± 0.84 [$P < .001$]). Compared with control group, both supervised ($P = .005$) and unsupervised trainees ($P = .005$) demonstrated a significant improvement. Between the supervised and unsupervised groups there were no statistically significant differences in composite scores.

Conclusions: Practice on low-fidelity simulators enabled trainees to improve on a broad range of skills; however, the additional effect of attending-level supervision is limited. In an era of increasing staff surgeon responsibilities, unsupervised practice may be sufficient for inexperienced trainees. (J Thorac Cardiovasc Surg 2015;149:12-7)

See related commentary on page 18.

Supplemental material is available online.

Simulation training is now a mandated component of cardiothoracic (CT) surgical training,[1] with a variety of low- and high-fidelity models available.[2,3] A tool developed by the Joint Council on Thoracic Surgery Education (JCTSE) for evaluating competence in coronary anastomosis performance has shown high interrater reliability and internal consistency for simulation models,[4] and is being used to assess simulation training effectiveness. Under the sponsorship of the JCTSE and Thoracic Surgery Directors Association, a nationwide invitation to CT resident physicians to participate in a simulation training competition culminated in recognition of the top winners at the 2013 American Association of Thoracic Surgeons Annual Conference.[5] However, an optimal method of teaching CT surgical tasks has yet to be demonstrated. We conducted a randomized study of medical students' performance of a coronary anastomosis on a low-fidelity simulator to determine the effects of supervision and practice.

From the Division of Cardiac Surgery,[a] Bluhm Cardiovascular Institute of Northwestern University, Chicago, Ill; and Center for Comprehensive Cardiovascular Care,[b] Saint Louis University, St Louis, Mo.

Disclosures: Authors have nothing to disclose with regard to commercial support.

Read at the 94th Annual Meeting of The American Association for Thoracic Surgery, Toronto, Ontario, Canada, April 26-30, 2014.

Received for publication April 20, 2014; revisions received Aug 16, 2014; accepted for publication Sept 5, 2014; available ahead of print Oct 22, 2014.

Address for reprints: Daniel H. Enter, MD, Division of Cardiac Surgery, Bluhm Cardiovascular Institute of Northwestern University, 201 E Huron St, Galter Pavilion, Ste 11-140, Chicago, IL 60611-2968 (E-mail: denter@northwestern.edu).

0022-5223/$36.00

http://dx.doi.org/10.1016/j.jtcvs.2014.09.029

MATERIALS AND METHODS

Participants

Forty-five medical students were recruited through a school-wide listserv E-mail, with all consenting respondents accepted until the sample size to achieve adequate statistical power was obtained. The primary end point was composite score after the training period. With 15 participants per group there is 80% power to detect a change of 1 standard deviation unit in the primary end point between the 3 groups at a 2-sided 5% type I error level. No compensation was offered and approval was granted from the Institutional Review Board at Northwestern University. The participant

Abbreviations and Acronyms

CT = cardiothoracic

JCTSE = the Joint Council on Thoracic Surgery Education

pool had a mean age of 23 ± 1.7 years, was 51.1% female, and the majority were first-year students (66.7% M1, 28.9% M2, and 4.4% M3).

Study Design

The task consisted of a simulated end-to-side coronary anastomosis using a low-fidelity simulator. All subjects viewed a short video demonstrating the steps of the task, after which they performed a task attempt that was video-recorded. Each of the 13 components of the JCTSE assessment tool were correctly illustrated and viewing was unrestricted. Using portable synthetic tube task stations, each participant was to anastomose a 5-mm synthetic tube (Limbs & Things, Savannah, Ga) in an end-to-side fashion onto a similar tube mounted on a plastic board (W.L. Gore & Associates, Inc, Flagstaff, Ariz). Surgical instruments were similar to those used in the operating room, and suture used was 5-0 polypropylene. All participants were administered a questionnaire documenting demographics, prior surgical experience, career plans, and level of interest in surgery, CT surgery, and simulation. Responses were measured using a 5-point Likert scale (where 1 = strongly disagree and 5 = strongly agree).

Students were then randomized in a 1:1:1 manner to a control group, unsupervised training group, or supervised training group after they completed the initial training session (Figure 1). The control group underwent no further training. The unsupervised training group was asked to perform 1 hour of training over 4 different sessions over a period of 1 month. This training consisted of anastomotic task completion on a provided at-home low-fidelity model. The supervised training group spent the same amount of time performing the task, but under the instruction of a cardiac surgeon or resident physician. Direct, constructive, verbal feedback was provided at each session for the supervised training group with a ratio of 3 to 4 students per instructor. At the conclusion of the 4-week training period, all participants returned for a posttest session where they were recorded performing an anastomosis and again completed the questionnaire.

These recorded anastomoses were then rated in a blinded fashion by 3 senior cardiac surgeons utilizing the JCTSE Assessment Tool. Each video recording was anonymous as to the identity of the participant. The raters were unrestricted in their time to grade the videos. The raters did not have direct, physical access to the simulated anastomoses performed. This rating scale consists of 13 assessment items on a 1 to 5 (low-high) scale as well as an overall pass/fail rating (Figure E1). The assessment items were averaged into a composite score, from which the arteriotomy skill was removed due to lack of rater agreement. In addition, the participants rated their own anastomoses performance using this scale at the pre- and posttraining sessions.

Statistical Analysis

Generalized estimating equations were used to compare the primary end point of rater-based pre- and posttraining composite score and pass/fail rates among the control, unsupervised practice, and supervised practice groups. Self-evaluations in pre- and posttraining composite score and pass/fail rates were assessed in similar fashion. The relationship between interest and training group was also modeled as a function of type of training (group), rater and the respective pretraining score using a 3-dimensional posttraining response (1 score per rater). Statistical significance was established at the 2-sided 5% alpha level and there were no adjustments for multiplicity. Composite scores are reported as mean ± standard deviation.

RESULTS

Performance Scores by Senior Rater Evaluation

All 3 groups demonstrated improvement in senior-rated composite scores (Table 1 and Figure 2); control: 1.93 to 2.45 (+0.52 ± 0.69 [P = .014]), unsupervised: 2.04 to 3.09 (+1.05 ± 0.48 [P < .001]), and supervised: 2.05 to 3.15 (+1.10 ± 0.84 [P < .001]). Both supervised (P = .005) and unsupervised trainees (P = .005) demonstrated a significant improvement compared with the control group. Both practice groups demonstrated significant improvement across the majority of the technical skills categories (Table 1). Significantly higher pass rates occurred in the supervised practice group (55%) and unsupervised practice group (58%) compared with control (36%; P = .005). Between the supervised and unsupervised groups there were no statistically significant differences in composite individual task, or pass/fail scores. The interrater agreements for pretraining and posttraining composite scores were 0.92 and 0.71, respectively, demonstrating good agreement between raters.

Performance Scores by Participant Self-Evaluation

Across groups and at both pre- and posttraining, participant self-rated composite scores were similar to the senior-rater scores (Figure 3). In pass–fail assessments, participant-rated pass rates were higher than senior-rated pass rates across all groups. This difference was greatest in the final evaluation of the supervised practice group, where 100% of participants determined themselves to have passed, versus a 55% pass rate by senior raters (Table 2). The gap between participant-rated and senior-rated final pass rates was smaller in the unsupervised (69% vs 58%) and control groups (53% vs 36%).

Interest in Surgery

Previously published data from this medical student cohort demonstrated that supervision led to an increase in interest in surgery (P < .028) compared with the control group.[6] We compared participants in the top 50% and bottom 50% of improvement (defined as initial score subtracted from final score), to determine if there was a relationship between score improvement and change in interest in a career in surgery. We found no relationship between improvement in score and interest in surgery (P = .47) based on the raters' evaluations. However, an increased interest in a career in surgery was found among the top 50% improvement group when self-evaluation scores were used (P < .048).

DISCUSSION

This prospective, randomized study with blinded expert raters demonstrated that practice on coronary simulators improves performance across a broad range of skills.