Synthetic Material Simulation Improves Performance of Laparoscopic Cholecystectomy in an Animate Model


Correspondence to:

Mika N. Sinanan, M.D., Ph.D.
University of Washington, Department of Surgery
Box 356410
1959 N.E. Pacific Street
Seattle, WA 98195-6410

This project was supported in part by grants from SAGES and United States Surgical Corporation.

Abstract

Background: Simulation for training videoendoscopic surgical skills has become increasingly important. While animate models used to predominate for simulating procedures has, we have developed several novel, synthetic material simulators that preserve functionality without the ethical issues and cost of animal surgery. The benefits of task and procedure practice seem clear, yet objective measure of simulation effects on clinical performance and the learning curve for technical skills have not been demonstrated.

Methods: Twenty-eight, surgically naive medical students were oriented to laparoscopic instruments and to the technique for laparoscopic cholecystectomy. They were then divided into three groups. Nine students had no training with simulation (S-0 group), 10 performed two cholecystectomy simulations (S-2 group), and another 9 performed repeated cholecystectomy simulations (S-X) until each individual reached a plateau on the learning curve for the simulator. The subsequent performance of each student during a porcine laparoscopic cholecystectomy was evaluated. Performance was measured with a composite score incorporating time, errors, and instructor coaching.

Results: The students who had reached the plateau in the learning curve performed significantly better clinically (S-X group, mean score 1873± 361) than the students without simulation (S-0 group, mean score 3513± 1141, p<0.001) or the students with only two simulations (S-2, mean score2780± 1284, p<0.05). Two simulations were not sufficient to change clinical performance (S-0 vs. S-2, p= 0.2). In the S-X group, significant improvement occurred by the third repetition but a plateau in the learning curve was not reached for all subjects until the ninth repetition. Regardless of initial performance, all subjects (S-X group) eventually reached the same level of performance (mean initial score 2314± 590 vs mean plateau 860± 223).

Conclusion: With adequate repetition to traverse the learning curve, training in a simulator improves videoendoscopic clinical skills.
Technical skills in the field of surgery have been taught using the apprenticeship model for the last 100 years. The advent and popularization of minimally invasive surgery brought about the use of simulation as a training tool. Videoendoscopic techniques are especially suited for simulation because the field of view is limited, peripheral vision is not possible, and the operative field is represented in two dimensions on a video monitor. Furthermore, the mentoring surgeon has a limited ability to direct the operation. Indeed, unlike open procedures, where the mentor may employ verbal and manual direction from the opposite side of the operating table, in videoendoscopic surgery verbal direction is the only basis for instruction and there is an increased reliance on the technical skills of the trainee. To achieve these basic skills, training in simulated operative environments has become increasingly common despite little evidence regarding the efficacy of these techniques.

Simulation may be divided into four main categories—skill stations, synthetic material, animal models, and computer environments. Skill stations have been designed to train on specific functions with laparoscopic instruments in a two-dimensional environment. The skills practiced in such simulators usually represent a component of a more complex procedure performed in the operating room. Synthetic material simulation uses man-made material to approximate tissues and anatomic relationships found clinically. Animate model simulation is based on finding an animal model similar to human anatomy for the respective procedure. This approach was used widely with the introduction of laparoscopic cholecystectomy. Computer environments vary in complexity from simple skill station tasks to anatomical recreations of common operative fields complete with realistic tissue deformation and force feedback.

Each of these simulation models has differing attributes as a training tool. Skill stations are inexpensive. They are easy to use, but at best, they only train the student in one limited aspect of the operation. Animals may provide an excellent model, but ethical and cost considerations require that they be used at a more advanced stage of the learning process. Computer environments are free of these ethical considerations but tend to be expensive and are currently incompletely developed. Synthetic material is robust, relatively inexpensive, may be similar to the operative environment and avoids ethical considerations.

Prior studies assessing the clinical impact of simulation training have been complicated because these skills were incorporated into a broader educational program in individuals of varying experience. Thus, objective improvement in the trainee's ability to perform a clinical skill after training on simulation has not been clearly demonstrated. Since we had developed a standardized method for training residents to perform laparoscopic cholecystectomy (Seven-Step Protocol for Laparoscopic Cholecystectomy, Ciné-Med, Woodbury, CT) using a synthetic model, we thought it would be important to determine, to the extent possible, the real impact that our method had on the ability of the residents to perform the operation being
taught. In order to avoid potential confounding errors in our measurement from other operative experience the trainees may have obtained, we chose a group of individuals who had not had any previous training. We used our model as the only training module to learn laparoscopic cholecystectomy and then measured the technical improvement achieved by determining the ability of these trainees to perform a laparoscopic cholecystectomy in the pig. These results were then compared to those obtained in a group of controls. The purpose of this study was to determine if synthetic material simulation could be used to improve the ability of an inexperienced, laparoscopic trainee to perform the analogous procedure in an animate model.

**Materials and Methods:**

**The simulation model**

Laparoscopic cholecystectomy and cystic duct cannulation was simulated using a synthetic material model from Simulab Corporation (Seattle, Washington). The base formed a hard plastic replication of the stomach and the liver. The gallbladder, cystic duct, cystic artery, common bile duct and surrounding connective tissue were made of latex and were provided as a single, replaceable unit that could be attached to the gallbladder bed of the liver with Velcro™ bands. The complete model was inserted into a nontransparent human torso model with an anterior abdominal wall made of rubber, which could be penetrated by laparoscopic trocars and ports (Figure 1). Procedures on the simulated gallbladder were performed using standard laparoscopic videoendoscopic equipment (Stryker, San Jose, CA; Storz, Goleta, CA) and instruments (US Surgical Corp., Norwalk, CT).

For the animate model, a laparoscopic cholecystectomy was performed in a standard fashion on pigs (60-80 pounds) under general anesthesia. Cystic duct cannulation was not attempted in the animate models. The protocol for animal use was approved by the Animal Care Committee of the University of Washington.

**The scoring system**

Performance of the test and control subjects in both model systems was scored in a similar fashion. Time required for the completion of the procedure was measured in seconds and additional 30-second penalties were added for each required or requested coaching event, and for each error (i.e. inadvertent entry into the gallbladder). Coaching varied between instructors as some opted to coach a student earlier, to prevent the commission of an error whereas others permitted the error to happen. Both error and coach penalties were weighed identically, so that the discrepancy evened out. The sum of time and penalty seconds resulted in a score. Time and score are reported as mean seconds ± SD.
The study groups

Twenty-eight fourth-year medical student volunteers were recruited and instructed in the technique of laparoscopic cholecystectomy with videotapes, which explained the anatomy, the instruments, and the steps of the procedure. The use of the laparoscopic instruments was then demonstrated and the students were given time to practice with instruments until they felt comfortable. The students were randomized into three groups. In the first group, 9 students received no additional training and proceeded with a laparoscopic cholecystectomy in the pig (S-0 group). In the second group, 10 students performed two laparoscopic cholecystectomy simulations with the synthetic material model and then proceeded to the laparoscopic cholecystectomy in the pig (S-2 group). In the third group, 9 students performed as many laparoscopic cholecystectomies in the synthetic material simulation model as were necessary to identify a plateau in the learning curve of each individual (S-X group). The plateau of the learning curve was defined as three consecutive total scores with variance ≤ 10%. Once it was apparent that the plateau had been reached, the student performed a minimum of two more simulated laparoscopic cholecystectomies on the following day to verify stability of their scores and then proceeded to the laparoscopic cholecystectomy in the pig. Coaching in this group was given to facilitate the learning process and not just to prevent errors.

Human Subjects Committee approval was obtained for the involvement of medical students in this study.

Statistics

Statistical analysis was performed with the two-tailed t-test with the level for statistical significance set at p<0.05.

Results

Animate model performance

The final scores for the animate model performance are shown in Table 1. The S-X group performed significantly better than the S-0 group, with mean scores of 1873± 361 and 3513± 1141 sec, respectively (p<0.001). The S-X group also performed significantly better than the S-2 group (1873± 361 versus 2780± 1284 sec) (p<0.05). Performance after only two simulations (S-2 group) seemed better than in the S-0 group (2780± 1284 versus 3513± 1141 sec), but this did not achieve statistical significance (p=0.2)

The mean number of errors and coaching events are also shown in Table 1. There was no difference between any of the groups.
Simulation model performance

In the S-2 group an improvement in the mean time and score was seen from the first simulation with a time of 1457 (± 379) and a score of 1562 (± 449) seconds to the second simulation with a mean time of 1355 (± 581) and a mean score of 1397 (± 601) seconds, but this was not significant (Table 2). Errors and coaching events did not differ between the simulations.

The S-2 and the S-X groups were comparable at the onset of the experiment with regards to time (1457± 379 sec versus 1860± 699 sec, p=0.15) and error values (1.4 versus 2.8, p= 0.26). However, the S-X group experienced a significantly higher level of coaching events by our protocol than the S-2 group (12 versus 2.1, p<0.0001), which raised the score for the first simulation of the S-X group compared to the S-2 group (2314± 590 versus 1562± 449 sec, p<0.01).

Learning curve

The nine medical students in the S-X group performed an average of 10.4 (range 9-12) simulated laparoscopic cholecystectomies before they performed the animate laparoscopic cholecystectomy (Figure 2). Time and score improved significantly from the first (1860 and 2314 sec) to the third simulation (1433 and 1666 sec) in this group (p<0.05). Coaching events subsided significantly after the first simulation (12 versus 8) (p <0.01) and the number of errors decrease significantly from 2.9 to 0.33 by the fifth simulation (p<0.05)(Table 3). The mean time, score, errors and coaching events continued to improve until the ninth simulation when a time of 817 sec and a score of 860 sec were achieved. This improvement from the third to the ninth simulation was also significant for time, score and coaching events (p<0.01). Although the ninth simulation also had fewer errors (2 versus 0.3 times), this was not significant (p=0.59). By the ninth simulation, a plateau in the learning curve was reached. For most subjects, time and score did not differ more than 10% in the subsequent simulations, leading to a final time and score of 775 and 805 seconds respectively (Figures 3 and 4).
**Discussion**

Simulation is routinely incorporated into programs designed to facilitate videoendoscopic technical skills training. The impact of this training is difficult to discern because different types of simulation are often employed during the same course, and clinical performance after the course has not been assessed objectively.

Some of the earliest and most thorough analysis of simulation training showed significant correlation between the performance at three skill stations and the ability to perform a laparoscopic suturing exercise on pig intestine. Similarly, Melvin showed that surgical residents who were evaluated with skill stations before and after a 6 hour suturing course did improve their ability to perform at the skill stations.

More recently, Fried showed through a series of seven simulators that concurrent repetition at multiple skill stations will improve the ability of the participant to perform each of the tasks. In this study, performance at four of the seven tasks was also improved in a control group of participants who had no repetition between testing. This study reveals two important points: repetition of a task will improve the ability to perform the task and some tasks require little repetition to see improvement (i.e. a very rapid learning curve).

Our study employed a complex simulation, which represented an entire procedure rather than a single technical skill. The goal of the study was to evaluate the impact of synthetic material simulation on clinical performance. The same scoring system used for both synthetic and animate model. Each student performed only one cholecystectomy in the animate model, which excluded the possibility of test repetition as a reason for improvement of the animate score.

The effect of repetition on performance was also examined. The evaluation of the 19 participants in the S-0 and S-2 group revealed that two simulation trials might not be sufficient to move the participants to a significantly different position on the learning curve. Although one could speculate that this observation only reflects the relatively small number of measurements and the considerable variation in time and scores between the individuals, we feel that it most likely reflects the fact that two simulations were not enough for the great majority of individuals. Indeed, as a group, their time, errors and score when measured in the animate model, did not differ significantly from those individuals who had had no simulator practice at all. The next set of participants was allowed to perform simulation trials until they were at a plateau in the learning curve. The results of this part of the study have many implications that deserve further analysis. First, and most surprising to us, we found that, when we allowed the trainees to perform as many simulations as needed "to achieve a plateau" (no change greater than 10% in 3 simulations, and no improvement on a subsequent day), they all achieved the same plateau. Despite the fact that, initially, there was significant difference
between trainees, they all eventually the same point. That is, they could all be trained. Some of them achieved the 800-1000 score within four simulations, some required up to nine simulations, but eventually they all reached it. If proven by other methods of training, the significance of this finding is considerable, in that, if the first test had been used to "select" those with best technical skills, an injustice would have been done. Secondly, the fact that, when thoroughly trained, all achieved the same plateau suggests that the model used is an appropriate one, one that gives a certain level of difficulty and one that allows the difficulties to be overcome with experience.

In the other part of this study in which the performance in the animate cholecystectomy was compared between the three groups, only those trainees showed significant improvement in their ability to perform a laparoscopic cholecystectomy on a pig who had had multiple repetitions with the synthetic cholecystectomy simulator. This is the first time that simulation has been shown to improve performance of animate laparoscopic procedures. The learning curve for synthetic material explains why we did not observe an improvement in the animate cholecystectomy scores for the S-2 group. Statistical improvement in scores did not occur in the S-X group until after the third repetition. This is similar to previous reports on the learning curve using animate models in which the third repetition of laparoscopic cholecystectomy in a pig model yielded significant improvement of gallbladder fossa dissection. In that study, the plateau of the learning curve was not attained at the end of the third procedure. Similarly, the full benefit of the synthetic material simulation was not realized until the ninth repetition.

The final scores in the first two simulation trials for the S-2 group were significantly better than the scores in the first two simulation trials for the S-X group. The explanation for this is found in the coaching for the S-X group. After expanding the study to include the third study group, we decided to coach the trainees as we would a resident in the operating room. This caused the penalties to be increased in the initial simulation trials for the S-X group. If only the times to complete the procedures are compared, there is no difference between the S-2 and S-X groups.

In summary, this study shows that the ability to perform a procedure in an animate model can be improved by practice on synthetic material. The full benefit of such complex simulation has a similar learning curve to that of animate models. Even after statistical improvement in ability is observed, further improvement in performance can be expected in speed, safety, and requirements for instruction.

The implication of these observations is that the learning curve for technical skills may be shifted out of the operating room and into the laboratory. This would be beneficial to the patients, trainees, mentors, and the health care industry. The cost of a synthetic material ($35/gallbladder) is much less than that of the operating room or animate models ($1,000/pig operation). The learning environment in the laboratory is less stressful and the consequences of errors are less significant. Other organ models could be developed to expand the role of synthetic material in the training of surgeons.
Table 1

<table>
<thead>
<tr>
<th></th>
<th>Animal Time (sec)</th>
<th>Animal Errors (n)</th>
<th>Animal Coach (n)</th>
<th>Animal Score (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-0</td>
<td>2939± 659</td>
<td>13</td>
<td>5.3</td>
<td>3513± 1141</td>
</tr>
<tr>
<td>S-2</td>
<td>2621± 1206</td>
<td>2</td>
<td>3.3</td>
<td>2780± 1284</td>
</tr>
<tr>
<td>S-X</td>
<td>1657± 315*</td>
<td>3.8</td>
<td>3.4</td>
<td>1873± 361*</td>
</tr>
</tbody>
</table>

*differs from S-0 $p<0.001$, differs from S-2 $p<0.05$

Table 1: Time, errors, coaching events and score for the animal laparoscopic cholecystectomy in the three student groups. The time and score in the S-X group are significantly better than in the S-0 group ($p<0.001$) and the S-2 group ($p<0.05$). Time and score are reported in mean ± standard deviation.

Table 2

<table>
<thead>
<tr>
<th></th>
<th>Time (sec)</th>
<th>Error (n)</th>
<th>Coach (n)</th>
<th>Score (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sim1</td>
<td>1457± 379</td>
<td>1.4</td>
<td>2.1</td>
<td>1562± 449</td>
</tr>
<tr>
<td>Sim2</td>
<td>1355± 581</td>
<td>1.4</td>
<td>3.3</td>
<td>1397± 601</td>
</tr>
</tbody>
</table>

Table 2: Mean time, errors, coaching events and score for the two simulations in the S-2 group. There is no difference between any of the parameters ($p>0.08$). Time and score are reported in mean ± standard deviation.
Table 3

<table>
<thead>
<tr>
<th></th>
<th>sim1</th>
<th>sim2</th>
<th>Sim3</th>
<th>sim4</th>
<th>sim5</th>
<th>sim6</th>
<th>sim7</th>
<th>sim8</th>
<th>sim9</th>
<th>sim10</th>
<th>sim11</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time (sec)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1860</td>
<td>1792</td>
<td>1433*</td>
<td>1331*</td>
<td>1069*</td>
<td>1148*</td>
<td>1092*</td>
<td>937*</td>
<td>817*</td>
<td>887*</td>
<td>909*</td>
</tr>
<tr>
<td>± 699</td>
<td>± 738</td>
<td>± 511</td>
<td>± 939</td>
<td>± 347</td>
<td>± 559</td>
<td>± 434</td>
<td>± 261</td>
<td>± 211</td>
<td>± 250</td>
<td>± 313</td>
<td></td>
</tr>
<tr>
<td>Score (sec)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2314</td>
<td>2083</td>
<td>1666*</td>
<td>1504*</td>
<td>1173*</td>
<td>1271*</td>
<td>1182*</td>
<td>980*</td>
<td>860*</td>
<td>947*</td>
<td>932*</td>
</tr>
<tr>
<td>± 590</td>
<td>± 688</td>
<td>± 651</td>
<td>± 1134</td>
<td>± 365</td>
<td>± 613</td>
<td>± 497</td>
<td>± 267</td>
<td>± 223</td>
<td>± 273</td>
<td>± 334</td>
<td></td>
</tr>
<tr>
<td>Error n</td>
<td>2.8</td>
<td>1.8</td>
<td>.6</td>
<td>1</td>
<td>.33*</td>
<td>.5*</td>
<td>.7*</td>
<td>.0*</td>
<td>.33*</td>
<td>.25*</td>
<td>0*</td>
</tr>
<tr>
<td>Coach n</td>
<td>12</td>
<td>8**</td>
<td>5.7**</td>
<td>5.2**</td>
<td>3.2**</td>
<td>2.7**</td>
<td>2.3**</td>
<td>1.4**</td>
<td>1.1**</td>
<td>1.8**</td>
<td>1**</td>
</tr>
</tbody>
</table>

* differs significantly from sim1 (p<0.05)

** differs significantly from sim1 (p<0.01)

Table 3: Time, score, error and coaching events in the S-X group for all the simulations. A statistical difference compared to baseline in time and score was found at the third simulation (p<0.05), for errors was at the fifth simulation (p<0.05) and for coaching was at the second simulation (p<0.01). The learning curve plateau, as defined by less than 10% variation of the time and score, was reached at the ninth simulation. Time and score are reported as mean ± standard deviation.

Figure 1: Synthetic laparoscopic cholecystectomy model.

Figure 2: Scores for each participant in the S-X group. A statistical difference in the score compared to baseline was found at the third simulation (p<0.05). The learning curve plateau, as defined by less than 10% variation of the time and score, was reached at the ninth simulation.

Figure 3: Mean time and score for each simulation in the S-X group. Error bars signify the Standard Error of the mean. A significant improvement occurred with the third simulation (p<0.05).
Figure 4: Mean errors and coaching events for each simulation in the S-X group. A significant improvement in coaching occurred with the second simulation ($p<0.01$). Errors improved significantly at the fifth simulation ($p<0.05$).

**Figure 1**

Simulab Part #'s LC 10, LC 20, ST 10
Figure 2: Scores for each participant in the S-X group. A statistical difference in the score compared to baseline was found at the third simulation (p<0.05). The learning curve plateau, as defined by less than 10% variation of the time and score, was reached at the ninth simulation.
Figure 3: Mean time and score for each simulation in the S-X group. Error bars signify the Standard Error of the mean. A significant improvement occurred with the third simulation (p<0.05).
Figure 4: Mean errors and coaching events for each simulation in the S-X group. A significant improvement in coaching occurred with the second simulation (p<0.01). Errors improved significantly at the fifth simulation (p<0.05).