Learning retention of thoracic pedicle screw placement using a high-resolution augmented reality simulator with haptic feedback

Authors

Cristian J. Luciano, Ph.D.,1,3 P. Pat Banerjee, Ph.D.,1-3
Departments of: Mechanical and Industrial Engineering, Bioengineering, and Computer Science,
College of Engineering, University of Illinois at Chicago;
Brad Bellotte, M.D.5
Department of Neurosurgery, Allegheny General Hospital, Pittsburgh, Pennsylvania;
G. Michael Lemole Jr., M.D.,7
Division of Neurosurgery, University of Arizona, Tucson, Arizona
Michael Oh, M.D.,5
Department of Neurosurgery, Allegheny General Hospital, Pittsburgh, Pennsylvania;
Fady T. Charbel, M.D.,4
Department of Neurosurgery, University of Illinois at Chicago
AND
Ben Roitberg, M.D.,6
Division of Neurosurgery, University of Chicago

Corresponding Author

Prof. P. Pat Banerjee, Ph.D

Mailing Address:
Departments of Mechanical and Industrial Engineering,
University of Illinois at Chicago
2039 ERF, M/C 251
842 W. Taylor, Chicago, IL 60607

Phone 312 996 5599
Fax: 312 413 0447
Email: banerjee@uic.edu

1 The research was supported in part by NIH NIBIB grant 1R21EB007650-01A1
**Background.** In this study we evaluated the use of a part-task simulator with 3D and haptic feedback as a training tool for a common neurosurgical procedure – placement of thoracic pedicle screws.

**Objective.** The purpose of this study was to evaluate the learning retention of thoracic pedicle screw placement on a high-performance augmented reality and haptic technology workstation.

**Methods.** Fifty one fellows and residents performed thoracic pedicle screw placement on the simulator. The virtual screws were drilled into a virtual patient’s thoracic spine derived from a computed tomography data set of a real patient.

**Results:** With a 12.5% failure rate, a two-proportion z-test yielded $P=0.08$. For performance accuracy, an aggregate Euclidean distance deviation from entry landmark on the pedicle and a similar deviation from the target landmark in the vertebral body yielded $P=0.04$ from a two sample t-test in which the rejected null hypothesis assumes no improvement in performance accuracy from the practice to the test sessions, and the alternative hypothesis assumes an improvement.

**Conclusions.** The performance accuracy on the simulator was comparable to the accuracy reported in literature on recent retrospective evaluation of such placements. The failure rates indicated a minor drop from practice to test sessions, and also indicated a trend ($P=0.08$) towards learning retention resulting in improvement from practice to test sessions. The performance accuracy showed a 15% mean score improvement and over 50% reduction in standard deviation from practice to test. It showed evidence ($P=0.04$) of performance accuracy improvement from practice to test session.

**KEY WORDS** • haptics • neurosurgical simulation • thoracic pedicle screw • virtual reality

**Background**

At the 2009 annual meeting of the AANS (American Association of Neurological Surgeons), the Young Neurosurgeons Committee continued its annual tradition of organizing a surgical competition using emerging simulators for residents and fellows in the exhibit hall. Thoracic pedicle screw placement was one of the techniques tested, and the results of this testing are reported in this paper. The performance of 51 fellows and residents was evaluated for learning retention of thoracic pedicle screw placement using the head-and-hand-tracked high-resolution and high-performance augmented reality and haptic technology workstation known as ImmersiveTouch® (ImmersiveTouch, Inc.) (Fig. 1).

The ImmersiveTouch augmented virtual reality system was developed at the University of Illinois at Chicago and combines real-time haptic feedback with high-resolution stereoscopic display. An electromagnetic head-tracking system provides dynamic perspective as the user moves his or her head. A half-silvered mirror is used to create an augmented reality environment that integrates the surgeon’s hands, the virtual instrument, and the virtual patient in a common working volume while eliminating image occlusions. The system has been evaluated for applications such as ventriculostomy and VP shunt placement.

The ImmersiveTouch system offers a number of convenient options for pedicle screw placement training. As shown in Fig 2, the user has the option of continuously monitoring the trajectory of the drill by simulated A/P, transverse and lateral fluoroscopic views as a means of image guidance. For
advanced training these views can be removed. The system offers tactile feedback as well as the simulated vibration of the virtual drill. A range of force feedback and drill frequencies are available. The system is sensitive to drill operation, e.g. if the user does not switch the drill on as simulated by its vibration and noise, the tip of the virtual drill can slip when the user first makes contact of the tip with the spine surface.

Insert Fig 2 here

Objective

The objective is to experimentally determine learning retention of thoracic pedicle screw placement. The user first undergoes training through a practice session with access to learning aids such as continuously updating fluoroscopic images. The user is subsequently asked to repeat the same task that was practiced through a test session with conditions more similar to those found in the OR, such as limited access to fluoroscopic images. The comparative evaluation of the practice session followed by the test measures the “learning retention”.

Methods

Selection Criteria

Participants consisted of 51 fellows and residents selected by the Young Neurosurgeons’ Committee at the 2009 AANS annual meeting through an open solicitation. They participated in a competition that featured a thoracic pedicle screw simulation running on the ImmersiveTouch. Learning retention of simulated open thoracic drilling and pedicle screw placement was evaluated. The ImmersiveTouch software utilizes a series of modules to acquire, process, and render the graphic and haptic data, which are then seamlessly integrated on the hardware platform. A virtual 3D volume of a human spine was created using a CT from a patient at the University of Illinois at Chicago Medical Center. The data were pre-segmented and assembled from a CT DICOM data set after the removal of all identifying personal data. The 3D polygonal isosurfaces corresponding to the skin, and underlying spinal column were extracted using the Marching Cubes algorithm.

Experiment

The participants were given approximately 5 minutes to practice on any pedicle of their choice from among the six pedicles: the left and the right T9, T10 and T11, respectively. As shown in Fig 3, a cutaway view is presented to the user with limited view of the spinal segment which is roughly equivalent to the view during an open surgical procedure. In the ImmersiveTouch system, an electromagnetic sensor attached to the stereoscopic goggles tracks head movements to compute the correct viewer’s perspective while the user moves his or her head around the virtual patient pedicles to locate or clearly view the landmarks on it by adjusting the relative position and orientation between the observer and the virtual patient.

Insert Fig 3 here

The practice session included continuously updating real-time computer simulated fluoroscopic images in A/P, laterals and transverse views as shown in Fig 3. This permitted the users to track all the
details of their targeting. Following practice these images were hidden during the test session based on repeating the skills learnt during practice.

The user may then use the cutting tool while moving and rolling the wrist as well as rotating the head to visualize the exact location of the screw and correlate the experience with technique (Fig. 4). This is achieved by means of a 6D sensor located inside the SpaceGrips (LaserAid) to track the position and orientation of the surgeon’s hand to define a cutting plane. For each participant, the final position and orientation of the screw were recorded by the computer.

Based on the position and orientation of the haptic stylus (SensAble Technologies, Inc.), collision between the virtual drill and the extracted 3D isosurfaces is detected, and corresponding force feedback is generated through the servomotors of the haptic device. The virtual drill is perfectly collocated with the haptic stylus, even when the user moves his/her head, because head movements are tracked by an electromagnetic sensor. The drill is started and stopped using the buttons on the SpaceGrip. Once the drill is perceived to reach its final destination, clicking one of the SpaceGrip button instantly places the screw at the final location of the drill. While drilling, the haptic stylus can be moved following a linear trajectory defined according to its orientation. A reactionary force is applied by the haptic device as the user deviates from that linear trajectory. This is intended to model good practice and it also simulates a firm feeling while drilling that is similar to the tactile sensation experienced during surgery. If trajectory is to be changed then the user has to take the drill out of the pedicle first by backtracking on the same linear trajectory before reinserting the drill following a different linear trajectory.

(see Video, Supplemental Video 1, which provides a visual highlight of the pedicle screw insertion simulation, 1 minute, 2.5MB)

Measurement Technique

A maximum of six pedicle screw drilling was permitted during both the practice and test sessions. The drilling action was simulated based on a typical electrical drill with simulated vibration felt during burr drill rotation and haptically simulated tactile resistance felt during drilling. The drill parameters are adjustable and have been set based on feedback from the experienced surgeons who are co-authoring this article to reflect actual sensation while drilling thoracic pedicle screws. The participants were informed at the beginning that they will be tested on the same problem following the practice session. The participants were allowed to ask questions during the practice session. Following the practice session, the data was collected through a test session in which the participants were asked to repeat the skills practiced on the same set of pedicles. Performance data are collected for a minimum of four different pedicles up to a maximum of six different pedicles.

Performance Accuracy Measurement: During the practice and the test sessions the participants were given a recommended landmark on each pedicle to start drilling from and another landmark at which to stop the drilling process. The score is based on the Euclidean distance in mm from these targets. If the screw is placed outside of the spinal body, then it carries a penalty of 200 (which is outside of the normal distance range) to the evaluation score to indicate a failure of the procedure. The aggregate score is the sum of all individual scores. The lower the score, the better is the performance.

Failure Rate Measurement: The failure rate is measured by detecting the collision between the burr drill and the virtual spine model (3D isosurface). The point of entry is detected through a collision
with the spine surface at the entry point. Once an entry is detected, it is flagged. This success flag remains on unless the tip of the drill exits the spine surface. The failure rate is the ratio of the number of failures to the total number of attempts.

Results

The failure rate analysis results are reported first. If the participant successfully inserts the screw into the pedicle then it is recorded with a score of 1, whereas if the participant is unsuccessful then the assigned score is 0. We have eliminated outliers, which indicate measurements for which there was a drastic difference in performance from practice to the test sessions. Outliers were determined by first finding the minimum, Quartile 1, median, Quartile 3, and the maximum from the collected data. The formula for the outlier value is Quartile 1-1.5*IQR and Quartile 3+1.5*IQR, where IQR indicates inter quartile range. IQR is the difference in values between Quartile 1 and Quartile 3. In our case, we get a lower outlier of -0.4542 and an upper outlier of 0.5569. This means that if the average score of a resident increases by more than 0.5569 or decreases by more than 0.4542 then that resident would be considered an outlier. The data collected for one resident was rejected because it turned out to be an outlier. Therefore, out of 52 participants who participated in both the practice and test sessions, we used the measurements from 51.

Out of 301 attempts, 51 failures were recorded during the practice sessions, thereby indicating a failure rate of 16.9%. For the test session, 28 out of 224 pedicle screw placement attempts ended up in failure for a failure rate of 12.5%. The reduced failure rate indicates a degree of positive learning retention from practice to the tests.

A two proportion z-test was conducted. The null hypothesis $H_0$ represents no improvement in the failure rate of the residents from the practice to the test sessions. The alternate hypothesis $H_A$ indicates that the failure rate increases from the practice session to the test session. The two-proportion z-test yielded $P=0.08$, which indicates a trend towards rejection of $H_0$. A trend towards learning retention from the practice session to the test session is thus established.

Following the results of failure rate analysis let us now present the results for performance accuracy. The performance accuracy metric in the experiment is the sum of Euclidean distance from starting landmark on the pedicle and the Euclidean distance from the target landmark. The average Euclidean distance deviations for all attempts are computed for each participant. The A/P view in Fig. 3 shows the starting landmark, while the lateral view shows the target landmark for each pedicle. The observed mean $\bar{x}_1$ and standard deviation $\sigma_1$ for the practice session are 27.58 mm and 15.08mm respectively. For the test session the observed mean $\bar{x}_2$ and standard deviation $\sigma_2$ are 23.42mm and 7.44mm respectively. The sample size is 51 after removal of outliers. A two sample t-test is performed for which the null hypothesis $H_0$ assumes no improvement in the performance accuracy from the practice to test samples, while the alternate hypothesis $H_A$ assumes an improvement. By obtaining $P=0.04$ from t-test, the null hypothesis can be rejected. The evidence shows that learning during practice improves the performance accuracy during the test. Furthermore, the mean $\bar{x}_2$ shows a 15.1% improvement over $\bar{x}_1$ and the standard deviation $\sigma_2$ is less than half of $\sigma_1$ which further corroborates the learning retention carried over from practice to the tests.

Conclusion

In this study we evaluated the use of a part-task simulator with 3D and haptic feedback as a training tool for a common neurosurgical procedure – placement of thoracic pedicle screws. Participants reported satisfaction with the realism of our simulation We found an error rate consistent with that
reported in the literature for pedicle screw insertion in clinical practice, and demonstrated some performance improvement when comparing the initial training session with the test session. Since an alpha value of within 0.05 is typically considered statistically significant, the performance accuracy improvement from practice to test is notable and the initial results in learning retention trends presented in the paper warrants further investigation through more follow-up experiments in future.

This work follows our previous experience with simulation of the placement of a ventriculostomy.3,4 We aim to develop a growing set of tasks that are simulated in a virtual reality environment and used for neurosurgical training. We chose to simulate placement of thoracic pedicle screws because their accurate insertion is considered an important training and safety goal. It was evaluated and discussed in multiple studies.9-28 Thoracic pedicle screw insertion appeared to have a substantial risk of breaching the bone even when using intraoperative image or fluoroscopic guidance. The rate of pedicle perforation varied between 4.6% and 47%.12,18 Middle thoracic levels were the most likely to have a perforation12. Even when using computer image guidance, 8.5% cortical violations were found in a four-year study by Youkilis, et al.27 Sagi et al.24 found that image guidance placed 92% of thoracic pedicle screws safely versus 90% of conventional (non-guidance) fluoroscopy. Similar rate of cortical violation was found in other studies, even when using C-Arm 3D fluoroscopic guidance.13 In a recent and more extensive study, Tian and Xu26 report examining the OVID, Springer, and MEDLINE databases consisting of 7,533 pedicle screws of which 6,721 screws were accurately inserted into the pedicles (89.22%). The median placement accuracy for in-vivo CT-based navigation subgroup was 90.76%, whereas with the use of 2D fluoroscopy-based navigation it was 85.48%. Based on a study of placing 150 pedicle screws in the T1-T3 levels using 3D image guidance, Bledsoe et al.9 conclude that 140 (93.3%) out of 150 screws were contained solely in the desired pedicle. Nottmeier et al.22 studied a total of 1084 thoracolumbar pedicle screws placed using either the BrainLAB Vector Vision (BrainLAB, Inc.) or Medtronic StealthStation Treon (Medtronic, Inc.) image guidance systems noted a breach rate was 7.5%.

Training and experience appear to have a role in improving outcomes. Kim et al.14 conducted a cadaveric study using a computer-assisted image guidance system for testing the accuracy of thoracic pedicle screw placement. The overall pedicle cortex violation was 23 of 120 pedicles (19.2%). Nine violations (7.5%) were graded as major and 14 (11.7%) as minor. A marked and progressive learning curve was evident with the perforation rates that decreased from 37.5% in the first cadaver to 4.2% in the last two cadavers. Schizas et al.23 studied the accuracy of upper thoracic screw placement without the use of fluoroscopy or image guidance using a modified Roy-Camille technique. The overall pedicle screw placement accuracy was 88.3% based on the performance of a single surgeon inserting 60 screws in 13 consecutive non-scoliotic spine patients. In this study, the authors found that inserting pedicle screws in the upper thoracic spine based solely on anatomical landmarks was safe with accuracy comparable to that of published studies using image-guided navigation at the thoracic level. These publications support the notions that 1) even with image guidance there is a substantial cortical perforation rate, and 2) training in insertion of thoracic pedicle screws can improve outcome. Therefore, even in the current era of improving intraoperative imaging training is important, and much effort has been expended on improving surgeon performance. Klein et al.15 introduced a CT-based patient-specific simulation software for pedicle screw insertion. However, it does not provide much user feedback for their system, which lacks haptics/graphics collocation. Similarly, the commercially-available BrainLAB Vector Vision and Medtronic StealthStation do not provide an interactive 3D environment with haptics/graphics collocation.

For our simulation we chose multidirectional fluoroscopic guidance under multiple settings. The number of views and fluoroscopy time can be selected by the trainee or a teacher, to simulate varying levels of difficulty and fluoroscopy availability. The accuracy of virtual pedicle screw placement achieved by participants using our simulator is comparable to the accuracy reported in recent retrospective evaluations of such placements. This similarity suggests that our simulator faithfully represents the “real-life” conditions, an important validation point.
The design of our study was limited by the circumstances of data collection. On the positive side, we used a relatively large group of trainees under controlled and uniform circumstances, provided by the Top Gun competition at the AANS meeting. We believe that this data set is valuable for learning about the potential of various simulation applications, getting realism and performance feedback and getting data that helps validate the simulator as a learning tool. On the other hand, the participants had a standard and limited time to learn the simulator, and could not have extensive repetitions. The uniform data was easier to analyze and compare, but we effectively eliminated one of the key advantages of computer-based simulation – the ability to perform unlimited repetitions until the task is learned. Given the brief and uniform training sessions, we were almost surprised to see evidence that learning occurred. Because the participants were primarily residents, some of them with limited or no experience, the error rates during both the practice and the test sessions may seem to be on the higher side. The entry and target landmarks were visible during the practice sessions but hidden during the tests, therefore the participants had to learn not only to identify the entry and target landmarks from the anatomy of the patient’s spine, but also to remember them for the tests. We believe that lack of experience contributed to the high error rates reported by this study. However, the learning of anatomy during the training session was one of the benefits of the training and may have contributed to the trend towards improved performance in the test session. In other words, for a beginner, even the most basic training demonstrating the anatomy and basic starting and target landmarks, can improve performance. Similar to real surgery we also observed that performance on the simulator is sensitive to the drill operation, e.g. if the user does not turn the virtual drill on prior to placing the virtual drill tip on the virtual spine surface, then the drill is likely to slip from the intended starting hole target when it is turned on. Learning virtual drill operation may also have contributed to improved performance on the test compared to the training session.

This study presents a set of data and feedback under particular circumstances. We believe that it is very valuable in order to demonstrate the potential of advances part-task simulator of neurosurgical procedures. To the best of our knowledge this is the most advanced haptics-based open surgical simulator currently available. Here we presented preliminary evidence validating our simulator as a training tool. Additional studies will include more extensive and individualized training of each subject, and expanded metrics of performance. Further studies comparing computer-based simulation with cadaver or other models for training can also be useful.

The opinions and preferred operating conditions by surgeons are extremely diverse. Our simulator is well suited for a variety of simulation settings, including training with or without image guidance under a variety of training conditions. We can also adjust multiple parameters according to the wishes of teachers or trainees. Therefore, many surgical tasks are well suited for practicing and testing using our simulator. Current research is directed at expanding the library of tasks that are simulated using our technology.

Acknowledgement: The authors would like to acknowledge the help of Naga Dharmavaram in organizing the data and in running some statistical tests.

Disclosure
Presented at the 2009 AANS annual meeting, this study was sponsored by the AANS Young Neurosurgeons Committee, which does not claim the superiority of the ImmersiveTouch® system over another system. The ImmersiveTouch technology has been licensed to ImmersiveTouch, Inc., by the University of Illinois at Chicago. Coauthors Dr. Charbel and Dr. Banerjee both own stock in ImmersiveTouch, Inc.

The research was supported in part by NIH NIBIB grant 1R21EB007650-01A1
References


**Figure Legend:**

Fig. 1. The *ImmersiveTouch®* workstation at AANS 2009 used for pedicle screw insertion simulation

Fig. 2. Drill with force feedback for locating screws

Fig 3. Practice session with continuously updated fluoroscopic images. The green landmarks on each pedicle visible in the A/P view indicates a recommended starting point, while the red landmarks visible in the lateral view indicates a recommended end point for the virtual drill

Fig. 4. Three-dimensional views of the pedicle screw for visual analysis of performance

**Supplemental Content:**

Supplemental Video 1. Video which provides a visual highlight of the pedicle screw insertion simulation

Author: Pat Banerjee
Videographer: Cristian Luciano
Length: 1 minute
Size: 2.5MB