Spring 2019

Orthopaedic surgical skills: examining how we train and measure performance in wire navigation tasks

Steven A. Long
University of Iowa

Follow this and additional works at: https://ir.uiowa.edu/etd

ORTHOPAEDIC SURGICAL SKILLS:
EXAMINING HOW WE TRAIN AND MEASURE
PERFORMANCE IN WIRE NAVIGATION TASKS

by

Steven A. Long

A thesis submitted in partial fulfillment
of the requirements for the Doctor of Philosophy
degree in Biomedical Engineering in the
Graduate College of
The University of Iowa

May 2019

Thesis Supervisors: Professor Donald D. Anderson
Professor Geb W. Thomas

ACKNOWLEDGEMENTS

There are many people that I would like to acknowledge for their help and support that have made this work possible. First and foremost, I would like to thank my advisors Don Anderson and Geb Thomas. We have come a long way these past years and I couldn’t have done this without your amazing guidance and support along the way. I would also like Dr. Matt Karam and Dr. Larry Marsh for their amazing clinical expertise that they bring to the team. Having good clinical collaborations is crucial to work like this and their advice and expert opinions are second to none. I would also like to thank the members of my committee Dr. Nicole Grosland and Dr. David Wilder for agreeing to participate in this process and help make the work stronger.

Much of this work was done with the support of members of the Biomechanics Lab and the GROK lab. Special thanks to Madison Chrisman, Anna Rodriguez, Marcus Tatum, Oliver Stroh, and Dominik Mattioli for helping on different trips to residency programs and conferences to help collect this data. Special thanks to Kevin Dibbern for accompanying me on the 18-hour drive to Orlando, Fl to collect crucial data for this document. And special thanks to Leah Taylor for collecting a large amount of data from the operating room that led the way for a lot of this work.

Lastly, I must thank my family, including my parents, sister, brother-in-law, new nephew Jack, future mother-in-law, and especially my extremely supportive and caring fiancée Kira for all their support throughout this entire process. I couldn’t have done it without you by my side.

ABSTRACT

Until recently, the model for training new orthopaedic surgeons was referred to as "see one, do one, teach one". Resident surgeons acquired their surgical skills by observing attending surgeons in the operating room and then attempted to replicate what they had observed on new patients, under the supervision of more experienced surgeons. Learning in the operating is an unideal environment to learn because it adds more time to surgical procedures and puts patients at an increased risk of having surgical errors occur during the procedure. Programs are slowly beginning to switch to a model that involves simulation-based training outside of the operating room. Wire navigation is one key skill in orthopaedics that has traditionally been difficult for programs to train on in a simulated environment. Our group has developed a radiation free wire navigation simulator to help train residents on this key skill.

For simulation training to be fully adopted by the orthopaedic community, strong evidence that it is beneficial to a surgeon's performance must first be established. The aim of this work is to examine how simulation training with the wire navigation simulator can be used to improve a resident's wire navigation performance. The work also examines the metrics used to evaluate a resident's performance in a simulated environment and in the operating room to understand which metrics best capture wire navigation performance.

In the first study presented, simulation training is used to improve first year resident wire navigation performance in a mock operating room. The results of this study show that depending on how the training was implemented, residents were able to significantly reduce their tip-apex distance in comparison with a group that had received a simple didactic training. The study also showed that performance on the simulator was correlated with performance in this operating

room. This study helps establish the transfer validity of the simulator, a key component in validating a simulation model.

The second study presents a model for using the simulator as a platform on which a variety of wire navigation procedures could be developed. In this study, the simulator platform, originally intended for hip wire navigation, was extended and modified to train residents in placing a wire across the iliosacral joint. A pilot study was performed with six residents from the University of Iowa to show that this platform could be used for training the other applications and that it was accepted by the residents.

The third study examined wire navigation performance in the operating room. In this study, a new metric of performance was developed that measures decision making errors made during a wire navigation procedure. This new metric was combined with the other metrics of wire navigation performance (tip-apex distance) into a composite score. The composite score was found to have a strong correlation (R squared = 0.79) with surgical experience.

In the final study, the wire navigation simulator was taken to a national fracture course to collect data on a large sample of resident performance. Three groups were created in this study, a baseline group, a group that received training on the simulator, and a third group that observed the simulator training. The results of this study showed that the training could improve the overall score of the residents compared to the baseline group. The overall distribution from resident performance between groups also shows that a large portion of residents that did not receive training came in below what might be considered as competent performance. Further studies will evaluate how this training impacts performance in the operating room.

PUBLIC ABSTRACT

Until recently, the model for training new orthopaedic surgeons was referred to as “see one, do one, teach one”. Resident surgeons acquired their surgical skills by observing attending surgeons in the operating room and then attempted to replicate what they had observed on new patients, under the supervision of more experienced surgeons. Learning in the operating is an unideal environment to learn because it adds more time to surgical procedures and puts patients at an increased risk of having surgical errors occur during the procedure. Programs are slowly beginning to switch to a model that involves simulation-based training outside of the operating room. Wire navigation is one key skill in orthopaedics that has traditionally been difficult for programs to train on in a simulated environment. Our group has developed a radiation free wire navigation simulator to help train residents on this key skill.

For simulation training to be fully adopted by the orthopaedic community, strong evidence that it is beneficial to a surgeon’s performance must first be established. The aim of this work is to examine how simulation training with the wire navigation simulator can be used to improve a resident’s wire navigation performance. The work also examines the metrics used to evaluate a resident’s performance in a simulated environment and in the operating room to understand which metrics best capture wire navigation performance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: BACKGROUND

Orthopaedic Residency Education

Surgical skills training has been undergoing a gradual culture shift for almost two decades now. The landmark study, "To Err is Human: Building a Safer Healthcare System" suggested that up to 98,000 deaths occur each year due to medical error and that "more than two-thirds (70 percent) of the adverse events found in this study were thought to be preventable, with the most common types of preventable errors being technical errors (44 percent)" [1]. Following this, working hour restrictions were implemented in both the United States and Europe, aiming to reduce the amount of errors made due to mental and physical fatigue [2, 3]. Those working hour restrictions however have potentially had the unintended consequence of limiting the number of opportunities for surgeons to gain valuable learning experiences during residency. A study by Bell et al. looked at the amount of experience surgeons were getting while in residency compared to the expected amount of surgical experience. The most notable statistic from that study was that of 121 common and required experiences expected before graduation, most residents had experienced none of them [4]. Similar studies have also raised questions about residents' preparedness for independent practice upon graduating from residency [5, 6].

In orthopaedics, there has been a large push in recent years to combat these concerns and improve the training model for residents. A mandate by the American Board of Orthopaedic Surgery (ABOS) in 2013 required laboratory-based training of basic surgical skills for first year residents [7, 8]. Early reports following this mandate examined a variety of simulation and training modalities to provide laboratory-based skills training [9-12]. Now, orthopaedic residency programs around the globe are making a push towards competency based medical education (CBME) [13]. CBME differs from traditional residency education models in that it is focused on demonstrating skill acquisition before graduation, as opposed to previous models that focused on a time-based graduation system. Although there are logistical issues to consider, competency-based education has several benefits, such as clearly defined learning objectives for trainees, a more confident and competent workforce of graduating surgeons, and the potential for better patient outcomes [14]. Several national organizations have begun slowly introducing this type of system. A pilot study was conducted in Canada that rotated residents based on their levels of competency following simulation-based training [15, 16]. This pilot study showed that residents who had participated in the competency-based training performed at a superior level in comparison to residents who were trained in a traditional manner. The Royal College of Physicians and Surgeons of Canada (RCPSC) has since announced that all orthopaedic surgery residency programs must be following a competency-based curriculum by the year 2022 [17]. In the UK similar efforts are being undertaken. A program called the Orthopaedic Competency Assessment Project (OCAP) has been established to provide a curriculum and assessment system for all surgical training programs [18]. This program includes electronic logbooks of surgeries, competency-based modules, and formal assessments. Simulation has an important complementary role to play both in training and performance assessment. The expectation is that at the end of training, an orthopaedic resident will be able to perform over 100 trauma procedures independently and will have been formally assessed for competence on six specific trauma procedures [19]. In the US, formal integration of a competency-based system has been somewhat slower. The Accreditation Council for Graduate Medical Education (ACGME) has introduced surgical milestones that orthopaedic residents are required to meet and document throughout their residency [20]. Residents are graded on a scale of 1 (ready to begin residency) to 5 (exceeds expectations for graduation) on these milestones and are expected to reach at least a level 4 by graduation. The milestones cover 16 clinical areas, 6 of which are related to orthopaedic trauma. Although progress is being made by many orthopaedic institutions toward systems of competency-based education, there is still no clear consensus on what system works best. For orthopaedic institutions, especially those in the US, to fully transition into competency-based education, there needs to be more clear evidence of which training methods and which assessment tools work best at both improving and evaluating orthopaedic surgical skills. Simulation will play an important role in these new competency-based systems.

Assessing Orthopaedic Resident Competency

Ultimately, a successful simulation program would clearly demonstrate that training on a simulated task improves objective, reliable, and valid metrics of performance on that task in the operating room (OR). This statement however relies on two assumptions. First, that objective, reliable, and valid metrics of performance in the OR exist, and second, that the training of the task is well understood and implemented correctly. That said, there are few published studies in orthopaedics that demonstrate the veracity of either of these assumptions. Most studies focus on tying performance assessments back to performance on a simulator or some other metric of surgical competency, rather than performance in the OR [21-23].

There are two main categories to examine when evaluating orthopaedic resident competency; orthopaedic knowledge and technical skills. In North America, the Orthopaedic In-Training Examination (OITE) is a commonly used exam to assess resident knowledge that has been shown to correlate with performance on Part 1 of the ABOS boards examination [24, 25]. That said, multiple studies have also shown that scores on the OITE do not necessarily correlate with surgical skills or clinical performance [26, 27]. This intuitively makes sense, given that technical skills and clinical knowledge are independent of one another. However, it also highlights the important need for valid metrics of assessing technical surgical skills.

This leads one to ask: what metrics are currently being utilized to assess technical surgical skills? One form of surgical assessments involves using global ratings and checklists. In arthroscopy, the Arthroscopic Surgical Skill Evaluation Tool (ASSET) looks at specific behaviors and tasks that should be performed during the procedure [28, 29]. Another commonly cited metric for assessing surgical skills during open surgical procedures that has also been adapted for orthopaedics is the Objective Structured Assessment of Technical Skills (OSATS) score [30]. These metrics often rely on expert surgeons observing a novice perform a task and grading the performance based on categories of behavior implicit to the surgery such as instrument handling, respect for tissues, knowledge of instruments, and flow of operation [31]. However, these categories do not reflect metrics that have an impact on the outcome of the quality of the surgical result. Prior work has argued that given the assessments are made based on what the surgeon can observe the novice doing, and not necessarily the inner workings of how a fracture is being put back together or a wire being driven in bone, that the OSATS score might not be effectively assessing the quality of the surgical result (**Figure 1**) [9]. A better strategy may instead be to evaluate performance based on measures from the radiographic images acquired during surgery. These fluoroscopic images contain rich information about the iterative process a surgeon goes through while trying to place an implant in bone, put a fractured ankle back together, or drive a guide wire through a bony corridor.

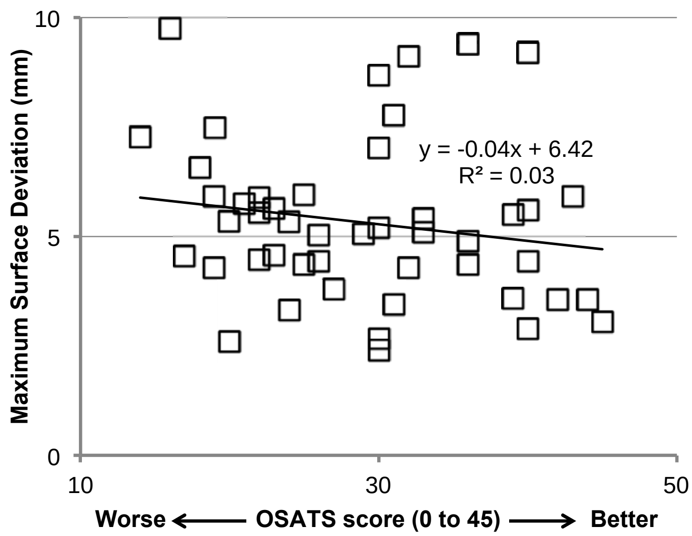


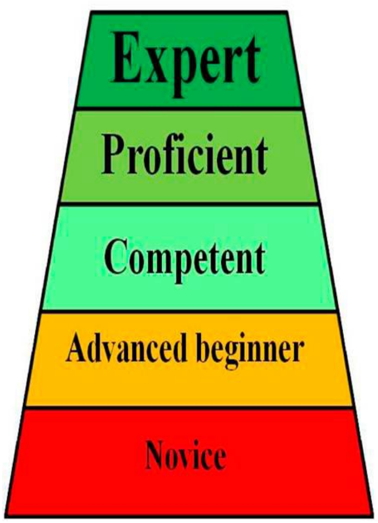
Figure 1: This graph shows the poor correlation between a resident's given OSATS score the quality of a surgical reduction on a simulated fracture repair. Image taken from [9].

Recently, a study published by the ABOS and Council of Orthopaedic Residency Directors (CORD) evaluated two online assessment platforms, the O-Score and the P-Score. The O-Score, or the Ottawa Surgical Competency Operative Room Evaluation, is a formative assessment tool intended to broadly assess any surgical procedure [32]. Like the OSATS score, the O-Score has various categories of evaluation upon which an attending surgeon must grade the resident. However, the O-Score grades based off how much input the attending surgeon had to provide during the surgery, going from a level 1 of instances where the attending surgeon had to do a portion of the procedure to a level 5 where the attending surgeon did not need to be present and felt the resident was independent and practice-ready. The P-Score is a single-question evaluation that allows the attending surgeon to give an overall grade of surgical performance and practice readiness. A score of 1 indicates a novice performance that required maximum assistance from an attending surgeon and a score of 5 indicates proficiency and the performance of an advanced surgeon. In the study comparing these two evaluation methods, 16 different residency programs participated over a six-month period to assess the feasibility of the two metrics. The conclusions from this study were that both scores were able to distinguish between entry, intermediate, and advanced level surgeons [33]. However, this study did not examine the level of agreement between attending assessments and given that evaluations were not blinded between the reviewer and resident, one has to wonder if bias played a role in the grading of surgical performance with an O-Score or P-Score. Further, the O-Score may depend on the style of the attending surgeon and some may feel more comfortable having novice surgeons lead than others. Although progress continues to be made on a variety of assessment methodologies, there is still a great deal of work to be done in properly evaluating and measuring competent orthopaedic surgical performance across a variety of surgical skills.

Skills-Training Approaches

The particular style in which training should be implemented is also a topic of discussion. Novice learners function on a very different set of principles than do experts. To properly design training paradigms that will produce improved surgical performance, one must understand the different models of learning and practice that have been developed over the years. A well-known model is Rasmussen's concept of skill-based, rule-based, and knowledge-based performance. Rasmussen introduces the notion that expert performance is a smooth and free flowing operation based on years of experience that have developed into skills. However, novice performance works off of rules and, in the most basic of settings, knowledge. Rules are developed over time as learners experience new environments and situations. When a learner encounters a new situation that they do not have a rule for, they instead have to revert back to first-principles and rely on their knowledge [34]. In an analogous manner, Dreyfus and Dreyfus created a hierarchy of characteristics which define different levels of performance (**Figure 2**). In their model, they stipulate that novices have little to no experience and therefore need rules to function while experts are instead a source of knowledge and experience and work primarily off of intuition [35].

In going from a state of knowledge-based or rule-based performance to skill-based performance, there are several phases that learners go through. Fitts and Posner model these in three phases, the early/cognitive phase, the associative phase, and the autonomous phase [36]. In the cognitive phase, the learner is trying to understand what the task is and what it demands, in essence they are relying on knowledge to learn and acquire the skills. In the associative phase, the learner is trying to piece together chunks of knowledge to build their rules. And lastly in the autonomous phase, similar to the Dreyfus model, the learner is working off intuition.



STAGE	PERFORMANCE CHARACTERISTICS
Expert	• Source of knowledge and information for others • Continually looks for better methods • Work primarily from intuition • Being forced to follow rules degrades performance
Proficient	• Seeks to understand larger context • Frustrated by oversimplification • Can self-correct performance • Can learn from experience of others
Competent	• Can troubleshoot problems on his/her own • Seeks out expert user advice • Develops conceptual models
Advanced beginner	• Starts trying tasks on his/her own • Has difficulty troubleshooting • Begins to formulate principles, but without holistic understanding
Novice	• Has little or no previous experience • Doesn't know how to respond to mistakes • Needs rules to function

Figure 2: The characteristics of performance that learners exhibit as they gain proficiency is laid out by Dreyfus and Dreyfus. Image taken from [37].

In understanding the phases of performance learners progress through, we can better design the type of training provided to them to help them in each phase of learning. For instance, in the early cognitive phase, online modules, videos, or textbook chapters may provide an adequate platform for helping learners develop their knowledge of the task. In an alternative methodology to "see one, do one, teach one", Sawyer et al. offer that a better model may be "learn, see, practice, prove, do, maintain", indicating that learning about the procedure and its indications, contraindication, and motor actions is a critical first step [38]. They further break down the progression of skill development into two main phases, a cognitive phase involving learning and seeing others do the procedure, and a psychomotor phase, involving practicing and assessment of the procedure.

The most well-known form of practice is termed deliberate practice, and it is a common form of training that has crossed many disciplines. In this methodology, trainees learn through repetition of a task while receiving direct and clear feedback that informs them of ways to improve their performance [39, 40]. Proficiency-based training is another teaching style which has recently gained momentum. In this style of training, residents must “reach at least the proficiency level (on two consecutive trials) before progressing to *in vivo* practice” [37]. The argument being that when residents move on to the real-world environment it will only be because they have reached a certain level of surgical skill proficiency rather than the more traditional notion of moving on because of a time-based curriculum. While similar to deliberate practice in that trainees receive feedback as they progress, proficiency training differs in that the overall task is broken down into its sub-components, and learners must become proficient on each component before moving on to the next. This type of training has been shown to be an effective tool in teaching trainees a variety of surgical skills [6].

Simulation-Based Skills Training in Orthopaedics

The ultimate goal of simulation-based training is to demonstrate that the training improves performance in the operating room. In general surgery, a training program known as the Fundamentals of Laparoscopic Surgery (FLS) has been shown by many studies to improve performance in the operating room [41-44]. In orthopaedics, strong evidence linking simulator training to improved performance in the operating room is still lacking. Arthroscopic surgery is one area of orthopaedic training that many researchers have focused upon. Although different from laparoscopic surgery in many ways, in some ways they involve learning similar techniques. Both surgical techniques involve using a camera at the end of an instrument to view the surgical field. Both techniques involve using a second instrument to make surgical corrections such as cutting, drilling, or tying knots. Perhaps this is one reason why arthroscopic surgery has been such an active area of simulation research.

Regardless, there are many lessons we can learn from the studies that have been published on arthroscopic simulation training. One of the most prominent arthroscopic simulators is the FAST system, or the Fundamentals of Arthroscopic Surgery Training (**Figure 3**). The FAST system is an arthroscopic training station that can be modified to train residents on a variety of arthroscopic techniques [45]. A curriculum developed by the ABOS, American Academy of Orthopaedic Surgeons (AAOS), and the Arthroscopy Association of North America (AANA) accompanies the simulator to guide residents through a series of basic and advanced training modules. A study by Pedowitz et al. examined the feasibility of the FAST station and a knot strength testing station at assessing and benchmarking suture knot strength of expert and senior resident surgeons [46]. The study found that faculty surgeons tied knots that failed at a rate of 21% when tested on a biomechanical testing platform. However, a very important feature of the study was that residents who had trained using a proficiency training methodology had knots fail at a rate of 11%. This study not only helped to set benchmarks of expert performance, but it also demonstrated the benefits of proficiency-based training.

Figure 3: The FAST system is shown here. Different boards can be placed inside the dome to train of a variety of arthroscopic skills. A clear dome or solid white dome can be used depending on the stage of the learner. Image taken from [47].

More recently, Schneider et al demonstrated construct validity (i.e., differences in performance on the simulation correlated positively with experience) on the FAST knot tying exercise [48]. This study had 73 participants split into three distinct groups based on experience; medical students, surgical residents, and attending physicians. The study showed that there was a significant difference in the number of successful knots tied on the FAST station between medical students and residents, between attending surgeons and medical students, but not between residents and attending surgeons. Establishing construct validity is an important part of validating a simulator platform because if it is going to be used to test for surgical competency, it must be able to distinguish between levels of novice and expert performance. It may be that in this case, the lack of difference between resident and attending surgeons lies in the binary metric of "failed" versus "did not fail" when grading the strengths of the knot. When possible, it is preferable to quantitatively evaluate the quality of an outcome to better distinguish performance.

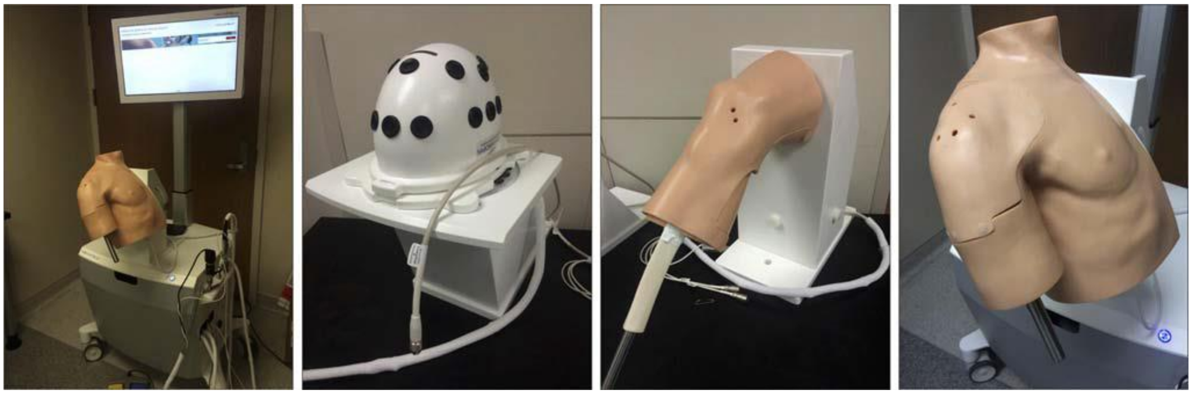


Figure 4: The FAST, Knee, and Shoulder arthroscopy stations of the VirtaMed simulator shown here. Image taken from [49].

Tofte et al. recently explored the learning curve associated with other modules of the FAST station [49]. This study examined residents and faculty at the University of Iowa on Knee, Shoulder, and FAST modules on a VirtaMed arthroscopy simulator (**Figure 4**). The study showed that there was a significant correlation between year in residency and performance on the FAST, Knee, and Shoulder exercises. However, on closer inspection, it appears that the average score of a first-year resident on the FAST module was around 61 (out of a possible 100) versus an average score of around 69 for a surgeon at the Fellow level. Although there might have been a significant correlation between levels of experience, this does not seem to be a meaningful difference in level of performance. However, there were clearly substantial differences between the first-year residents and Fellow surgeons on the shoulder exercise. This might suggest that although the FAST modules can be used as training tools, other tasks are better at differentiating between levels of performance. On a broader note, this shows that the exercises used to train residents different surgical techniques may not be the same exercises used to test for surgical competency.

The study by Tofte et al. brings another topic to the discussion which is, what is the best format for delivery of simulation training? Does the platform need to be a high-end, ultra-realistic simulation model? The VirtaMed simulator used in the Tofte article is listed as a $125,000 product [49]. In 2011, Karam et al. surveyed residency program directors and found that 62% of respondents would pay between $1000 and $15,000 for a surgical simulator [50]. Although there might be great value in a high-end simulator like the VirtaMed product, it may not be feasible for most programs to acquire. At the other end of the spectrum, some programs use ultra-low-tech setups for their simulated training. A group at the Albert Einstein College of Medicine, led by Martin Levy, have become very creative and use a variety of fruits and vegetables for their skills training exercises. In their paper, *The Grapefruit: An Alternative Arthroscopic Tool Skill Platform*, a grapefruit platform was used to train and assess on a variety of arthroscopic surgical skills (**Figure 5**) [51]. This study, like others, demonstrated construct validity between medical students, novice residents, and senior residents. Apart from the outcomes of this study, it is interesting to think of how such a model could be replicated at other institutions or even on a national scale. Could consistency be maintained between grapefruits used in California versus those in Chicago or New York? How does the ripeness of the grapefruit impact the difficulty of the arthroscopy tasks? These questions illustrate that although using grapefruits and other fruits and vegetables is a very cheap and creative solution, it likely would be difficult to implement and standardize across a national stage.

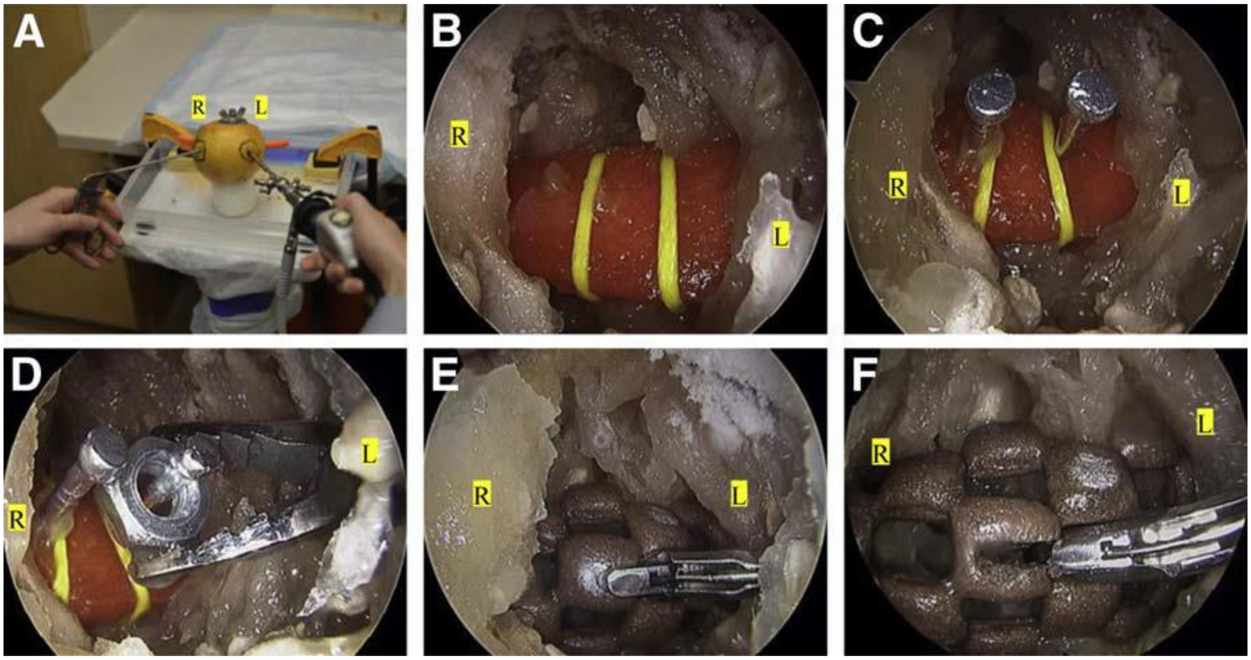


Figure 5: The grapefruit platform used to assess arthroscopic skills is seen in image A. Images B, C, D, E, and F show the setup for different arthroscopic skills being tested. Image taken from [51].

In the Canadian system, which is primed to implement competency-based education, they have implemented a boot camp model to train their residents. In these boot camps, residents train on a variety of cadavers, sawbones, and plaster-based models to practice skills of open-reduction and internal-fixation, drilling, sawing, cutting, wound closure, etc. Studies showed that the residents who participated in these lab-based trainings exhibited significantly higher OSATS scores than residents who received a traditional form of resident training [52-54]. Cadavers clearly offer a large benefit in their level of face validity. However, they can be very costly. One of the largest criticisms of the Canadian boot camp system was its large cost to programs. Before implementing the competency-based curriculum (CBC), the University of Toronto spent $1259 USD on cadavers for training residents. After implementing the CBC, the University of Toronto spent a total of $68,495 USD on cadaver specimens [16]. When including other costs such as the rental of fluoroscopy equipment, sawbones material, and the cost of labor, the total cost of the CBC program at the University of Toronto is $158,050 USD annually [16]. As the programs become more refined, these costs may decrease. However, it is clear that relying so heavily on cadaver models, although very beneficial for training, presents a tremendous financial burden.

Benchtop training models have also been used to teach basic orthopaedic skills [21, 22, 55, 56]. Lopez et al. introduced a do-it-yourself benchtop model called the Fundamentals of Orthopaedic Surgery that was shown to have good construct validity on a variety of tasks including fracture reduction, three-dimensional drill accuracy, depth-of-plunge, and suture speed (**Figure 6**)[55]. Burns et al. evaluated fracture fixation skills using benchtop models to assess depth of plunge, drilling direction, and placing a side plate on a sawbones model (**Figure 7**) [56]. Coughlin et al. used similar low tech hardware to develop a benchtop model for training and evaluating basic arthroscopic skills (**Figure 8**) [21]. While the benefit of these training models is clearly that they are inexpensive, low-tech, easy to use platforms, there are also several issues with them. First, they all have low elements of face validity. Although this is not critical to the success of a simulation, it can often be important in getting successful engagement/participation from residents. Second and more importantly, these models have yet to demonstrate that training with them can lead to improved performance on other applications or in the operating room. Although models like the FLS had success in providing abstract training that did translate into the OR, some of these models seem too basic or not generalizable enough to extend into the OR. Clearly, more work remains to be done.

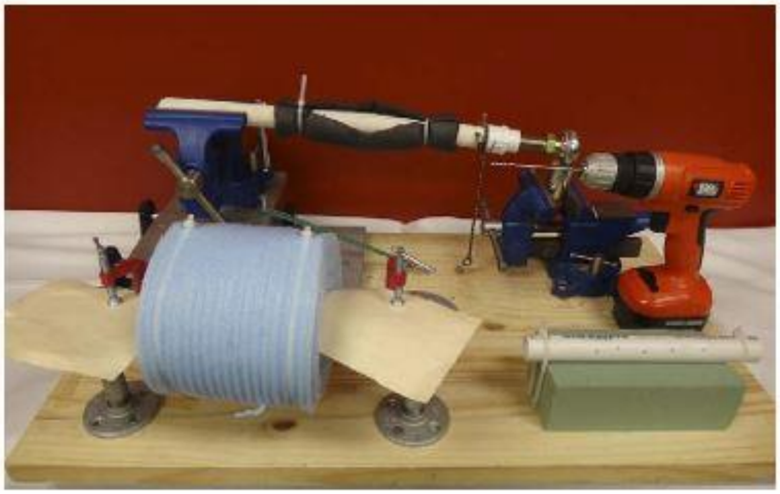


Figure 6: The Fundamentals of Orthopaedics Surgery board is shown here. The board is put together with common items found at hardware stores. Image taken from [55].

Figure 7: The benchtop stations shown from the Burns study. Stations were setup to have residents practice depth of plunge accuracy and drilling direction. A separate module (not shown here) involved placing a plate on a Sawbones mechanical model. Image taken from [56].

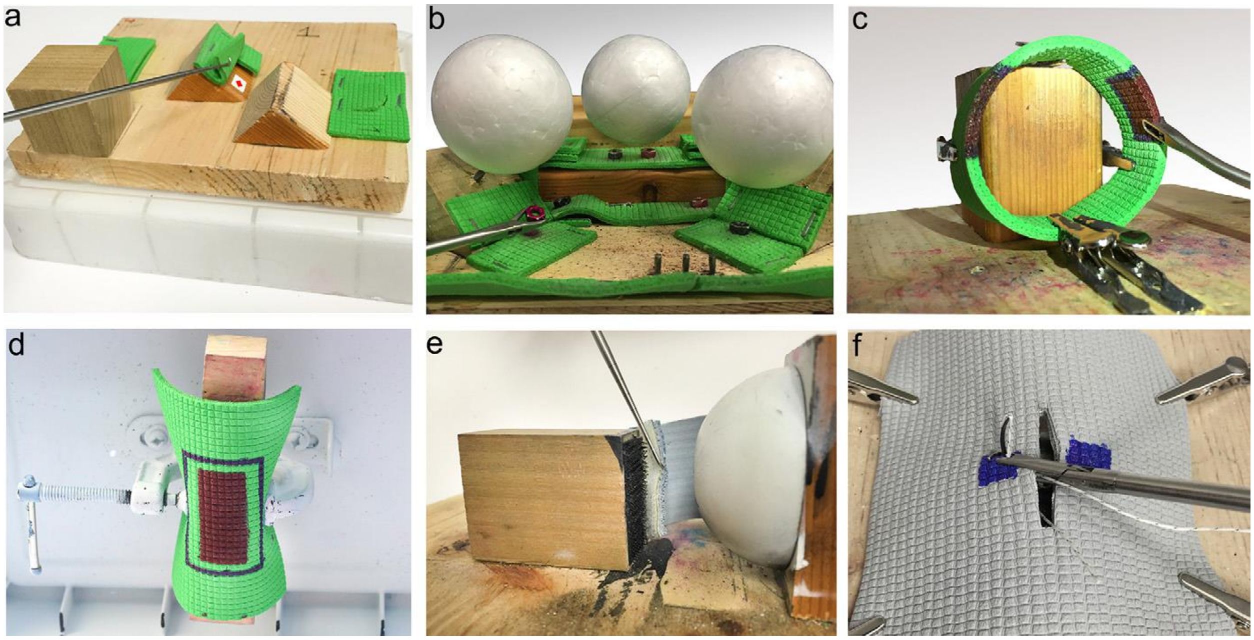
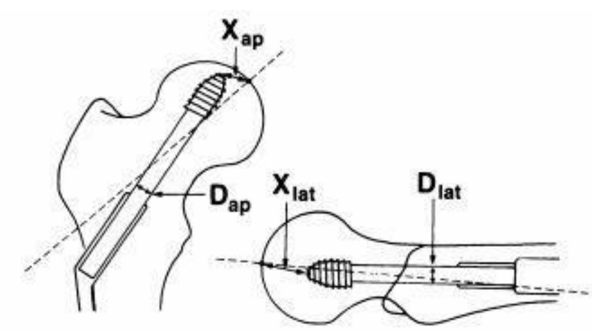


Figure 8: Benchtop stations shown from Coughlin et al. evaluating the platform for training and assessment of basic arthroscopic skills. Image taken from [21].

Simulation for Wire Navigation Tasks

One skill that is central to many applications of orthopaedic surgery, yet has traditionally been difficult to simulate, is fluoroscopic wire navigation. This skill involves interpreting intra-operative fluoroscopy images to understand where the trajectory of a wire is headed into bone. It is used often in the course of repairing certain fractures, and the guide wires placed during wire navigation can either be the final piece of hardware used for fixation or can be used to place a cannulated screw as the final fixation device. The challenge in training this skill is primarily the difficulty of obtaining or having a C-arm fluoroscopy unit readily available. A survey conducted by Karam et al. showed that only 22% of skills labs reported having a C-arm [50]. Additionally, radiation exposure is another barrier to limitless use. Practicing wire navigation skills should take place in an environment that does not continually expose residents to radiation as they home in their skills.

Several studies have examined using simulation to train wire navigation. Most of these studies have focused on wire navigation for the treatment of intertrochanteric hip fractures. The ability to surgically navigate a wire is a critical first step in the operative management of hip fractures and therefore provides an excellent application for training orthopedic residents. Hip fractures are common injuries, and Post-graduate year 2 (PGY2) and year 3 (PGY3) residents are often involved and indeed participate in the wire navigation aspect of these procedures as part of their surgical training. There are roughly 300,000 hip fracture cases each year [57], and that number is soon expected to increase to 500,000 cases per year [58]. Several metrics of surgical performance that have been shown to affect clinical outcome can also be measured in the operating room (OR). The tip apex distance (TAD – Figure 9) is one of the most cited metrics of performance in hip wire navigation. It is a measure of the distance between the tip of the implant and the apex of the femoral head [59]. It is widely cited due to the strong relationship between the TAD value and the likelihood of implant failure as shown by Baumgartner et al [60]. Taylor et al. demonstrated that measurements of a surgeon's wire navigation time and their TAD correlated with greater surgical experience in the OR [61].



$$TAD = \left(X_{ap} \times \frac{D_{true}}{D_{ap}}\right) + \left(X_{lat} \times \frac{D_{true}}{D_{lat}}\right)$$

Figure 9: This image describes the calculation of the TAD. It is a sum of the distances in both the anterior-posterior and lateral views.

Of the studies that have examined hip wire navigation, most have utilized some form of a haptic based VR system [62-68]. Both Akhtar et al. and Christian et al. demonstrated construct validity on the haptic VR system developed by Swemac, a Swedish orthopaedic implant manufacturer (**Figure 10**) [63, 64]. Like other studies examining construct validity, these studies focused on medical students, residents, and attending surgeons. As other studies have shown, the simulator was able to distinguish between medical students and residents, but not between residents and attending surgeons. Christian et al. also examined how scores on the simulator varied based on the amount of time they had been practicing surgery and found a strong correlation (R square of 0.61) between surgical experience and the wire placement in bone.

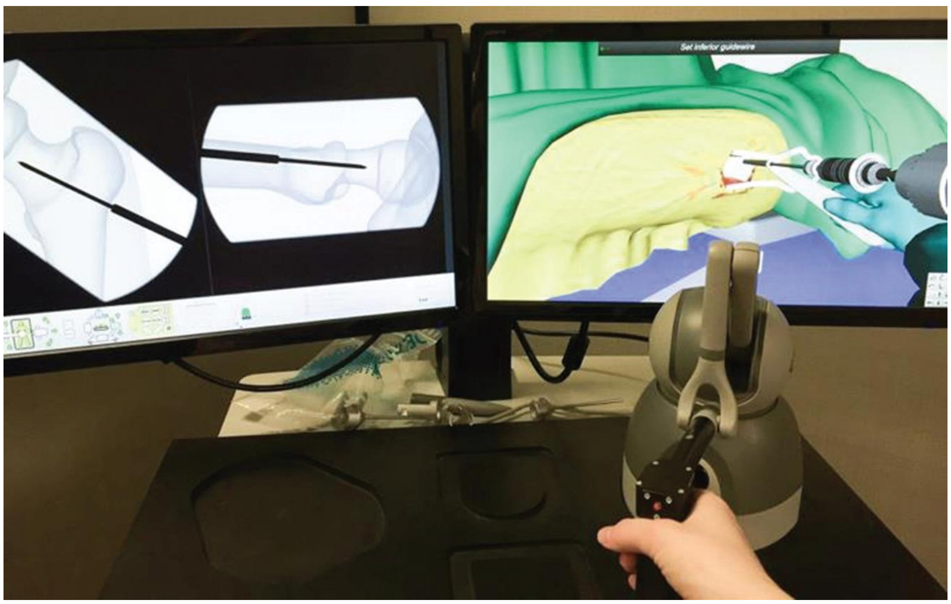


Figure 10: The haptic based VR system developed by Swemac is shown here. The system uses a haptic wand to track a virtual guide wire as it is drilled into a virtual bone. Images showing the surgery and fluoroscopic images are displayed on a screen above the haptic system. Image taken from [64].

One study published in 2018 did not use a haptic system and instead used a vision based system to track the guide wire and generate artificial fluoroscopic images [69]. Shown below in **Figure 11**, the system, titled FluoroSim, uses a set of stereo positioned cameras to track two large spheres permanently affixed to a wire to calculate the wires position relative to a fixed bone and generate artificial fluoroscopic images showing the wire position in bone. In their study, Sugand et al. recruited 26 surgeons that were split into 3 performance groups, novices, intermediates, and experts. They found a statistical difference between the novice group and the other experienced groups, but not between the intermediate and expert groups. A criticism of this study would be that because of the spheres that are placed on the guide wire, the surgeons must use a fixed drill to drive the wire in, rather than a pin collet that would normally be used in the OR and allows the resident to slide the drill up the axis of the wire as they are making adjustments. This is an important feature to make note of because using the fixed drill, as shown in **Figure 11**, forces the drill to be very far away from the bone that the surgeon is drilling into. Wire navigation can be best controlled when a surgeon can keep the drill up close to the bone, thus limiting the moment arm between the hand of the surgeon and the tip of the wire.

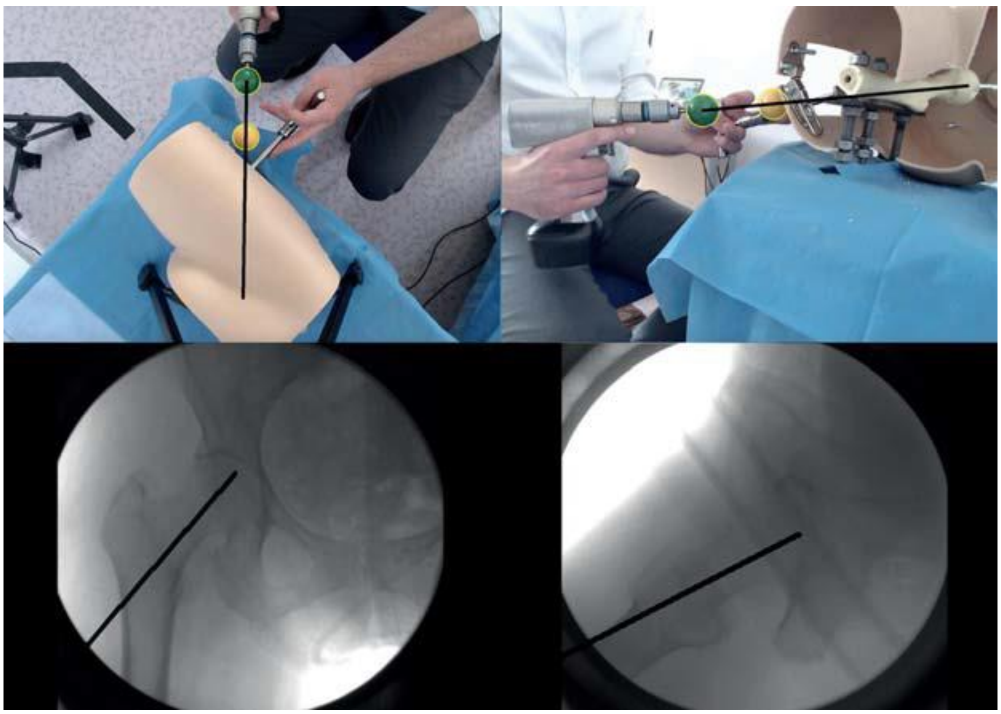


Figure 11: The vision based FluoroSim System is shown here. The system works by tracking two large beads that are permanently affixed to the wire to project the wire position into bone. Artificial fluoroscopic images are then created and shown to the user. Images taken from [69].

In previous work, our team at the University of Iowa developed a hybrid reality wire navigation simulator (Figure 12) [70]. This simulator takes advantage of an optical system to track a guide wire relative to a fixed bone model, not unlike the simulator developed by Sugand et al. Unlike the work of Sugand et al, our system tracks a barcode etched into the guide wire and calculates the depth of the wire in bone based on the section of barcode visualized. This allows a resident to use a standard surgical driver and to place the wire close to bone while they are drilling. When a resident wants to see a fluoroscopic image, they simply retract the drill an inch or two to expose more of the barcode pattern. This simulator is an extension of a prior wire-navigation simulator that used electromagnetic tracking to generate artificial fluoroscopic images. This previous simulator was able to establish construct validity between a novice group, first year orthopaedic residents and engineering students, and an expert group, fourth year orthopaedic residents and several faculty surgeons, in a hip wire navigation task [71].

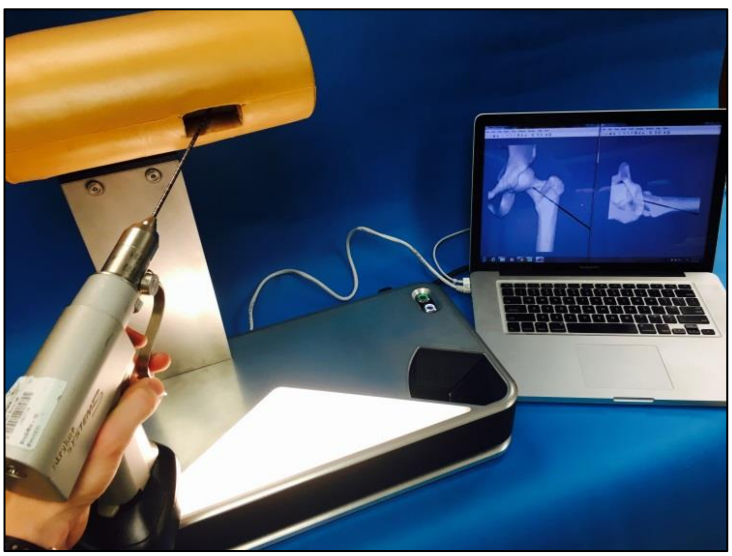


Figure 12: The wire navigation simulator shown on the left. The laptop on the right demonstrates an example of computer generated fluoroscopic images that may be presented to a resident.

Summary

Simulation based skills training has come a long way in the last two decades. It is clear that competency based medical education (CBME) is the direction that the orthopaedic community both wants to head in and is working towards. However, there is still a great deal of work that must be done before CBME is widely adopted. To date, no study, whether examining arthroscopic skills or wire navigation skills or others, has demonstrated that training on a simulator improves or impacts performance in the operating room. There are many different metrics being tested to evaluate operating room performance on a variety of orthopaedic surgical techniques, but there is yet to be a clear measurement which objectively, reliably, and repeatedly can evaluate performance in the operating room. The body of work described in this dissertation aimed to accomplish four main goals to help advance the mission towards competency based medical education. The first aim was to demonstrate that training on the wire navigation simulator can improve performance in a higher-level simulation; in essence, to demonstrate transfer validity. The study presented in Chapter 2 which addresses this aim has been accepted for publication in the Clinical Orthopaedics and Related Research (CORR) journal. The second aim was to extend the wire navigation platform to be able to train residents on a multitude of wire navigation applications. The study presented in Chapter 3 which addresses this aim has been accepted for publication in the Journal of Medical Devices. The third aim was to evaluate wire navigation performance in the operating room to better understand the current learning curve associated with hip wire navigation. The fourth and final aim was to broadly disseminate the wire navigation training and demonstrate that it can be used to improve orthopaedic surgeons' performance on the skill of wire navigation.

CHAPTER 2: SKILL TRANSFER FROM WIRE NAVIGATION TRAINING ON A SIMULATOR TO PERFORMANCE IN A MOCK OPERATING ROOM ENVIRONMENT

Introduction

Surgical skills training and simulation are playing an increasingly important role in modern orthopaedic surgical education. Following a 2013 mandate by the American Board of Orthopaedic Surgery (ABOS) to include laboratory-based training of basic surgical skills for first year residents, programs have moved to incorporate simulation into their surgical skills education for residents [7]. Early reports following this mandate examined a variety of simulation and training modalities [10-12, 64, 72]. Most of these published studies focused on demonstrating construct validity, typically defined as the capability to differentiate between the performance of novices and experts on the simulated task. Demonstrating construct validity is an important first goal for a task simulation, but the ultimate goal of a training simulation is to improve surgical skills in the operating room (OR). Evidence of simulator training improving OR performance is known as transfer validity; it represents the strongest evidence of the benefit of training with a simulator. The vast majority of simulation-based validation studies in orthopaedics have focused narrowly on arthroscopic surgical skills, just one of many skills required of orthopaedic surgeons [73].

The focus on arthroscopy may be a consequence of the relative accessibility of the relevant simulation technology. A review article published in 2014 examined 14 papers that had studied arthroscopic skills training, citing that "arthroscopic education provides a unique area for the evaluation of simulation as a teaching tool because … [it] can be easily mimicked in an artificial setting" [73]. A different skill that is equally central to orthopaedic surgery, yet has traditionally been difficult to simulate, is fluoroscopic wire navigation; the task of directing a surgical wire through a bone under fluoroscopic guidance. The ability to surgically navigate a wire under fluoroscopic visualization is a critical step in the operative management of hip fractures, making that surgical task an excellent application for training orthopedic residents. Hip fractures are common injuries: there are roughly 300,000 in the U.S. each year [57]. Many times, post-graduate year 2 (PGY2) and year 3 (PGY3) residents are involved, and indeed participate, in the wire navigation aspect of these procedures as part of their surgical training. Several meaningful metrics of surgical performance that have been shown to affect clinical outcome can be readily measured in the OR. Taylor et al. demonstrated that measurements of a surgeon's wire navigation time and tip apex distance (TAD) correlated with greater surgical experience in the OR [61]. Being able to quantify surgical performance in the OR in an objective and reliable manner is imperative in establishing a connection between performance in a simulated environment and performance in the actual OR.

An important logistical hurdle many residency programs face in training wire navigation is gaining sustained use of a C-arm fluoroscopy unit and an appropriately trained fluoroscopy technician. Also, hazards related to radiation exposure may create logistical hurdles. Practicing wire navigation skills should ideally take place in an environment that does not unduly expose residents to radiation as they hone their skills. In an attempt to surmount these challenges and provide a readily available, radiation-free training environment, our group has developed a hybrid reality wire navigation simulator. This simulator uses an optical system to track a guide wire relative to a fixed bone model [70]. A prior version of this wire-navigation simulator that used electromagnetic technology to track the guide wire was shown to possess construct validity in differentiating performance between a novice group (first year orthopaedic residents and engineering students) and an expert group (fourth year orthopaedic residents and several faculty surgeons) in a hip fracture wire navigation task [71].

The step from showing construct validity to establishing transfer validity is important. Transfer validity indicates effective utilization of the simulator as a training tool, i.e., integrating the simulator into a beneficial teaching curriculum. The benefit a resident gets from a training experience with a simulator, the magnitude of the effect that might be transferred to the OR, may be dominated by the productive feedback and facilitation of that resident's extended, deliberate practice with the simulator. Deliberate practice involves task repetition with direct and clear feedback that informs a trainee of ways to improve their performance [39, 40]. Proficiency-based training demarcates improvement by defining goals that a trainee attain, such as "reach at least the proficiency level (on two consecutive trials) before progressing to *in vivo* practice" [37]. Proficiency training has been shown to be an effective tool in teaching trainees a variety of surgical skills [6]. Maximizing the effectiveness of the training is likely to improve the likelihood of being able to demonstrate a simulator's transfer validity.

With this in mind, the present study aims to answer two lingering and important questions. First, which training method lead to the greatest improvement on a wire navigation task? Second, is a resident's performance on a wire navigation simulator predictive of their performance on a higher-level simulation task? Answering these questions would help other residency programs better understand how to most effectively use a simulator to train their residents on wire navigation tasks. This study has been accepted for publication in the Clinical Orthopaedics and Related Research (CORR) journal.

Methods

For this study, 55 PGY1 orthopaedic surgery residents were recruited to participate from four different residency programs in the Midwest Orthopaedic Surgical Skills (MOSS) Consortium. Participants included residents from the University of Iowa, the University of Minnesota, the Mayo Clinic, and the University of Nebraska. Residents were divided into three groups, each receiving different methods of training (**Figure 13**).

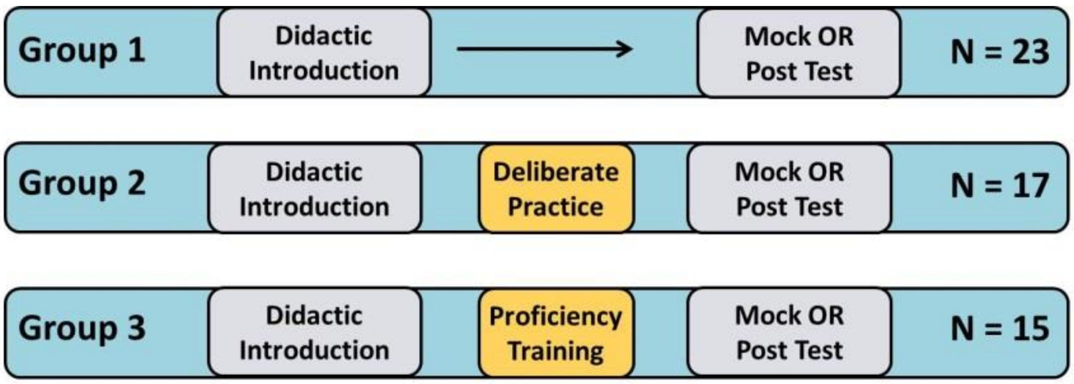


Figure 13: The study design of the experiment is shown here.

Traditional Training

The Traditional Training Group consisted of twenty-three residents. Residents received what was considered to be traditional training followed by a performance assessment. The traditional training began with a didactic presentation that described the methods for placing a guide wire in the treatment of an intertrochanteric hip fracture followed by a video from the ABOS module on *Fluoroscopic Knowledge and Skills* demonstrating the proper technique [74]. After the didactic presentation, residents proceeded to a performance assessment in a mock OR (**Figure 14**). A radiopaque left proximal femur (Sawbones part # 1130-21-22) was placed inside an artificial soft tissue envelope to simulate a patient's anatomy in an intertrochanteric fracture case. A C-arm was used to provide fluoroscopic visualization of the bone and of the guide wire as it was drilled into bone. Residents received instruction to place the guide wire in a center-center position with the goals of minimizing the tip apex distance (TAD), wire navigation duration, and the number of fluoroscopic images requested. Residents had discretion in switching between anteroposterior (AP) and lateral images at their request. The viewpoints of these images were standardized for all residents. Performance feedback was not provided during the task. Completion of the task was determined by the resident participant when they felt that they had achieved an acceptable wire position. The final AP and lateral fluoroscopic images were recorded for later measurement of performance.

Figure 14: A resident is shown placing a guide wire using fluoroscopy in the mock OR environment.

Deliberate Practice

The Deliberate Practice Group consisted of seventeen residents. This group received the same traditional training as the Traditional Training Group. However, this group then also participated in a deliberate practice session with the wire-navigation simulator. Following focused practice, they proceeded to the mock OR for performance assessment, following the same procedure as the Traditional Training Group.

The simulator training session involved 30 minutes of practicing the hip wire navigation task on the simulator (Figure 12). The simulator provided real-time feedback to the residents on their wire position relative to the ideal while they practiced navigating the wire (**Figure 15**). The interface provided a generic target zone, overlaid in green, onto the computer-generated AP and lateral images of the femur. This provided residents with immediate visual feedback on how their position compared to the ideal center-center position. Additionally, residents could see what their TAD was as they were driving the wire into bone. Each resident went through the practice module of the simulator a minimum of three times. Following this, the residents practiced driving the wire with the simulator one time without any feedback. This was designed as a transition between the simulated environment and the mock OR environment. Upon completion of the practice session, residents placed a guide wire in the mock OR with the same instructions provided to the Traditional Training Group.

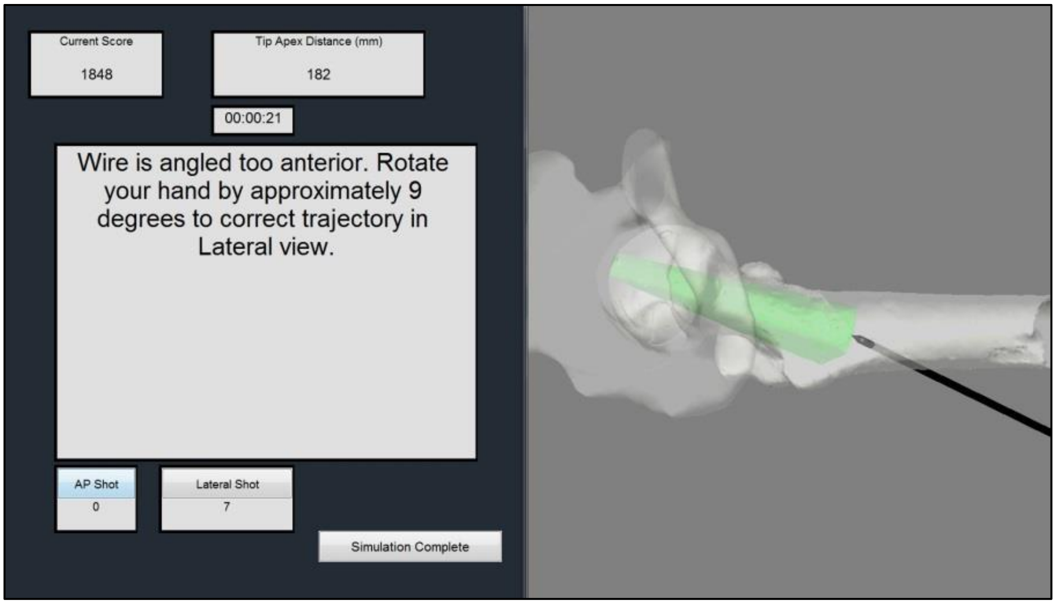


Figure 15: An example of feedback given to a resident using the simulator in Group 2.

Proficiency Training

The Proficiency Training Group consisted of fifteen residents. This group received the traditional training, then a proficiency-based training, and finally the mock OR performance assessment. The proficiency-based training consisted of a series of tasks implemented first in an online format and then on the wire navigation simulator. These tasks emphasized identifying visual cues and achieving two- and three-dimensional wire positions believed to be essential skill elements for wire navigation. Residents were required to achieve specified levels of proficiency before they could advance to the next training exercise.

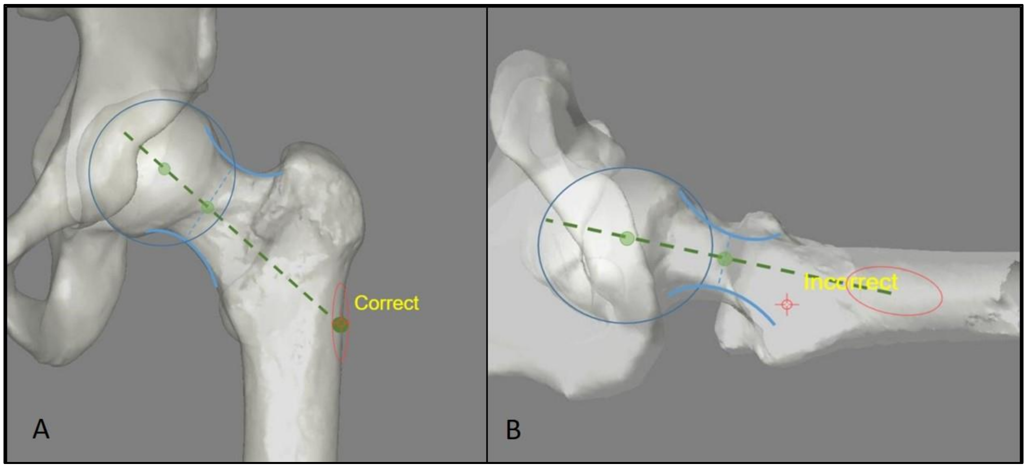


Figure 16: The feedback given to residents in the online module. A circle fit to the femoral head helps identify the center of the femoral head. A green circle illustrates the center of the femoral neck. Connecting these two points shows the trajectory that establishes where the wire should enter the lateral cortex of the femur. Images A and B show a variety of fluoroscopic perspectives that were shown in the online training.

In the first set of proficiency-based tasks, residents were taken through a series of AP and lateral images of left proximal femurs in an online module. In these images, residents were asked to identify where the guide wire should ideally enter the femur and where the tip of the wire should ideally end up in the femoral head. When a resident clicked on the image a feedback screen would immediately appear showing the resident if they had chosen the right locations on the femur (**Figure 16**). A circle fit to the femoral head helped identify the center of the femoral head. Lines outlining the femoral neck help identify the center of the femoral neck. Connecting these two points provided the trajectory needed to identify the proper start and end points of the wire for the procedure and achieve a center-center position. In total, residents reviewed twenty different fluoroscopic viewpoints. By showing subtle variations in AP and lateral fluoroscopic projections, residents were forced to respond to the viewpoint variations by accommodating their starting points and trajectories, a frequent challenge in the actual surgical setting. Residents were required to correctly identify start and end points on 80% of the images prior to moving on to the next task.

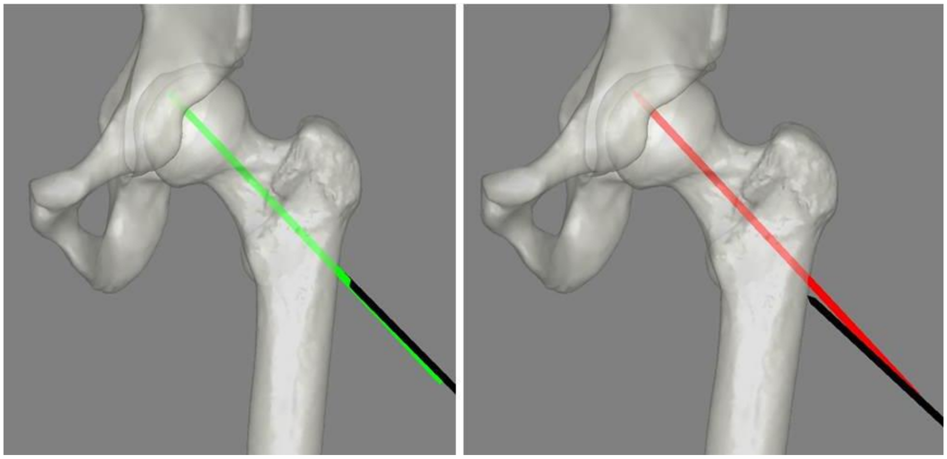


Figure 17: Examples of feedback images displayed during subtask training are shown. The left image shows feedback that reinforces the correct trajectory of the black wire. The right image shows feedback given when a wire is on the wrong trajectory.

Following the online module, residents worked on the simulator, where they were taken through a series of wire navigation sub-tasks, an approach distinct from the simulator training for the Deliberate Practice Group. These sub tasks involved interpreting the computer-generated fluoroscopic images to properly place the guide wire on the Sawbones femur so that it was at the correct starting point in the AP plane, the correct starting point in the lateral plane, had the correct trajectory in the AP plane, and the correct trajectory in the lateral plane. This was intended to be an extension of the online training; however, it included testing the resident's knowledge and ability to correctly move their hands in 3 dimensions. For each sub task, the resident was required to correctly position the wire on bone, either looking at wire starting point or wire trajectory, a total of 3 times before they could move on to the next task. A starting point was deemed correct if the wire tip was within 5mm of the ideal placement. A wire trajectory was considered correct if the wire was angled within 3° of the ideal wire angle. This information was measured and collected automatically by the simulator (**Figure 17**).

The physical Sawbones remained fixed on the simulator throughout the training, so to simulate the different viewpoints from the online module, the virtual position of the femur was instead moved with each new attempt. In this way, residents were forced to rely entirely on the fluoroscopic images presented to them and move their hand and wire accordingly. Once a resident had completed all the sub task trainings, he or she was given one final trial at driving the wire into bone. Residents in this group did not receive the general practice environment that included wire positioning and TAD feedback that were a part of the deliberate training for Deliberate Practice Group. Upon completion of the practice session, residents went into the mock OR and placed a guide wire with the same instructions provided to the other groups.

In all the tasks, whether on the simulator or in the mock OR, residents were graded on their wire navigation time, the number of fluoroscopic images requested, and their TAD. In the data taken from the mock OR, the TAD was measured using the methods laid out in Johnson et al [59]. The TAD measurements were made by two independent reviewers who examined the final AP and lateral images of the residents in the mock OR. The average of their independent measurements was recorded as the final measured TAD. Work by Taylor et al. has shown this methodology of measuring the TAD to be reliable and reproducible between different users [75]. The TAD from the simulator trials was automatically computed as part of the simulator software. Additionally, for each trial that residents completed on the simulator, the wire tip position and trajectory were stored for each image requested. This allowed for a post training analysis of how residents directed their wire through bone to reach the apex on the simulated model.

Results

The data in **Table 1** summarize the performance of the residents during their final wire navigation attempt on the simulator and in the mock OR. The main comparison of interest between the three groups was in their mock OR performance following the different types of training. One-way ANOVA tests were performed in SAS Enterprise (version 7.12) to examine the variance between the three groups on TAD, use of fluoro, and total time. All three tests were significant (TAD: F = 11.15, $p < 0.0001$; use of fluoro: F = 4.38, $p = 0.017$; total time: F = 3.37, $p = 0.042$ for total time). Table 1 tabulates the means and standard deviations of each performance metric for each data collection and presents the pair-wise comparisons of each of the three groups [mock OR data] with post-hoc, student's two-sided T-test for each of the performance metric. Both the Deliberate Practice Group and the Proficiency Training Group had a significantly lower TAD than the Traditional Training Group ($p = 0.001$ and $p < 0.001$). The Proficiency Training Group used significantly more fluoroscopic images in the mock OR than both the Traditional Group ($p = 0.041$) and the Deliberate Practice Group ($p = 0.012$).

Table 1: Summary Performance Measures - values are displayed as mean (standard deviation)

Group	Wire Navigation Time (minutes)	Number of Fluoroscopic Images	TAD (mm)
Traditional Training: Mock OR Data (n=23)	4.03 (2.07)	21.61 (11.97)	23.58 (7.34)
Deliberate Practice: Simulator Data (n=17)	3.03 (1.30)	15.71 (4.98)	11.53 (5.16)
Deliberate Practice: Mock OR Data (n=17)	3.83 (1.67)	19.12 (7.85)	15.88 (5.35)
Proficiency Training: Simulator Data (n=15)	4.44 (2.21)	30.67 (14.23)	10.80 (5.10)
Proficiency Training: Mock OR Data (n=15)	5.99 (3.75)	31 (14.35)	15.30 (4.21)

Testing for Significant Differences in Mock OR Performance:

Traditional vs. Deliberate	P = 0.747	P = 0.471	**P = 0.001**
Traditional vs. Proficiency	P = 0.089	**P = 0.041**	**P < 0.001**
Deliberate vs. Proficiency	P = 0.061	**P = 0.012**	P = 0.742

Another relationship of interest was between the performance of residents on the simulator and the same resident's performance in the mock OR. Residents TAD values had a correlation coefficient of $R^2 = 0.15$ ($p = 0.030$) between these two environments. The wire navigation duration had a correlation coefficient of $R^2 = 0.61$ ($p < 0.001$). Finally, the number of fluoroscopic images requested had a correlation coefficient of $R^2 = 0.43$ ($p < 0.001$).

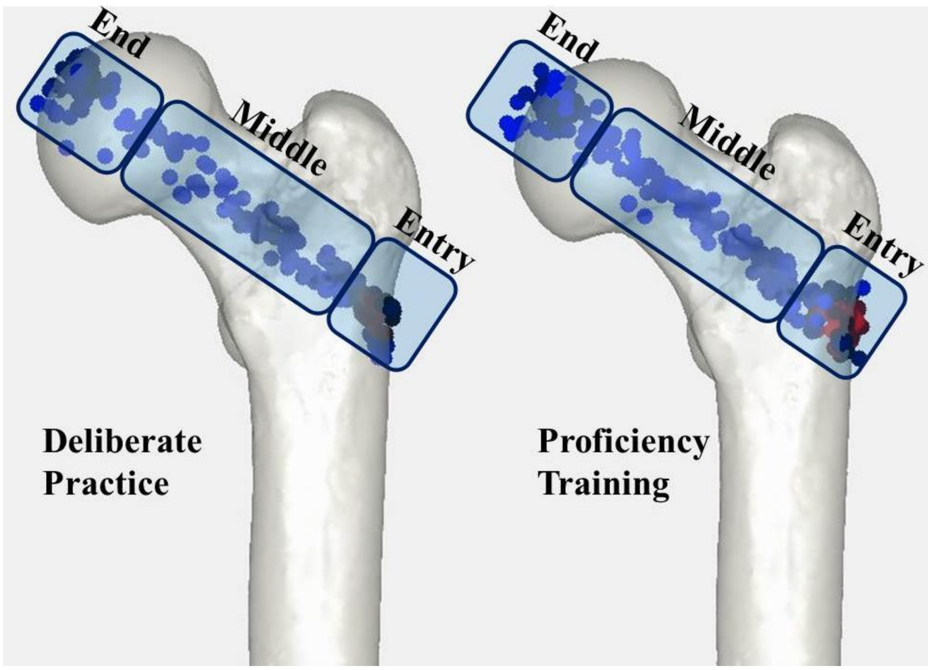


Figure 18: The wire position of each requested image for all residents during the transition trial on the simulator are plotted. The densities of where more wires were placed in bone are represented by red areas. Light blue areas represent fewer wires in that position of bone.

The residents' performance on the simulator task was also analyzed to better understand differences between the two groups that received simulator training. In the trial that immediately preceded the mock OR, where both groups placed a wire on the simulator without any added feedback, the position of the guide wire for each image requested was recorded. **Figure 18** shows the graphical representation of the wire positions for the trials of the residents in both groups. The wire positions revealed that residents in the deliberate practice group used an average of 3.7 images in the entry area of the wire navigation region (26% of total image count), 6.8 images in the mid area (47% of total image count), and 4 images in the end area (27% of total image count). In contrast, residents in the proficiency training group used an average of 11 images in the entry area (38% of their overall images), 11.6 images in the mid area (41% of total image count), and 6.1 images in the end area (21% of total image count). Running these data through a one-way ANOVA showed that there was a significant difference between the proportion of images used in the entry area between the Deliberate Practice and the Proficiency Training groups (F value = 4.81, p = 0.036). There was no significant difference, however, in the proportion of images used in the mid or end region between these groups.

Discussion

The "see one, do one, teach one" training model is challenged by the desire to provide safe and reliable care to patients. An alternative, simulation-based training model more closely supports the "first do no harm" mantra [76]. A new, "learn, see, practice, prove, do" training model may more closely align the goals of increased patient safety with the need to cultivate the next generation of surgeons [38]. With this model, residents first learn about procedures, see them performed in the OR or in a video, practice them in a simulated environment, and demonstrate that they have achieved a certain level of proficiency on the task, all before first attempting the procedure in the OR.

We found that residents who practiced the wire navigation task on the simulator achieved a lower TAD in a mock OR environment compared to residents who simply learned about and watched a video of the procedure. Wire navigation is a complex task that involves coordinating hand movements with 2D fluoroscopic images to control the 3D position of a wire in a patient. This may explain why residents who only practiced the cognitive portion of the task did not perform as well as the residents who had practiced both the cognitive and psychomotor portions of the task.

While it makes intuitive sense that simulation training would produce improved performance in comparison to traditional didactic training, a question of how to best implement the simulation training also exists. In this study, we explored the difference between two, well-documented training methodologies: deliberate practice and proficiency training. A somewhat surprising result from this study was the difference in performance between residents in the Deliberate Practice Group and residents in the Proficiency Training Group. Broadly speaking, these two groups both improved their TAD when compared with traditional training; however, the residents in the Proficiency Group used significantly more fluoroscopic images and took more time to achieve this result than those in the Deliberate Practice Group. Our expectation was that residents in the proficiency training program would improve not only their TAD, but also use fewer images and take less time to perform the procedure than those in the deliberate practice training. One explanation may be that the residents in the Proficiency Training Group developed a more structured technique for placing the guide wire. Going through the process of finding the proper starting point in the AP plane, adjusting the trajectory in the AP plane, and repeating this all again in the lateral plane may have taken more time during this procedure. Alternatively, it may be that the additional complexity of the varying fluoroscopic viewpoints that was introduced in the online training for the Proficiency Training Group led them to be more attentive to positioning details that participants in the Deliberate Practice Group may have ignored. If so, this effect might be more related to a subtle difference in our training materials than to the broad training approaches themselves. Nevertheless, perhaps with a more advanced surgical technique, like placing an iliosacral screw, these residents may be better equipped to take on the task than those who serendipitously developed their own procedure for the wire navigation.

Figure 18 illustrates that participants in the two groups spent their images in different ways. Residents in the Proficiency Training Group spent 38% of their images in the vicinity of their entry point, indicating that they were following the procedure learned during the proficiency training. Residents from the Deliberate Practice Group, on the other hand, only spent 26% of their images in this region. Although this was not the intended result of the training, it does illustrate two potential strategies for placing the guide wire. Taking more images before the wire enters bone could potentially be beneficial for a patient with poor bone quality that may not withstand redirecting the guide wire very well. On the other hand, this may lead to taking more images overall and more radiation exposure to the patient as it may be harder to judge the wire trajectory so far from the ultimate target of the femoral head apex. In any case, it is clear that the style of training had an impact on the performance of the residents and the strategies they used. If we wanted to incentivize achieving a low TAD while using fewer images that may simply require a different set of proficiency standards or training tasks.

One of the main strengths of this study is the assessment platform used to test residents on their wire navigation skills. Many simulation studies have evaluated the construct validity of the simulator, but few have established a transfer of skill [11, 55]. In this study, our group specifically focused on testing in a higher-level simulation platform, which we refer to as the mock OR. This is a key point to note because it helps establish the transfer of skills from the wire navigation simulator that the residents trained on to the mock OR. Further, in examining the relationship between a resident's performance on the simulator and in the mock OR, we found a significant correlation between the TAD, number of fluoroscopic images, and overall time of the wire navigation procedure. This suggests that in future evaluations the simulator itself could be used as the testing platform, instead of needing to test in the mock OR environment. Another strength would be that the metrics used to assess the residents are all objective and able to be measured in the OR, as well. Having performance metrics that can be comparably measured in a simulated environment and in the OR is one of the keys to eventually demonstrating that simulation training has a positive impact on performance in the OR.

A potential weakness of this study may be the sample size of the residents trained with the simulator. Although there were 55 participants in total, the two groups that received training only contained 17 and 15 residents, respectively. This is a common issue throughout all of surgical skills training literature as each program only has limited numbers of residents in each year. Our group focused on first year residents since they have no experience in the OR and would receive the most benefit from training. That said, our group endeavored to increase the number of participants by collaborating with multiple programs around the Midwest. Another potential critique of this study might be that the bone model used in this study, both on the simulator and in the mock OR, did not have a fracture as part of the model. We felt that reducing and maintaining a fracture reduction in the hip was a separate task to that of driving a wire through the center of the femoral neck and into the femoral head. Although fracture reduction is a key skill in orthopaedics, it was not the focus of this study nor the simulator.

Wire navigation is a key skill in orthopaedics that has a broad range of applications. Future studies will aim to address the multitude of clinical scenarios in which wire navigation is a key component. Navigating a wire into the femoral head as part of surgically treating an intertrochanteric femur fracture is a common procedure that is also relatively simple and created a solid foundation to begin studying training residents on this skill. More complex and challenging procedures, such as placing a wire for an iliosacral screw, intramedullary nail, or multiple wires for a pediatric elbow fracture, may help in training residents as they advance in their residency programs.

This study successfully demonstrated transfer validity from a wire navigation simulator to a mock OR. Further work will need to be done to show that the skills developed on the simulator can be further applied to the actual OR and real patient cases. Regardless, this is an important first step that will help bridge the gap between orthopaedic resident simulation training and improved surgical performance in the OR.

CHAPTER 3: AN EXTENSIBLE ORTHOPAEDIC WIRE NAVIGATION SIMULATION PLATFORM

Introduction

Over the past two decades, the demand for simulation-based surgical skills training in orthopaedics has steadily grown. This demand has been driven by numerous factors. In 2003, work hour limits for residents were implemented, creating a safer work environment for employees and patients, but potentially limiting the amount of surgical experience a resident can gain during their time as a trainee [2]. In July 2013, the American Board of Orthopaedic Surgery, the body responsible for certifying orthopaedic surgeons as competent to practice, required residency programs to dedicate time and space to training first year residents for "skills used in the initial management of injured patients and basic operative skills to prepare residents to participate in surgical procedures" [8].Some countries have even begun transitioning to a competency-based training program that advances residents when they demonstrate adequate surgical skills on various simulated tasks [15].

As simulation in orthopaedics has become more common, the number of tools available to train surgeons has steadily increased. However, a majority of the simulators available to programs today focus on a narrow range of topics, mainly those involved in arthroscopic surgery [11, 12, 73]. One task that has been difficult to train residents on in a safe and cost-effective manner is wire navigation, an orthopaedic surgical skill used in many surgical procedures. This skill enables surgeons to drill a surgical wire, typically a long, small diameter steel pin, along a specified path through bone, visualized using 2D intra-operative fluoroscopic images. Wire navigation requires sophisticated visuospatial cognitive skills because the surgeon must visualize and navigate the patients' anatomy. In some situations, placing the wire is the final task; in others the wire serves as a guide for subsequently placed cannulated implants. Regardless of the situation, the placement of the wire in the bone directly influences the surgical result for the patient. To highlight the importance of this skill in the field of orthopaedics, the American Board of Orthopaedic Surgery includes wire navigation as one of the core competency skills [7].

Our team previously presented the design of a wire navigation surgical simulator dedicated specifically to hip wire navigation [70]. Hip wire navigation was chosen as an initial application because of the many incidences of hip fractures in the US each year [57]. However, many other surgical procedures, including some outside of orthopaedics, rely on wire navigation. Because the visualization challenges faced by the surgeon are largely task-specific, training for many of these procedures requires dedicated simulators. Instead, we have modified the existing simulator to accommodate new surgical procedures, while still exercising the underlying skill of wire navigation. This paper will explore the criteria used to develop new simulator applications, introduce an example of developing a new application, and present data from orthopaedic residents who trained with the new application. This study has been accepted for publication in the Journal of Medical Devices.

Methods

The simulation platform

The wire navigation simulator is an optical system that uses image processing algorithms to track the position of the surgical instrument, in this case a guide wire, relative to a fixed artificial bone [70]. A stereo camera system feeds image pairs to a repository on a connected laptop. When a resident requests a fluoroscopic image, the latest image pair is pulled from the repository and processed. The cameras are fixed in space, so the guide wire trajectory can be calculated in 3D space from any stereo image pair. A parallel line pattern etched around the circumference of the wire allows the algorithm to estimate the axial position of the wire. From these two pieces of information, the wire trajectory vector and the axial position, a simulated fluoroscopic image can be rendered. A foam shell encases the artificial bone that is mounted on the simulator, thus obstructing the bone geometry from the surgeon. This forces the trainee to focus primarily on the simulated fluoroscopic images when navigating the wire, as is the case in the operating room. Residents use the same drill to drive the pin into the artificial bone as they would use with a real patient in the operating room.

Criteria for Developing New Applications

Expanding the simulator platform to simulate other surgical procedures involves comparing the capabilities of the simulator's optical system with the needs of a specific application. The guide wire or other surgical instrument must be viewed and focused in both camera views. The overlapping, in-focus region is called the working volume. Currently, the simulator has a frustum-shaped working volume that expands out from the entry location on the surface of the artificial bone. The frustum has a width of approximately 102 mm (4 inches) at the surface of the soft tissue envelope, a depth of approximately 76 mm (3 inches), and a height of approximately 102 mm (4 inches) (**Figure 19**).

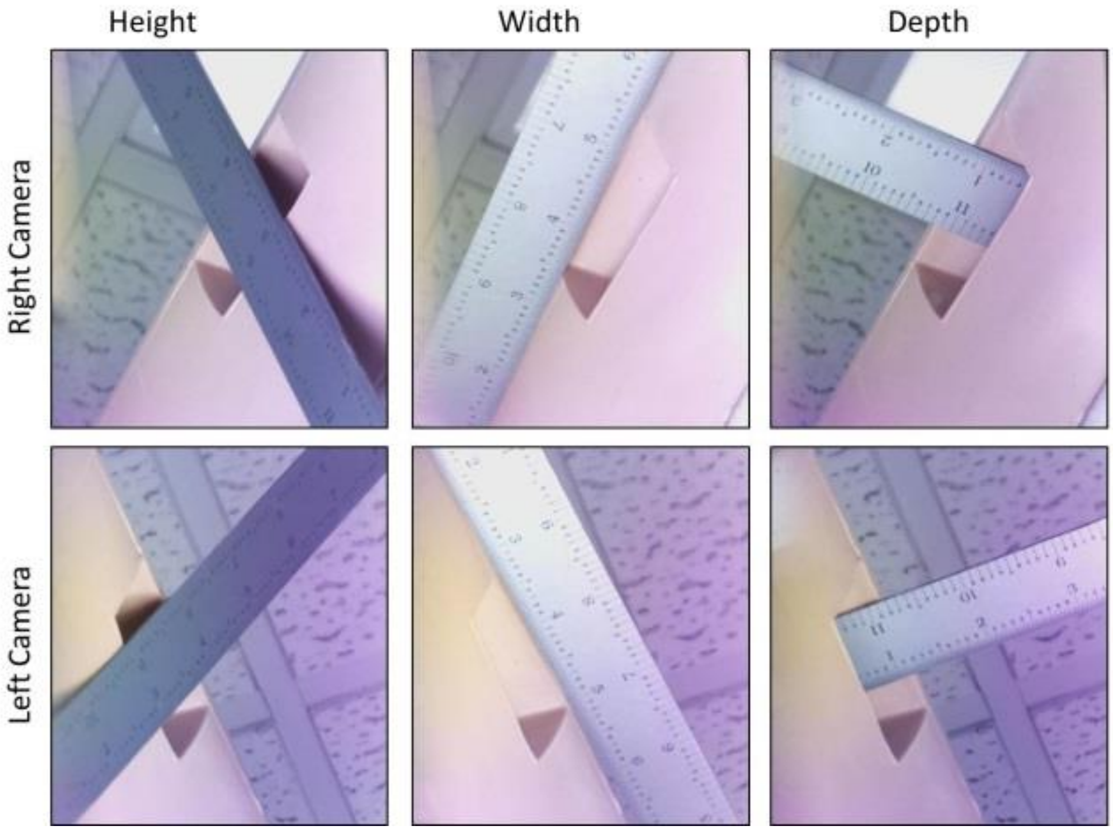


Figure 19: The dimensions of the working volume are shown here. The ruler is positioned against the soft tissue which encloses the bone of the simulator. The space beyond the soft tissue defines the region of the image in which the wire can be seen by the camera system to calculate the wire position in bone.

A second optical system characteristic is the precision with which the wire can be tracked. Some procedures may require more precise tracking than others. For example, the placement of a surgical wire in a pediatric elbow requires greater precision than in an adult hip because of substantial differences in the sizes of the pediatric elbow and the adult hip. Also, more advanced surgeons may be sensitive to minor position adjustments; so different types of users may require different levels of performance. To quantify the simulator's accuracy, we compared known wire locations with the locations calculated by the optical system. The reference wire locations were defined by pre-drilled holes in two target bones. The holes were placed so that the trajectories of wires placed in the holes spanned the simulator working volume. Copper tube liners with an inner diameter matching the test wire outer diameter were placed in each hole to ensure that repeated wire insertions would not alter the hole dimensions. The copper tubes were plugged at their bottom with a metal stopper to ensure repeatable wire insertion depth (**Figure 20**).

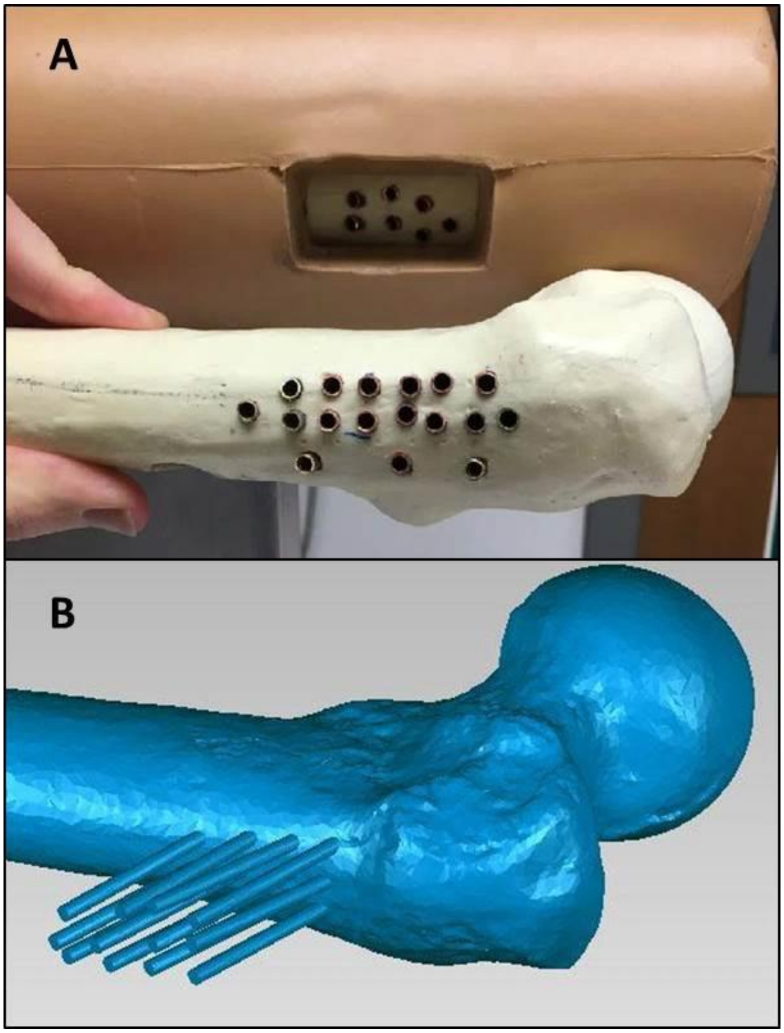


Figure 20: (A) The copper tubes used to create reproducible wire positions are shown here. The tubes are cemented in bone to create a reproducible slot that a wire can be inserted to for testing. (B) A laser scan of one of the bones with the wire positions for each slot. The laser scan defined the wire trajectory and position for each copper tube.

A laser scan of the simulator-mounted bone was acquired using a FaroArm laser scanner (FARO LLC). This defined the 3D location of each wire relative to the simulator frame (Figure 20). The FaroArm has a scanning precision of 0.024 mm (0.00094 inches) [77]. A short wire of a known length was placed into each lined hole and individually scanned to create a dataset of reference trajectories and tip positions for each prepared hole. These 29 reference positions were then compared with the associated wire trajectories and axial positions calculated by the optical system. Ten optical measurements were taken for each reference position. Because the system is sensitive to background light and image noise, this test was done in an ideal laboratory setting, as well as at 4 other locations selected as being typical for standard operation.

A third consideration to be evaluated when expanding the system to other wire navigation applications is the tactile feedback the user experiences when drilling into high density cortical tissue on the bone's exterior versus the soft cancellous tissue in the bone's interior. Oftentimes, detecting the transition from cortical to cancellous drilling is an important component of a surgical task. For example, surgeons often use this transition as a cue to know that the wire has passed through the bone and is at the cortical wall on the opposite surface.

The first new wire navigation application was selected in consultation with a team of orthopaedic surgeons. Together we developed a list of potential wire navigation procedures, including iliosacral screw fixation, distal radius fracture fixation, the fixing of pediatric elbow fractures, the placing of pedicle screws for spinal fusion, fixing proximal humerus fractures, and tibial plateau fractures. The iliosacral screw fixation procedure was selected because it was similar in scale to the hip wire procedure (but it demands higher precision from the trainee), it introduces imaging projections that are not orthogonal to the floor (demands a deeper geometric understanding), and because we believed that the residents would be excited to try this challenging, yet relatively rare procedure [78]. In addition, there isn't currently a widely-adopted model for training residents on this task. A group out of the Orlando Regional Medical Center published on using a Sawbones pelvis model with a C-arm (the machine that generates fluoroscopic images) to train residents [79]. However, access to a C-arm can be limited and expensive, not to mention the exposure to radiation presents a hazard of training. Other current options for training residents on this task include cadaveric training, which can often be very expensive and would also include the use of a C-arm and radiation exposure [80].

Iliosacral Screw Fixation

The surgical technique for iliosacral screw fixation is a minimally invasive procedure that involves placing a guide wire perpendicular to the sacroiliac (SI) joint and beyond the midline of the sacral body [81]. There are several factors that make this surgically challenging. First, the corridor for placing the implant properly can be quite narrow. The corridor is surrounded by several vulnerable nerve roots and the spinal cord. Improper placement of the implant has been reported to be between 18 to 24% [82-85]. Also, pelvis anatomy is highly variable between patients, requiring the surgeon to adjust the trajectory from patient to patient and even rejecting some patients for the procedure.

A literature search revealed the parameters for ideal implant position within the anatomic variability typically encountered by surgeons. One study analyzed a set of pelvic scans to define an ideal path and acceptable corridor parameters relative to the anatomic variability [19]. This analysis defined the working volume as a cylindrical corridor through the sacral body that does not intersect with the nerve roots. The analysis indicated that across multiple pelvic anatomies, this corridor has an average diameter of 14 mm (0.55 inches) and an average length of 150 mm (5.91 inches) [86].

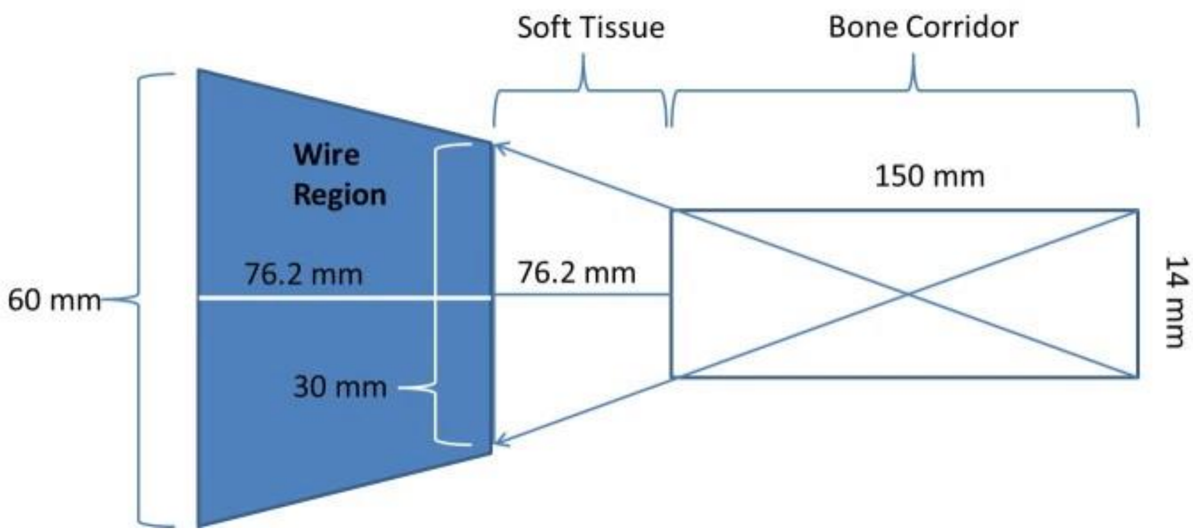


Figure 21: This shows the potential working volume for the iliosacral screw placement on the simulator. The area in dark blue defines the potential wire regions that would be used to successfully guide a wire into the iliosacral corridor.

This corridor defines the wire position within the bone, but the surgeon must also guide the wire percutaneously through the surrounding soft tissue. Therefore, to fully define the working volume for this procedure, the path of the wire taken through the soft tissue to reach the bone surface must also be considered. For our implementation, we designed a soft tissue envelope with roughly 76 mm (3 inches) of material separating the surface of the bone from the area of wire visible to the camera. The resulting space occupied by the wire to reach this corridor in bone is then a frustum with a front width of 30 mm (1.18 inches), a depth of 76.2 mm (3 inches), and a back-end width of approximately 60 mm (2.36 inches) (Figure 21). This working volume allows the system to accurately track any wire that is within the safe corridor through the sacral body.

A mounting bracket was designed to hold an artificial hemi-pelvis, (Sawbones, model #1294-29) so that the working volume needed for the iliosacral screw procedure intersected with the working volume of the simulator's optical system. The mounting bracket consists of 3 main components: an L bracket to provide a flat and rigid surface to mount onto, a 3D printed shell matching the bottom surface of the sacral body anatomy, and a 3D printed shell matching the top surface of the sacral body anatomy (**Figure 22**). Two specially designed surfaces reproducibly position the hemi-pelvis on, and clamp the hemi-pelvis to, the simulator.

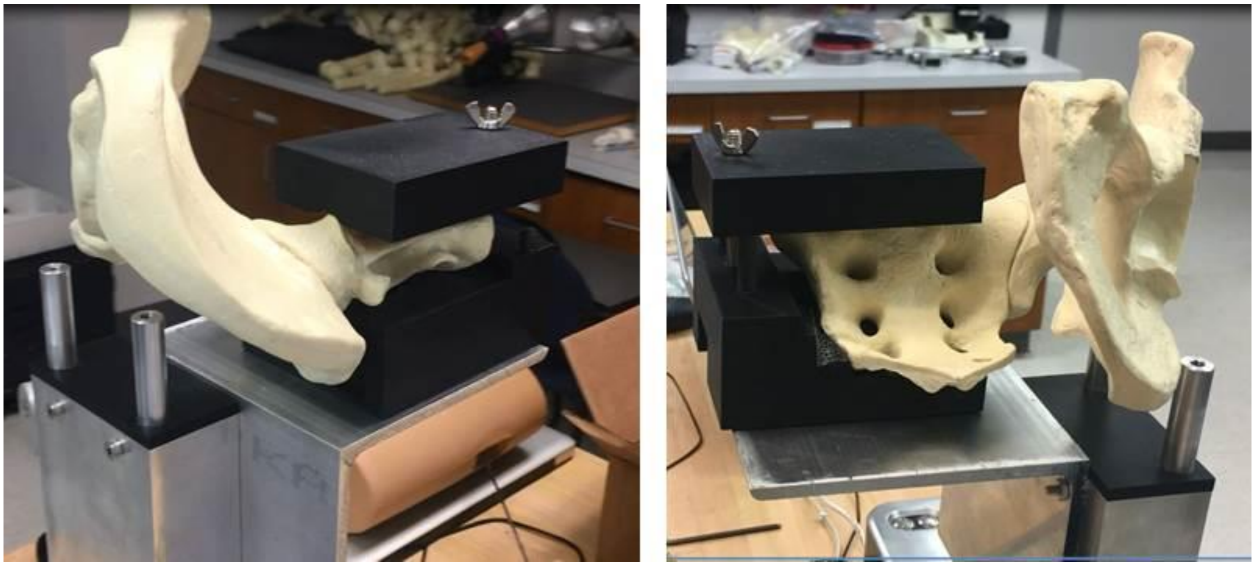


Figure 22: These oblique views show the pelvic bone mounted to this simulator base. The soft tissue is not shown here. The black 3D printed materials act as a vice to hold the pelvic model in place while residents drill into the bone.

The simulator's software was modified in order to display the proper images to the residents. CT data from the Visible Human Project® (https://www.nlm.nih.gov/research/visible/visible_human.html) was used to generate the fluoroscopic images used in the procedure (**Figure 23**). The procedure uses three main perspective views: an outlet view, inlet view, and lateral view. The inlet and outlet views are the main views used to navigate the guide wire. These two views are at an oblique angle relative to the patient, making it difficult to interpret wire trajectory and necessary adjustments. Learning to properly assess the wire trajectory in these views is considered one of the main skills of this procedure.

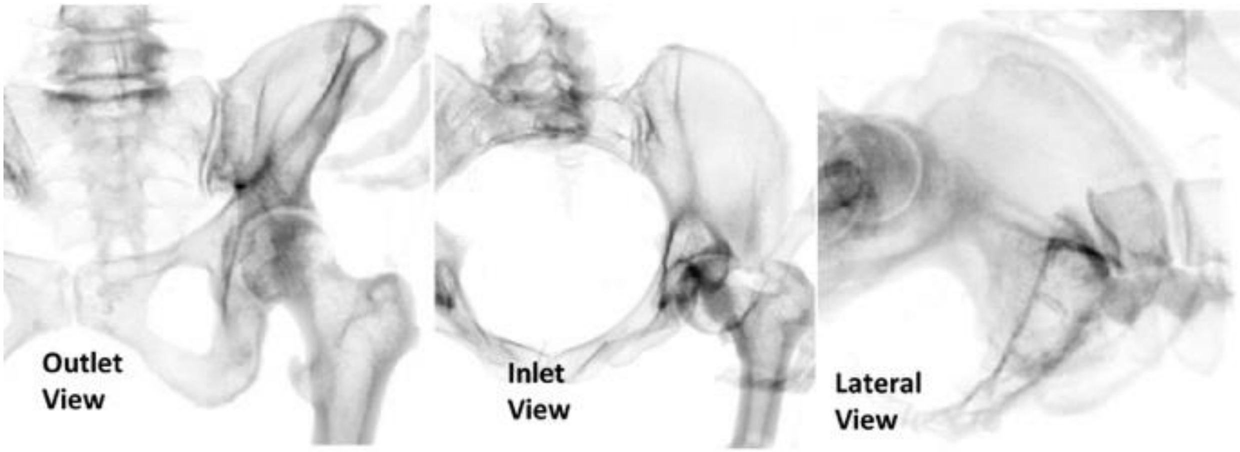


Figure 23: The three views used in the iliosacral procedure are shown here. The inlet and outlet views are taken at oblique angles to a patient, and the lateral view is taken directly along the corridor of the iliosacral joint.

Iliosacral Screw Module Implementation

The new iliosacral simulator was pilot-tested with six, first-year orthopaedic residents. Residents were consented to participate in the study as part of an IRB approved research experiment (IRB ID #201409755). The residents were each presented with a task designed to train them to navigate their wire in the outlet view, inlet view, and in a combined 3D setting. The task involved a series of target bubble pairs overlaid on rendered fluoroscopic images (Figure 24). Residents held a surgical wire on the hemi-pelvis and oriented it so that a line projected along the wire intersected with both targets inside the bone. The task was intended to simulate the step in the surgical procedure when the surgeon orients the wire before drilling. For each series of target bubbles, the resident could try up to 5 wire adjustments to properly position their wire. If the wire was correctly positioned, a new pair of targets were presented, this time with a smaller bubble diameter. In the event that the resident failed after 5 tries, the new targets had a larger diameter. The starting target size had a diameter of 5 mm (0.2 inches). Residents were awarded 5 points for each accurately placed wire. Bonus points were awarded if multiple targets were hit in a row, and for hitting a target with a diameter less than the original target size.

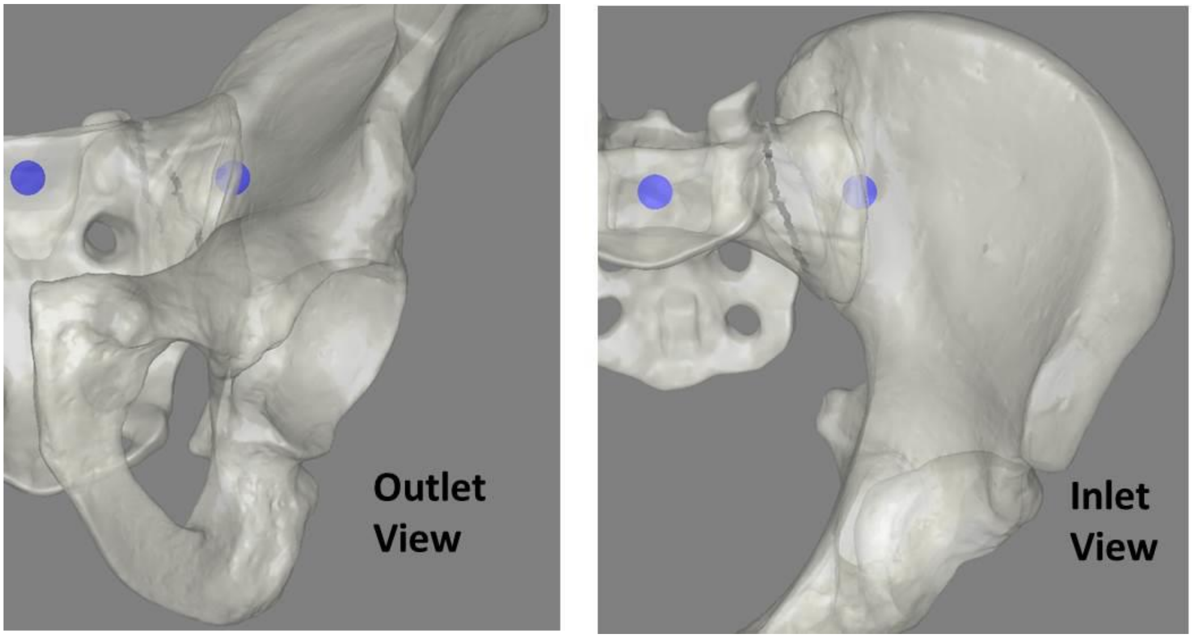


Figure 24: Starting views in the bubble game shown here. In the game cartoon depictions of the pelvic anatomy were used to make the targets more clear to the trainees. The blue circles act as targets for the residents to pierce with their wire during the training exercise.

As the residents used the simulator, data were also collected on how the simulator itself was performing. If an erroneous wire position was calculated or there was a problem in the simulator algorithm that resulted in no wire being shown to the resident, the raw image that was used for that requested fluoroscopic shot was saved to a repository. These images were later examined to understand if the errors were from a mechanical issue, like the simulator focus or working volume, an algorithmic issue, or a user error, such as having a hand obstructing the wire.

Results

When quantifying the simulator accuracy, the two main metrics used were: wire tip accuracy and wire trajectory accuracy. In comparing calculated tip placements to known tip placements, the wire tip accuracy varied between 0.25 mm (0.01 inches) and 4.85 mm (0.19 inches) of error. On average the wire tip error was 1.7 mm (0.067 inches). The wire angle error varied between 0.04° and 4.3°. The average angle error was 1.31°. Across all tests, the average time to compute the pose was 1.05 seconds, well within the range of a typical fluoroscopic image acquisition in surgery. Table 2 shows the average error measurements of the 29 reference positions for the tip location and trajectory measurements across the different testing locations.

Table 2: The simulator accuracy is shown here. For each location, the simulator computed the wire tip position and trajectory in bone for all 29 laser scanned reference points. The average errors values for each of the 29 reference points at each new testing location is shown below.

Location	Average Tip Error (mm)	Average Angle Error (°)	Average Computation Time (s)
Desktop	1.53	1.15	1.12
Dry Lab	1.59	1.31	1.03
Conference Room	1.61	1.12	1.01
Wet Lab	1.97	1.18	1.08
Library	1.81	1.31	0.99
Average	**1.70**	**1.31**	**1.05**

The system accuracy variation across the working volume of the simulator is another factor in consideration. For each wire reference position in bone, the accuracy was queried to examine any trends that might exist. Most the wire positions had an accuracy in tip placement of 1 to 2 mm. The largest error was in a wire position near the edge of the working volume of the simulator, with an average error value greater than 3mm. Six wire positions had average errors of less than 1 mm which all trended towards the center of the working volume.

Table 3: Bubble Target Practice Scores

Resident	Inlet View	Outlet View	Combined 3D Task
1	20	25	11
2	19	18	5
3	39	52	24
4	31	45	15
5	30	36	17
6	61	32	17

The performance of the residents participating in the bubble target practice exercise on the iliosacral module can be seen in Table 3. The average score on the inlet task was 33.3, meaning on average resident could hit the target at least 6 times in the inlet view during the training (5 points were awarded for each time a target was hit in that 2D plane). On the outlet view, the average score was 34.6, again suggesting that residents were able to hit the target at least 6 times on the outlet view during training. When looking at the combined 3D task however, the average score was 14.8, suggesting residents were only able to hit the 3D target 2 to 3 times during training. This suggests that correctly positioning the wire in a 3D position is a more challenging task than correctly positioning the wire in a single plane, which makes intuitive sense. What is not clear at first, however, is which plane is more difficult to maneuver in or contributes more heavily to the ability to correctly drive the wire in a 3D position. The correlation between each resident's performance in the inlet view and their performance on the 3D task had an R squared value of 0.36. Likewise, the correlation between each resident's performance in the outlet view and their performance on the 3D task had an R squared value of 0.79 (Figure 9).

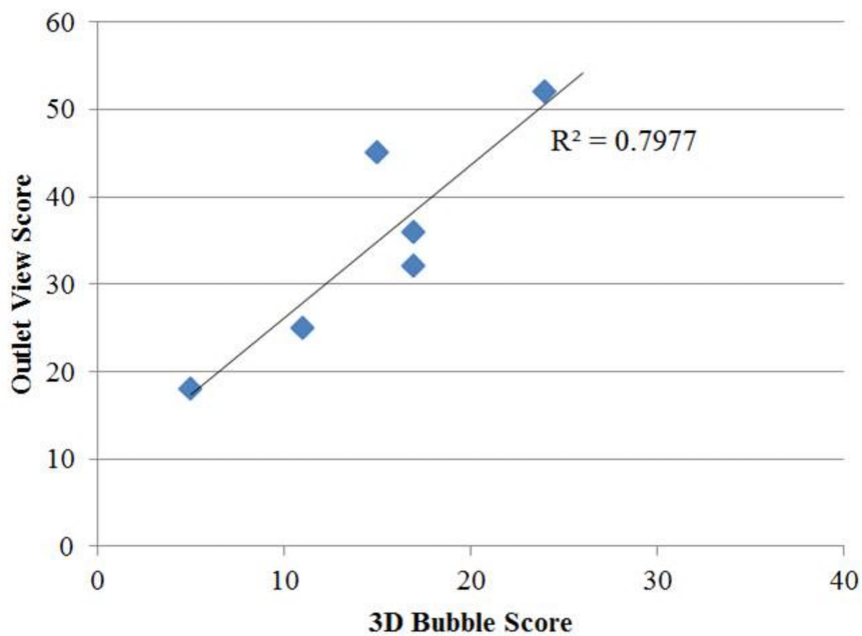


Figure 25: This graph illustrates the strong relationship between a resident's performance on the target practice in the outlet view with the 3D target practice. Each residents' performance on the outlet view bubble task is graphed on the y axis and the corresponding score on the 3D bubble task is shown on the x-axis. A strong correlation of R^2 = 0.79 can be seen between these two metrics.

Over the course of training with six residents on the iliosacral target task, a total of 948 images were taken. Of these images, 74 (7.8%) triggered a failure in the system. Of these failures, 35 were due to a resident's hand obstructing the view of the wire to the camera, 15 resulted from the wire being outside of the working volume of the simulator, 4 resulted from a poorly focused image, and 20 resulted from a flaw in the underlying simulator algorithm that was subsequently fixed.

Discussion

Wire navigation is a central skill in the field of orthopaedics that has many applications. Developing a training tool that can encompass a variety of these applications has obvious benefits to the community of orthopaedic surgeons and ultimately the patients receiving the operation. Given the positive feedback our group has received when training residents with the wire navigation simulator for applications in the hip, we endeavored to explore the simulator platform's ability to incorporate other wire navigation tasks. In this study, the simulator accuracy, working volume definition, and tactile feedback were the main criteria used to create a design space around which new simulator applications could be developed. We found that the simulator can accurately position the wire within an average of 1.7 mm (0.067 inches), 1.31°, and display the resulting image to a resident within 1.05 seconds. As the skill level of surgeon using the simulator increases, or the wire navigation application changes, this level of accuracy may be more or less sufficient. In the application of an iliosacral screw, residents are attempting to place the wire within a corridor of approximately 14 mm (0.55 inches) in diameter. Although no qualitative feedback was formally collected, it appeared that the accuracy of the images presented to the residents provided sufficient realism that the task was accepted as valid. In other applications, however, greater precision may be needed to convey the same level of realism. For instance, fractures in the hand can often be treated with a wire navigation technique. In these instances, the corridor a surgeon is trying to place the guide wire in may only be slightly larger than the simulator can accurately place the wire. In these instances, design parameters of the simulator, such as the resolution of the optical system, may need to be altered to provide the necessary precision.

Apart from characterizing the accuracy of the simulator, this study also examined how the accuracy varied across the working volume. Figure 8 illustrates that although there is not a large bias or pattern observed, the accuracy of the simulator decreases slightly as one moves to the edges of the optical viewing frame. The wire locations that had an average precision of less than 1 mm (0.039 inches) are generally located at the center of the working volume. Only two testing locations placed the wires with accuracy error greater than 2 mm (0.079 inches) and both of these locations are near the far edge of the working volume. This result is somewhat expected, as the center of the working volume would be where the camera is most in focus, and at the edges of the working volume the camera would begin to lose focus. This may be an important design element to consider in future application developments as varying levels of precision and working volume are required.

Implementing the iliosacral targeting task was an important step in testing the design of the new module. Although the theoretical model satisfied the design criteria, there are always unexpected issues that only present themselves when real users engage with a device. In this instance, the analysis of the flagged system failures during use illustrated two unexpected findings. First, most errors came from residents placing their hand on the wire while requesting an image. This type of behavior illustrates the disconnect that can sometimes occur between behavior in the operating room and in a simulated environment. In the operating room, it is common for surgeons to support the guide wire with their hand to ensure that they are aiming in the trajectory that they intend to drill the wire. Given a simulator should never stifle behavior that relates to operating room, this presents a design challenge that may need to be tackled in future iterations. However, placing a hand in the view does come at the expense of additional radiation exposure to the surgeon. The second behavior that was unexpected was wires being placed outside of the working volume of the simulator, accounting for 15 of the 74 system failures. Although the theoretical working volume of the simulator should have been more than sufficient in comparison to the working volume of the procedure defined during the literature review, it may be that learners have behaviors that take them outside these boundaries. Additionally, the soft tissues provided around the artificial bone gave little indication as to the limits of the optical system. It could be that if some sort boundary line was drawn on the soft tissue, this would prevent these types of errors from occurring. However, that may be used as information to guide the residents in their placement of the wire that is not available in the operating room.

The results of the iliosacral targeting task showed a strong link between a resident's ability to direct a guide wire on an outlet view and their ability to direct a guide wire in a 3D targeting task (Figure 9). This suggests that the outlet view may be more difficult to interpret or maneuver in, and thus practice in this plane holds the key to resident improvement on the iliosacral screw task. Although both views are critical to the success of the procedure, this result could help reduce complications during this type of procedure, as the outlet view is the main view that allows the surgeons to see where their wire is traveling relative to nerve root paths. Perhaps by placing more emphasis on learning to navigate in this plane, surgeons will have better place implants across the sacroiliac joint and will lead to fewer complications for patients.

A potential limitation of this study would be the small sample size of residents who have used the iliosacral module on the wire navigation simulator. As more residents are exposed to this task, some of the patterns of behavior seen in this study may change or become less evident. Additionally, all the residents who interacted with the module were first year novice surgeons. Given that iliosacral screw placement is a relatively advanced task, it would make sense to target a more advanced level of surgeon. In this instance, the more experienced surgeons may not so easily accept the accuracy or resolution of the system and may find it does not adequately reflect their experience in the operating room.

This study has successfully demonstrated the ability to adapt an existing wire navigation platform originally developed for a hip application to new applications such as an iliosacral screw placement. Through a careful examination of the existing capabilities of the simulator and needed capabilities of other procedures, the simulator could be modified to present the new surgical application for iliosacral screw placement. Further, testing with orthopaedic residents has demonstrated the simulator reliability in the new application and begun to illustrate a new relationship between navigating in a single plane and a 3D space that may exist. Further testing with additional residents will be needed to validate the iliosacral module as an effective training tool for orthopaedic surgeons.

CHAPTER 4: DECISION MAKING IN WIRE NAVIGATION

Introduction

The research thus far has shown the benefits of training with a wire navigation simulator. We have demonstrated that by practicing on the wire navigation simulator, we can improve a resident's ability to minimize their tip apex distance in a mock operating room environment. However, recent work has brought into question whether the tip apex distance captures the entirety of a surgical performance. There is no doubt that there is a connection between a low tip apex distance and less likelihood of implant failure as shown by Baumgartner et al [60], however this metric only examines the final implant position and does not take into account the path that a resident took to achieve that position. **Figure 26** displays wire positions from trials recorded on the simulator. Both residents achieved a tip apex distance of 19 mm, however, the resident on the left took a drastically different path through bone to achieve that result.

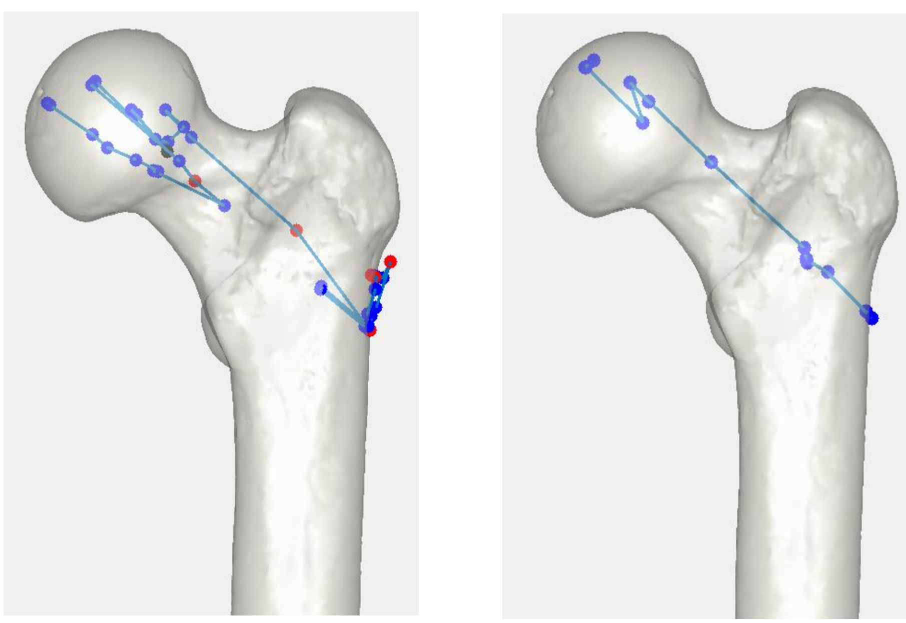


Figure 26: The above graphic shows two very different paths taken in placing wires with identical tip apex distance (TAD).

From this graphic, it is clear that there is an iterative process that residents/surgeons take to adjust their wire position, correct for drilling that would result in a poor TAD, and ultimately place their wire in its final position. This iterative process is one that likely contains a lot of information about the skills of the surgeons and the decisions they are making along the way to achieve a satisfactory result. In medical education, examining decision making skills has been used as a way of assessing the competency of physicians [87, 88]. These studies often referred to decision making in a high level context, such as deciding whether or not to do a surgery or whether or not to use a bag to extract a gallbladder [89]. In this context, we will define decision making as the intentional change in guide wire position between two consecutive fluoroscopic images taken during surgery. That said, similar principals may still apply in this context. Decision making is a promising area to examine as there has been a negative correlation demonstrated between decision making errors and technical-errors in junior residents [90].

Similar to understanding how learners progress through phases of performance, we must develop an understanding of how surgeons normally make decisions in the operating room so that we can understand how to improve and aid their decision-making capacity. Cristancho et al. developed a model of intraoperative decision making that involved three main stages; assessing the situation, the reconciliation cycle, and implementing the planned course of action [89]. Assessing the situation is just as it sounds, the surgeon begins by evaluating the information at hand and comparing it to an existing planned course of action. Following this is an iterative process that involves gaining information, weighting that information, and projecting future steps. And lastly the surgeon implements their plan of action. If we apply this model to wire navigation, a surgeon may plan for a particular wire path through bone, drill the wire into bone, acquire a fluoroscopic image to see how the wire position fits with their original plan, and then decide to continue forward or adjust their trajectory. The last step described here is potentially the crucial step as it is the technical component of the procedure. A skilled surgeon may be able to flawlessly execute his or her plan, making few errors along the way. However, a novice may not yet have the technical skill to implement his or her desired plan in a smooth and seamless fashion. The novice may make more technical errors along the way, resulting in continued assessments and adjustments of their plan until they finally achieve the wire position they cognitively understood to be correct. This type of data and analysis can only be found by examining the intra-operative fluoroscopic images acquired during a surgery. If acquired and analyzed, we may be able to better assess skilled wire navigation performance versus poor wire navigation performance. Therefore, the objective of this study was to analyze fluoroscopic image data taken from our simulator environment, mock operating room environment, and operating room to create a model of decision making and assess its ability to quantify wire navigation performance.

Methods

Data Collection

Data to assess wire navigation decision making was gathered from three environments; the wire navigation simulator, the mock operating room, and the operating room. When residents use the wire navigation simulator, each time they request an image the wire tip location and wire vector are saved to a folder containing the residents data for a given trial. This allows for a retrospective analysis to be performed examining the adjustments residents made when placing their guide wire. Over the course of training with residents from several institutions (the University of Iowa, Mayo Clinic, the University of Nebraska, the University of Minnesota), a total of 176 wire navigation trials were collected. These trials varied from training-based performances where feedback was given to the resident as they drilled into bone to assessment-based trials where the resident was relying only on the fluoroscopic images to place their wire.

Data was also sampled from residents' performance in the mock OR environment that was collected during the transfer validity study. The mock OR was used as a surrogate OR environment for testing first year residents on their hip wire navigation skills. Residents were able to drill a guide wire into a radiolucent sawbones femur and see their wire position through real fluoroscopic images provided by a C-arm. For a subset of cases, all fluoroscopic images were saved when the residents were placing their guide wire as part of the exercise to place a center-center wire as they would for a compression hip screw in the operating room. These images were then analyzed using a DICOM viewing software, Osirix, to locate the wire tip, femur apex, and wire trajectory in bone. The apex of the femur was located using a standard method validated by Johnson et al [59]. This data, similar to the data collected from the wire navigation simulator, was taken through a retrospective analysis to understand the resident decision making during their wire navigation in the mock OR.

Data was also collected from hip wire navigation performances in the actual operating room. From August of 2014 to February of 2016 and again from August of 2018 to February of 2019 residents and faculty at the University of Iowa were instructed to save all fluoroscopic images that were acquired during cases that involved placing a center-center guide wire for the treatment of intertrochanteric hip fractures. A total of 21 cases were successfully collected during these time periods. In the same process used with the data from the mock OR, the images were analyzed in Osirix to identify the apex of the femoral head, wire tip, and wire trajectory.

Categorizing Decision Making Errors

From these three environments of hip wire navigation, images were analyzed to understand the types of behaviors that surgeons operate under when placing a guide wire. A schematic diagram below shows how the logical flow of a wire should proceed when drilling towards a target, in this case the apex of the femoral head, and also the types of errors that surgeons can potentially make along the way.

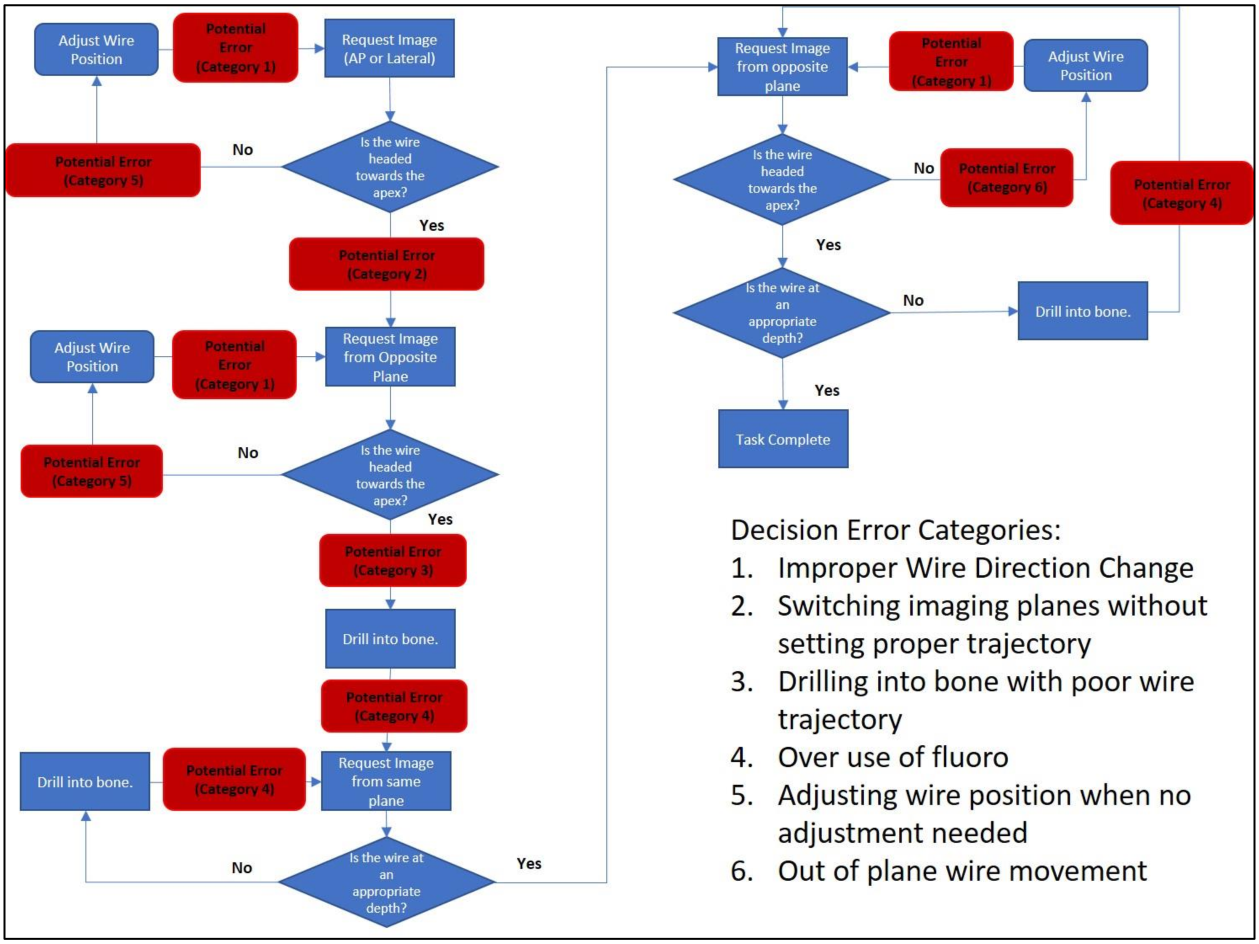


Figure 27: This diagram shows a schematic of logic used to lay out how a wire should be driven in bone to place a center-center guide wire. At each decision point there is a potential for error from the surgeon. These errors have been categorized into 6 potential errors as shown here.

There are six main error types that are proposed from this analysis. The first type of error occurs when a surgeon identifies that their wire is not headed toward their desired target.

However, rather than making an adjustment that would move towards the intended target, the surgeon makes an adjustment that directs the wire further away from the intended target. An example of this can be seen in **Figure 28** where a resident clearly makes an adjustment that worsens their wire trajectory. This type of behavior is indicative that the resident is still learning how to adjust their hand and drill position relative to the patient to achieve the desired goal of placing a center-center guide wire. The number of degrees that a resident worsens their wire position can also be quantified as a way of assessing how severe the error was. More experienced surgeons may make similar mistakes throughout a procedure, but it is likely that the magnitude of those mistakes decreases with experience.

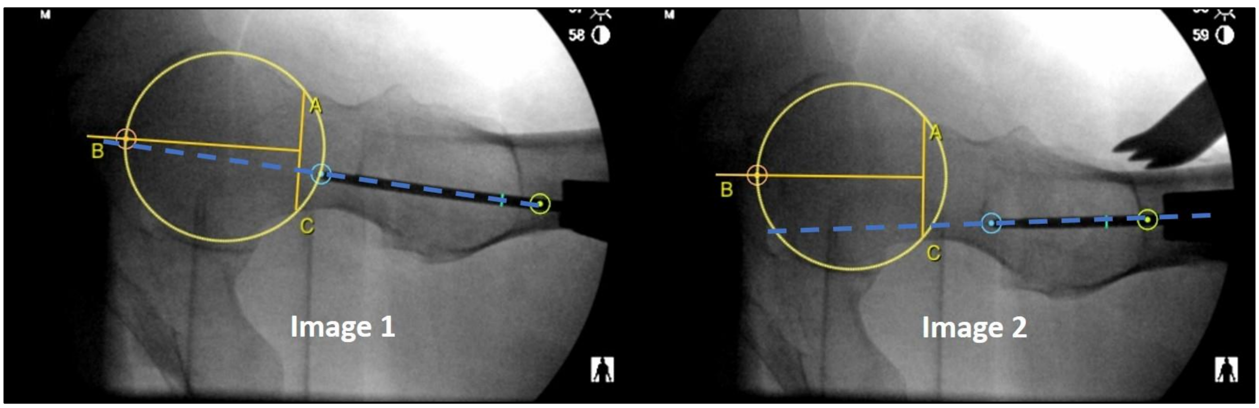


Figure 28: A series of images taken from an OR case can be seen here. In image 1 the resident wire is headed towards the apex target. In image 2, the resident has tried to improve the wire position but instead has adjusted the wire so it points away from the apex.

The second type of error occurs when a resident switches between asking for an AP image to a lateral image (or vice versa) without first properly setting the wire trajectory. An example of this can be seen in **Figure 29**. This may occur either do to an error in judgement on the part of the resident in being able to assess when his or her wire is pointed in the direction of the target. The risk of switching between imaging planes without first properly setting the wire in one plane is that the surgeon will likely have to come back and reset the wire trajectory in the first imaging plane. This would likely result in added images and radiation to the patient that are unnecessary.

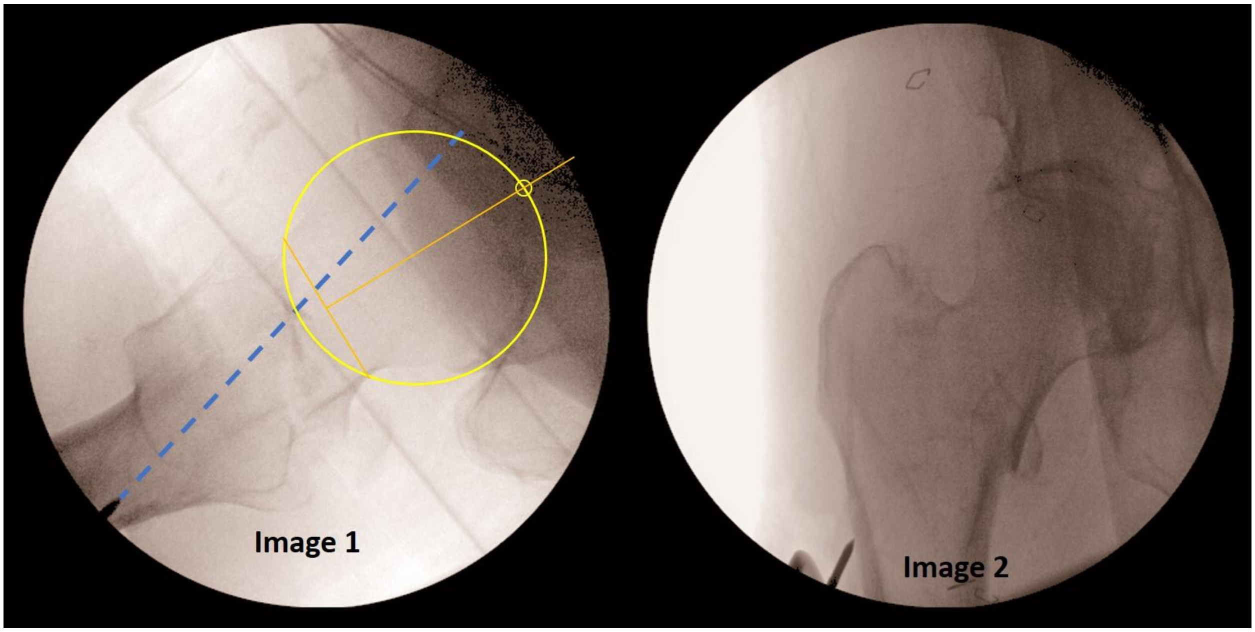


Figure 29: An example of when a resident improperly switched between asking for a lateral image to asking for an AP image is shown here. In the first image, the wire trajectory is clearly not set on target before they decided to switch to the AP image.

The third type of error occurs when a resident drills into bone with an incorrect trajectory. It is often common for surgeons to drill into bone to see how they are advancing toward their target. It is perfectly acceptable to drill in a small amount to assess the wire trajectory, however at a given point the surgeon should realize that they are going on a poor trajectory and adjust the wire angle. For the decision-making analysis, a wire depth of 1.5 centimeters into bone was chosen as a threshold for depth at which the trajectory would be assessed. The depth was calculated based on converting pixel coordinates to millimeters based on the known width of the wire. **Figure 30** shows an example of a resident drilling nearly halfway into bone before they realized that they were on an improper trajectory and eventually had to reverse the wire and redirect their angle of approach. As residents gain experience, they may be able to better assess the projected trajectory that as wire is headed to prevent these types of errors and make corrections earlier in the process, limiting the number of paths taken in bone in the amount of radiation exposed to the patient.

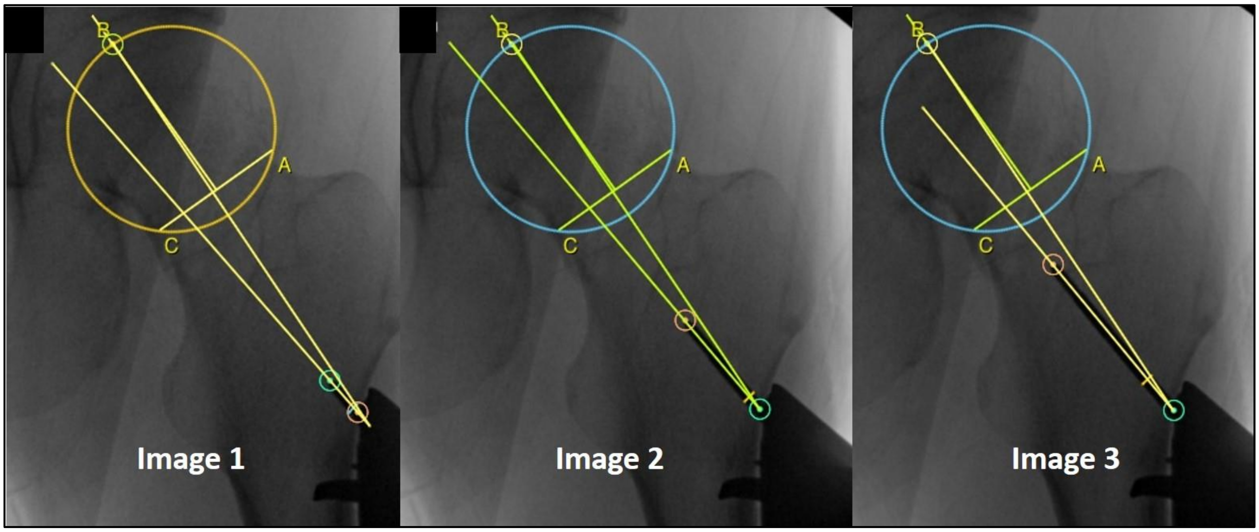


Figure 30: A series of three images here illustrates how sometimes residents drill into bone with an improper trajectory that ends up having to be readjusted later in the procedure.

The fourth type of error is labeled as an over use of fluoroscopy. This occurs when a resident takes an excessive number of fluoroscopic images while progressing the wire very little into bone. For instance, a resident may take an image, drill forward a centimeter, take another image, drill forward another centimeter, and so on. When the wire is outside of the bone, it is a good strategy to use fluoroscopic images to see the changes in wire trajectory. However, once a wire is drilled into bone, the trajectory of the wire will not change as the resident drills into bone. Therefore, the resident should be able to advance the wire sufficiently before requesting a new fluoroscopic image. An example of this can be seen in **Figure 31**.

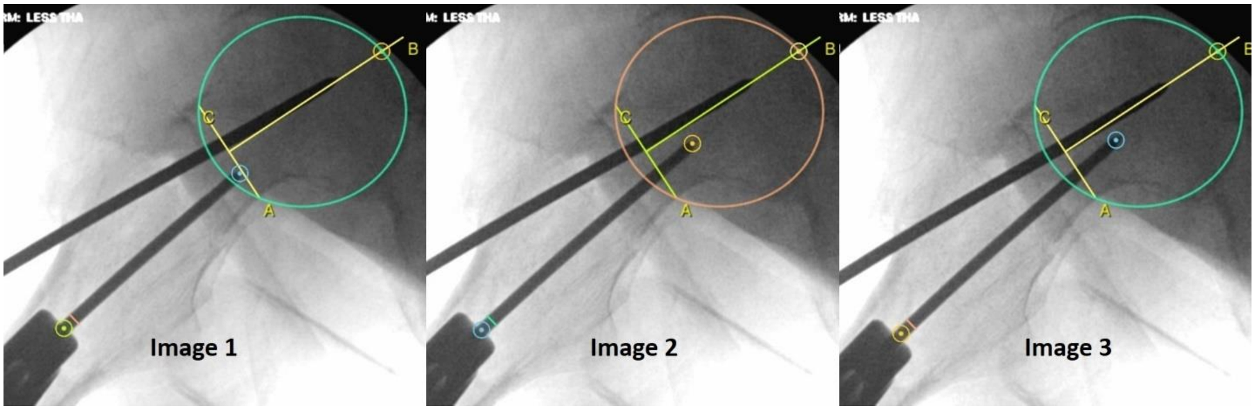


Figure 31: A series of three images here show the wire trajectory staying constant in all three images, yet the wire advances into bone very little. Image 2 and Image 3 show almost no difference in the wire position in bone.

The fifth type of error is making an adjustment to a wire trajectory when the wire trajectory is already satisfactory. Part of good wire navigation is knowing when the wire position is good enough. Making too many adjustments exposes the patient to further radiation, increases the number of drill paths taken in bone, and adds time to the surgery. In the decision-making analysis, a wire trajectory that is within 3.5 degrees of the ideal wire path is considered satisfactory as this would result in a TAD less than 25mm, the threshold set by Baumgartner to be mechanically safe from screw cut-out. If a resident is within this range on either a series of AP or lateral images, they should not be making corrective adjustments. **Figure 32** shows a series of 6 images where a resident attempts to improve on a satisfactory wire position, only to end up with the same wire position 6 images later.

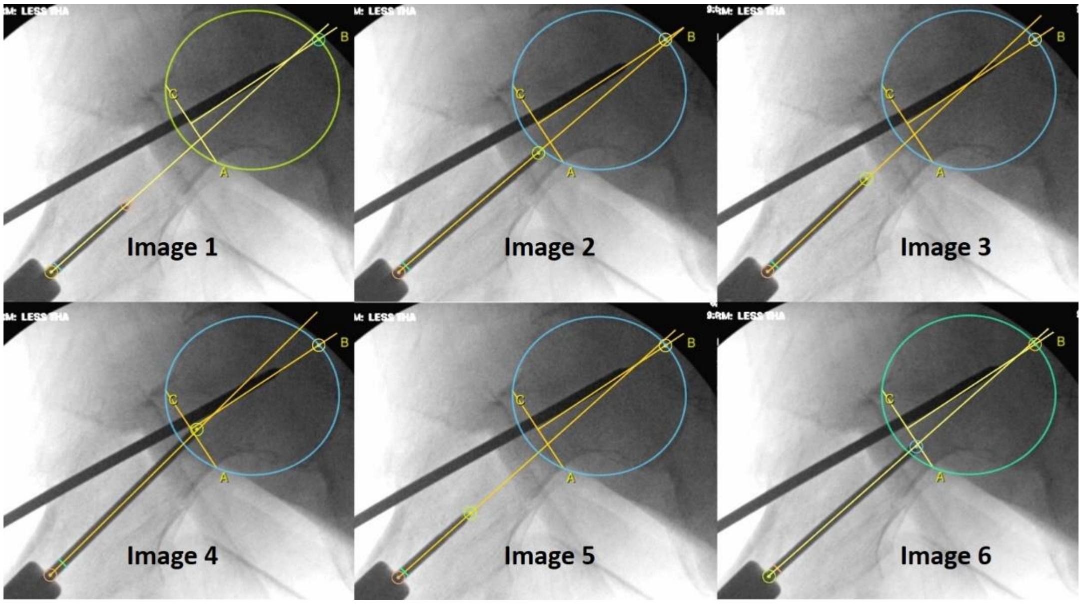


Figure 32: A series of 6 images here show a resident over correcting the wire position. In image 1 and 2 the wire trajectory is pointing in the direction of the apex of the femoral head. The resident reverses the wire in image 3 to achieve a better trajectory but sees in image 4 that this is a worse trajectory. At the end of image 6 the resident has returned to the original trajectory.

This sixth and final type of error is categorized as out of plane wire movement. This occurs when a resident is attempting to make an adjustment to correct a wire trajectory in one imaging plane but inadvertently also changes the wire trajectory in the opposite imaging plane. Some surgeons may argue that they can adjust the wire in both directions at once. For this reason, this type of error is only graded when the out of plane movement worsens the trajectory in the opposite plane. If a surgeon is able to move the wire in both imaging planes at once and improve the wire trajectory in both planes, this is not deemed an error. On the simulator, this can easily be measured given the 3D position of the wire is known for every image acquired. This however is not the case in the operating room. To account for out of plane movement in the OR, when a resident switches between imaging planes, say going from the AP to lateral image, and then returns to the original plane, say going back to the AP image, we can compare the wire angle before and after switching imaging planes. As an example, **Figure 33** illustrates how a resident's wire trajectory can worsen after making adjustments in the opposite plane. In this instance, the residents wire trajectory was on target in the first image, however after making a few adjustments when looking at lateral images, they return to the AP image to see they are no longer on target. Being able to make wire adjustments in a single plane is a difficult skill, but one that is likely indicative of a competent surgeon.

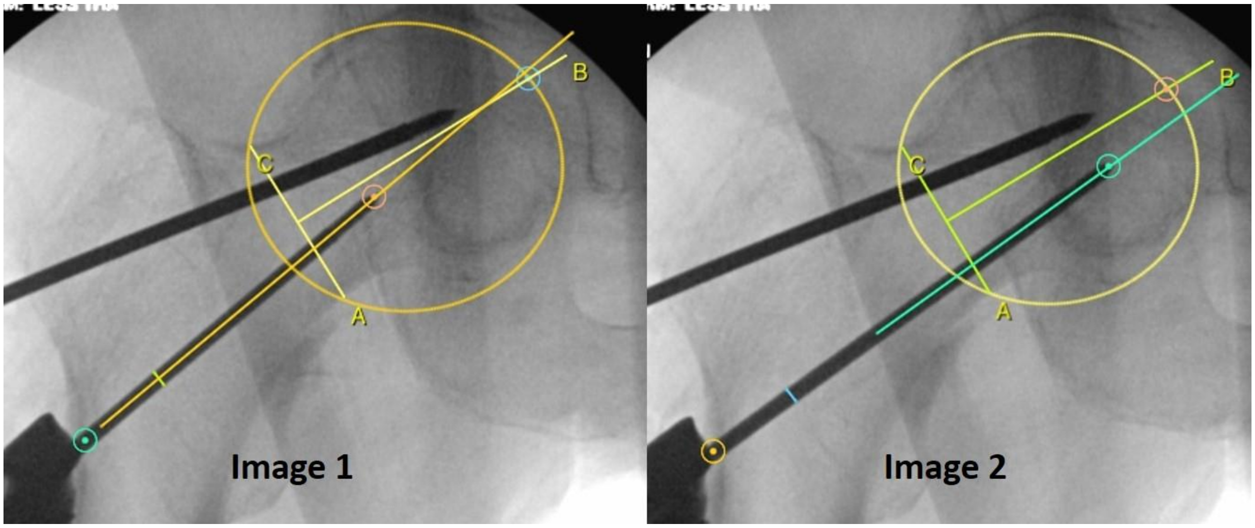


Figure 33: These images were not taken sequentially in the OR. Between image 1 and image 2 a series of lateral images were taken. However, it can be seen that when the resident adjusted their trajectory from the lateral images, they also accidentally worsened the trajectory in the AP image, as seen in image 2.

To analyze wire navigation decision making on the simulator, a MATLAB script was written to read in data from a resident's performance, examine the wire trajectory changes between each requested image, and calculate the number of times any one of the 6 type of decision errors occurred (see APPENDIX A for details). Each data point read into the script was labeled as an AP or lateral image, so it could be calculated which plane the resident was intending to move in compared to which plane the resident actual adjusted the wire in. The apex of the femoral head was consistent throughout all the simulator cases and was used as the point towards which the wire trajectory was compared. A custom script was also written to analyze the 2D data from images taken during cases in the mock OR and actual OR (see APPENDIX A for details). This script was written to assess the same errors of wire navigation decision making that was used in the analysis of the simulator data. Given that all the decision points are made based off 2D information that is presented to a resident (intraoperative fluoroscopic images), the same logic could be applied as in the 3D data of the simulator. In addition to calculating the decision-making errors, the TAD was also calculated and the number of fluoroscopic images taken was recorded. To compare this data with surgeon experience, residents case logs were queried to see how many previous hip wire navigation cases the residents had completed to date. For data from the mock OR, it was assumed that residents had led on no cases in the OR given they were all first-year residents.

Results

Simulator Results

To assess the value of the decision-making analysis, the number of decision errors detected were correlated with other metrics of wire navigation performance on the simulator. A total of 176 wire navigation trials were analyzed. When looking at the relationship between decision-making errors and the number of fluoroscopic images requested, there is a strong positive correlation that can be seen (**Figure 34**). On the other hand, when looking at the relationship between decision-making errors and the TAD, there is almost no correlation between the two metrics (**Figure 35**). This is also true when looking at the average angle error, that is the average angle a resident adjusted the wire in the wrong direction, in relation to the TAD (**Figure 36**).

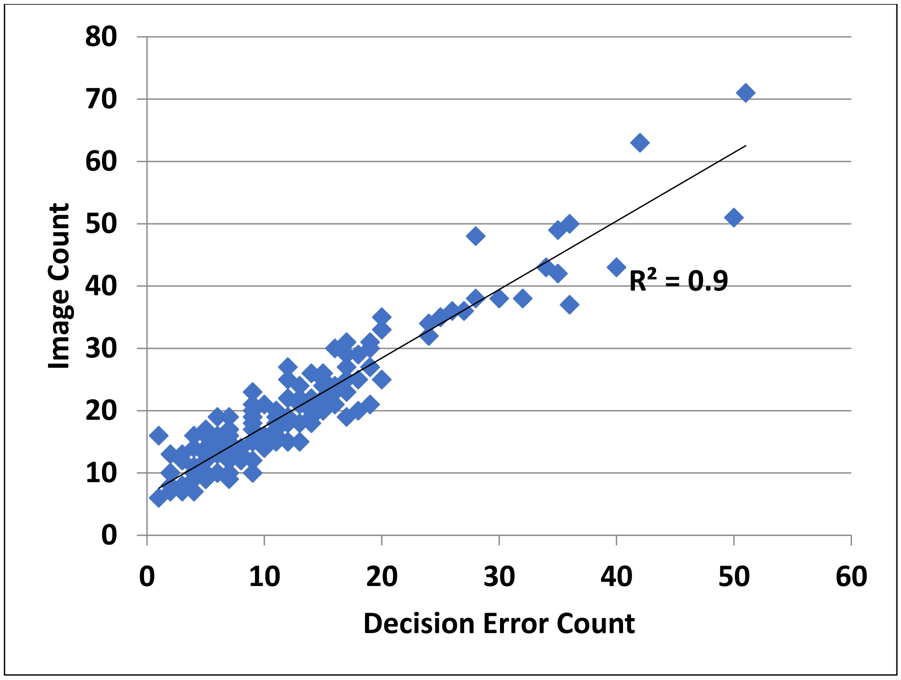


Figure 34: A strong correlation between the number of images taken on the simulator and the number of decision errors is shown here.

Figure 35: A very weak to no correlation can be see between the TAD a resident achieves and the number of decision errors they make.

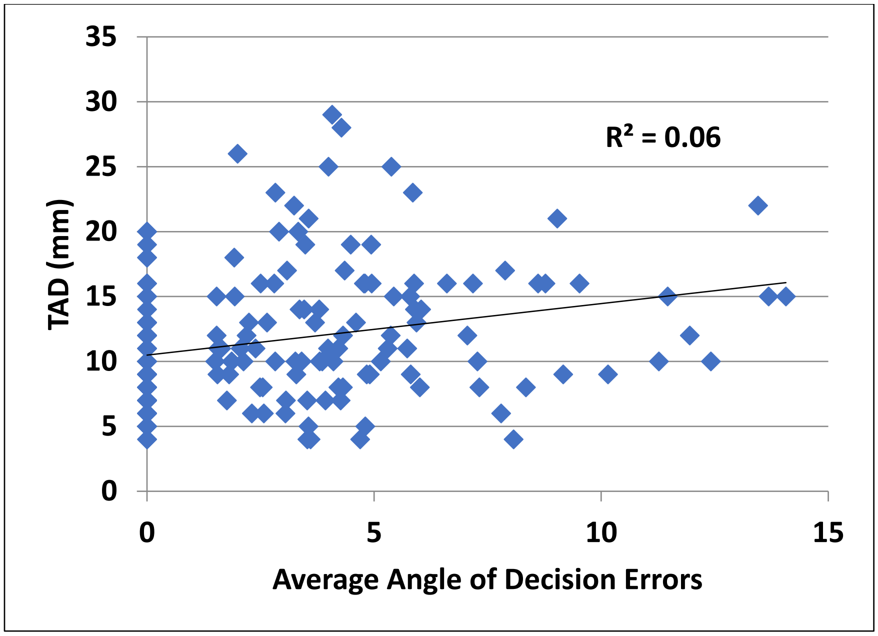


Figure 36: A very weak to no correlation can be see between the TAD a resident achieves and the average angle of their wire adjustment errors.

OR and Mock OR Results

Data from the OR and mock OR were grouped into one dataset. A total of 29 cases were collected and analyzed, 21 cases from the OR and 8 cases from the mock OR. The average resident year was 2.9 years, ranging from first year residents to a faculty member that was categorized as a seventh year resident (the faculty member was two years post residency). The average number of cases logged was 6.13 cases, ranging from 1 case, meaning it was their first case, to 25 cases, (the number of cases completed by the faculty member). In looking at the relationship between the decision making metric and other metrics of wire navigation performance, similar relationships can be observed in the OR that were seen on the simulator. A strong correlation can be seen between the number of decision errors made and the number of fluoroscopy images requested (**Figure 37**). A weak to no correlation exists between the number of decision errors made and a residents TAD (**Figure 38**). And a weak to no correlation exists between the average angle of errors made and a residents TAD (**Figure 39**).

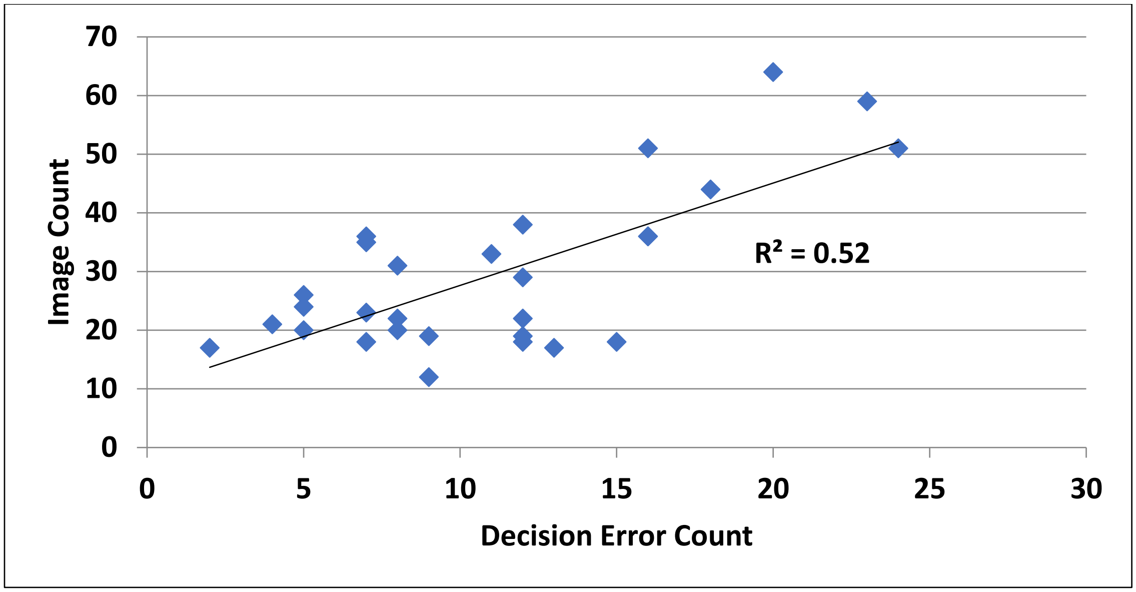


Figure 37: A strong correlation between the number of images taken in the operating room and the number of decision errors is shown here.

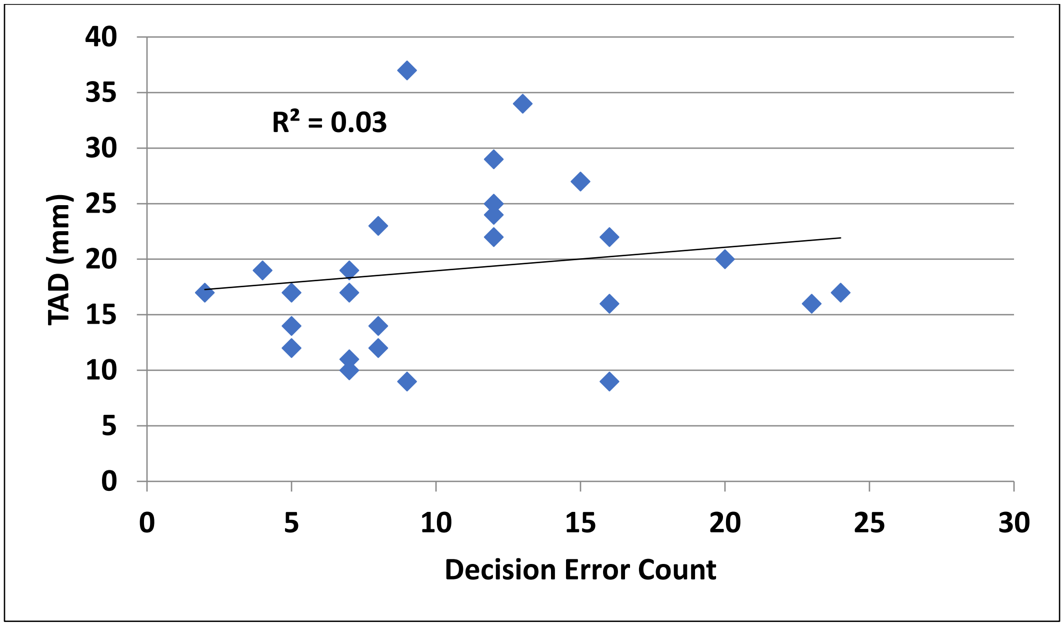


Figure 38: A very weak to no correlation can be see between the TAD a resident achieves and the number of decision errors they make in the operating room.

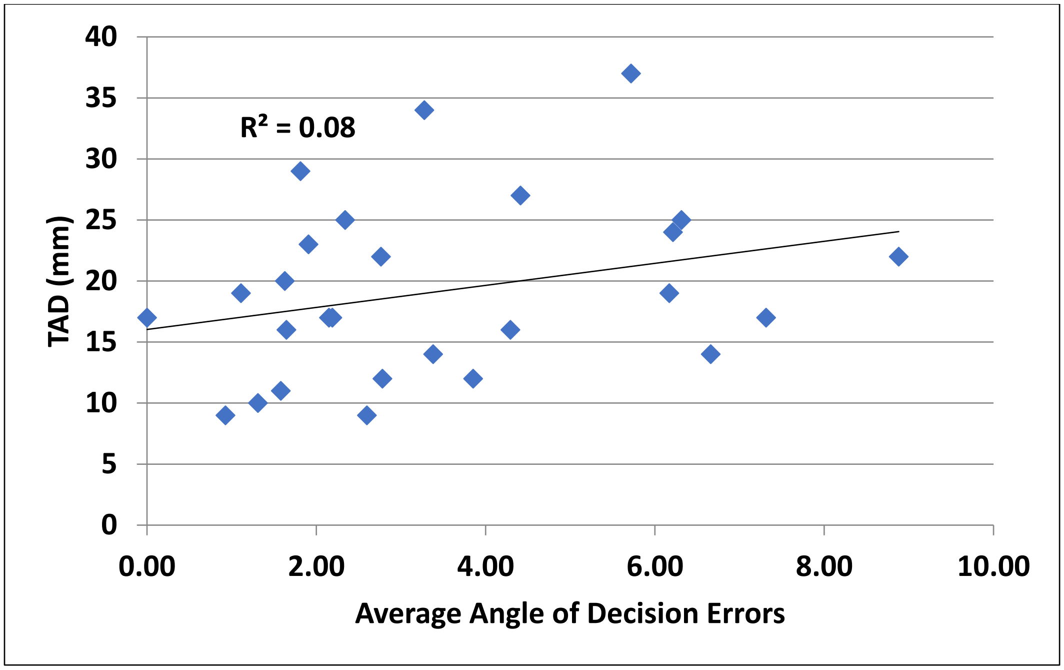


Figure 39: A very weak to no correlation can be see between the TAD a resident achieves and the average angle of their wire adjustment errors in the operating room.

The number of decision errors, the average angle of decision errors, and the TAD metric were all analyzed to see their relationship with the number of cases a resident has logged in the OR. The decision error count had a moderate correlation with the number of hip wire navigation cases completed of R-squared equal to 0.14 when using a logarithmic curve fit to the data (**Figure 40**). This correlation was statistically significant ($p = 0.046$) when running a correlation test using the SAS statistical software. The average angle of decision errors had a stronger correlation with the number of hip wire navigation cases completed of R-squared equal to 0.29 when using a logarithmic curve fit to the data (**Figure 41**). This correlation was statistically significant ($p = 0.003$) when running a correlation test using the SAS statistical software. Lastly, TAD metric had a strong correlation with the number of hip wire navigation cases completed of R-squared equal to 0.62 when using a logarithmic curve fit to the data (**Figure 42**). This correlation was statistically significant ($p < 0.001$) when running a correlation test using the SAS statistical software.

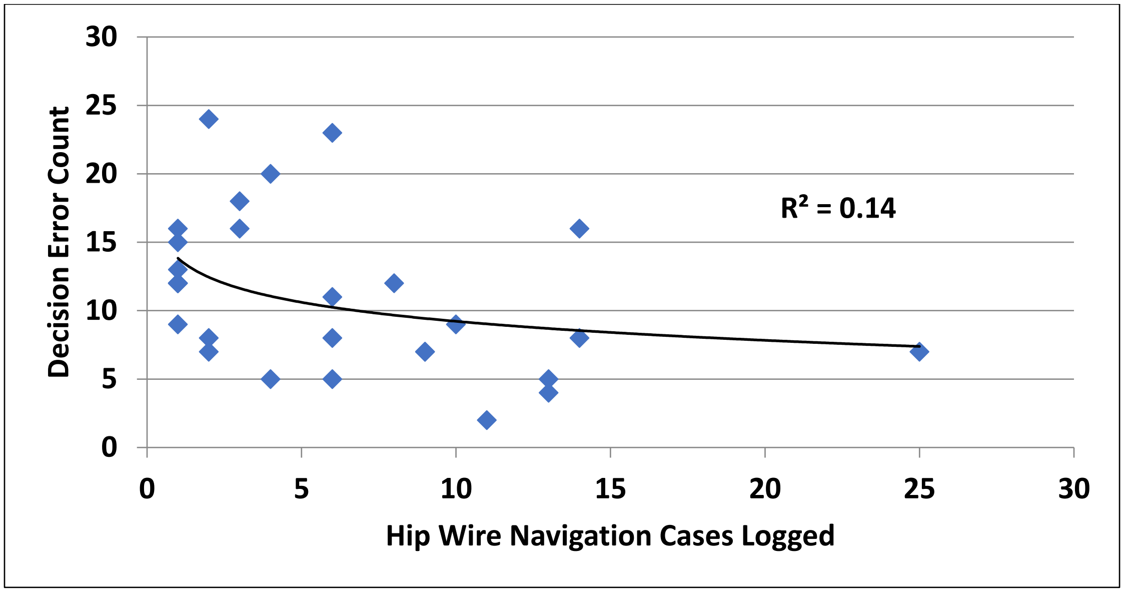


Figure 40: This graph shows the correlation between the number of decision errors made and the number of hip wire navigation cases logged.

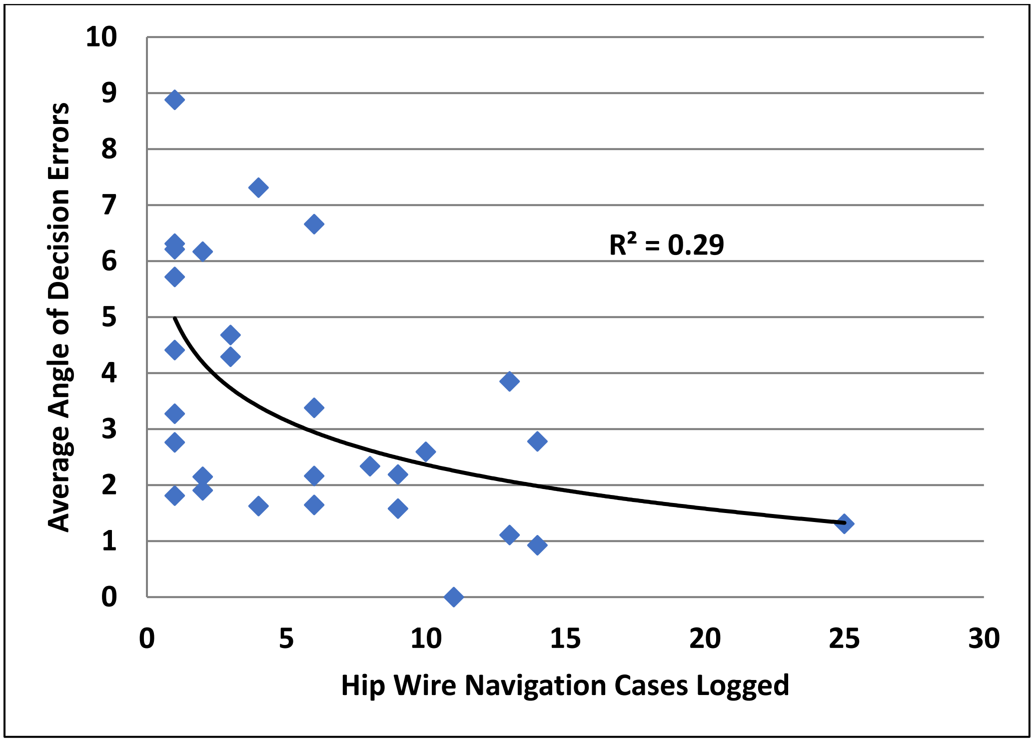


Figure 41: This graph shows the correlation between the average angle of decision errors and the number of hip wire navigation cases logged.

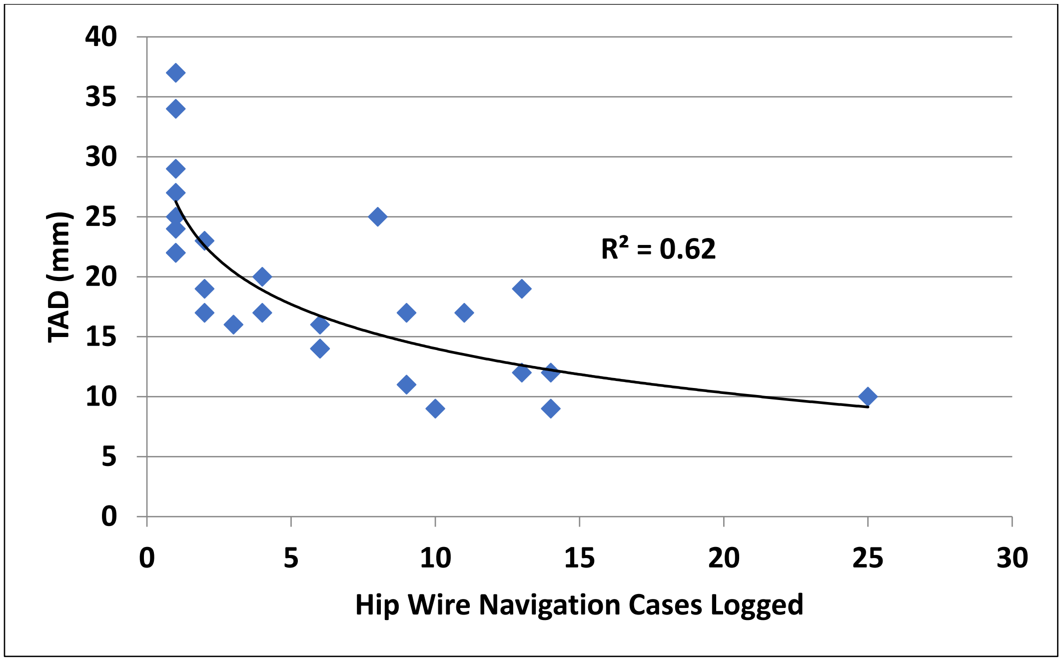


Figure 42: This graph shows the correlation between the TAD a resident achieved and the number of hip wire navigation cases logged.

A composite score was also calculated to summarize the resident's overall performance on the wire navigation task. This score was a combination of the TAD metric, the number of decision errors made, and the average angle of the decision errors. Each metric was first normalized based on the overall population mean and standard deviation before the metrics were summed to create the composite score. In this composite score, a higher score indicates better performance. A score that is equal to zero would indicate that it matches perfectly with the average performance across all subjects. As can be seen in **Figure 43**, a strong correlation exists between the overall composite score and the number of hip wire navigation cases logged by a resident. This correlation was statistically significant ($p < 0.001$) when running a correlation test using the SAS statistical software. A table containing all performance metrics from the operating room and mock OR can be found in APPENDIX B.

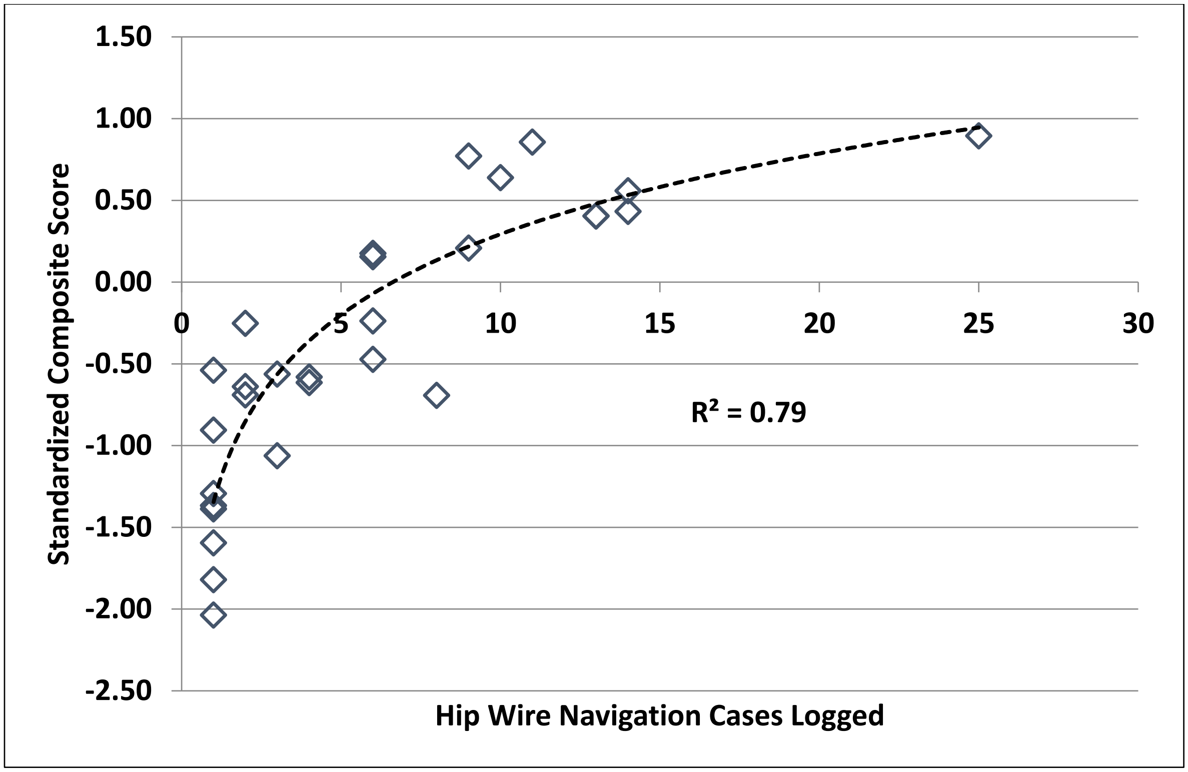


Figure 43: This graph shows how a resident's overall score improves as they complete more cases and gain more experience. The score is a metric that combines the TAD, number of decision errors, and the average angle of decision errors.

Discussion

Assessing surgical performance in the operating room is an important component in moving towards a system of competency-based education. Current assessment metrics like the OSATS score or O-Score use expert ratings of resident performance based on different categories of assessment such as tool use, respect for soft tissue, or surgeon independence [30, 32, 33]. These metrics have been shown to distinguish between levels of novice and expert performance and have been used to validate a variety of simulation tools [31]. However, it has also been shown that these metrics do not necessarily relate to the quality of the surgery as it relates to more mechanical factors than can influence patient outcomes [9]. Additionally, there is concern that some residents feel intimidated when asking for evaluation from different faculty members, so there may be a selection bias in some of this data [33].

Examining intraoperative images taken during a given surgery provides an opportunity to gain unique insight into the decision-making ability and technical abilities of surgeons with data that is already being acquired. It is important to have metrics which can be readily measured in both a simulated environment and in the operating room so that performance on a simulator can be linked to performance in the operating room. In this study, the decision-making analysis was successfully implemented with data from both the simulator environment and the operating room environment. Additionally, the relationships that were observed between the decision making metric and other metrics of wire navigation were found to be similar in both environments. The was a strong relationship between decision error made and the number of images taken on both the simulator and in the operating room. Also, the number of decision errors appears to be independent of a surgeon's tip apex distance on both the simulator and in the operating room. It is promising that these two environments have similar relationships to the decision-making metric because it suggests that the behavior in one setting is linked to the behaviors in the other setting.

Given that the number of decision errors and the average angle of the decision errors appears to be independent of a surgeon's tip apex distance, this led us to believe that combining these metrics into a composite score would result in a greater prediction of surgical skill or surgical experience. Taylor et al. implemented a similar metric when looking at hip wire navigation performance in the operating room and found that combining a surgeons TAD, number of fluoroscopic images, overall time, and amount of intervention required from a supervising surgeon had a correlation of R squared = 0.43 when related to the number of hip wire navigation cases logged by a resident [61]. When combining the TAD, number of decision errors, and average angle of decision errors into a composite metric there was a strong correlation of R squared = 0.79 with relation to the number of hip wire navigation cases logged by a resident.

When breaking down the composite score into its components, the TAD does appear to account for a large amount of the variance in the data. The TAD metric had a correlation of R squared = 0.62 in relation to surgeon cases logged. However, given the decision-making analysis is independent of the TAD metric, it can account for roughly an additional 20% of the variance in the surgical performance. Furthermore, the strong TAD correlation may be weighed by data from the mock OR in which first year residents were placing a guide wire without supervision from an attending surgeon. Of the mock OR data included in this study, four of the eight residents had a TAD greater than 25 mm. It is likely if they were forced to keep working until they had achieved a clinically acceptable TAD that they would have made more decision errors along the way. In any case, these three metrics of wire navigation performance, TAD, number of decision errors, and average angle of decision errors, clearly provide a strong assessment of wire navigation performance.

The decision-making analysis also has the potential to be very beneficial for training and feedback purposes. Karam et al. demonstrated the value feedback from coaching based on video of resident's performance during a fracture reduction task [72]. However, getting video into the operating room for feedback later can be logistically challenging and also has patient anonymity concerns. The series of intraoperative fluoroscopic images taken during a surgery provide a frame by frame summary of a resident's surgical performance, not unlike a video. There is potential value in using these images to debrief with residents after a surgery and look back on what went well and what could have gone better. Without going through the full analysis of quantifying someone's decision making performance, there is a great deal of qualitative information and feedback that could be given through reviewing these images. For instance, an attending surgeon could help a resident understand that they should have switched to an AP or lateral image sooner in the procedure before drilling far into bone which would have made future angle adjustments easier. The examples are endless.

In addition, this analysis has the potential to be applied to a variety of different wire navigation applications. In surgeries like placing a wire for an iliosacral screw a decision-making analysis could be very beneficial at evaluating a surgeon's ability to drive a wire in a particular plane. This is likely a very important and challenging task for the iliosacral screw because of the orientation that the fluoroscopic images are taken from as well as the complexity of the pelvic anatomy. In placing a guide wire for an IM nail, this could also be used to evaluate how a resident is able to find the proper starting point for the wire. **Figure 44** shows a qualitative example of a resident struggling to place their wire in the proper entry point and at the proper angle. Given the large number of wire navigation applications in orthopaedics, this new analysis and metric has the potential to provide great benefit.

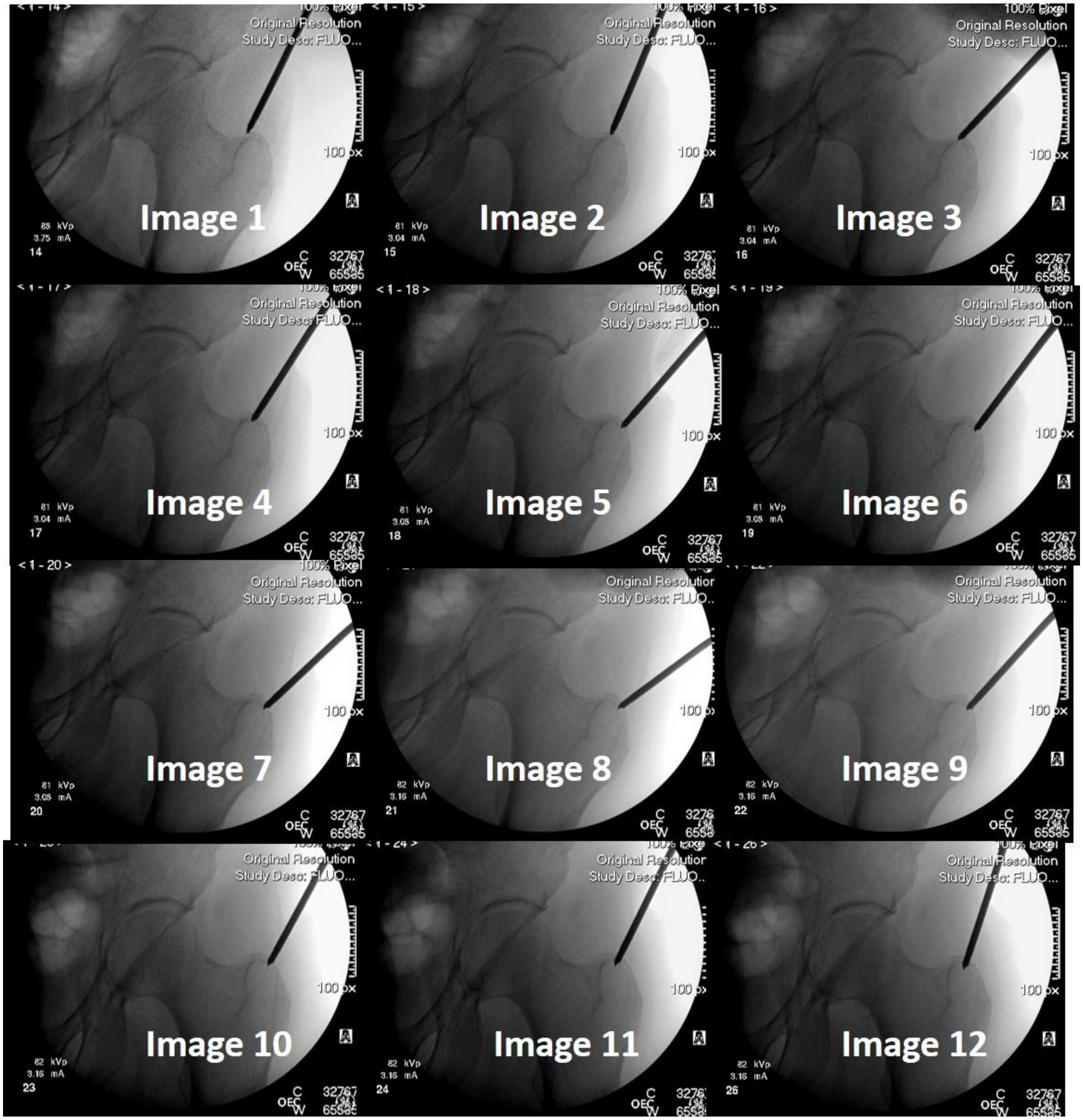


Figure 44: A series of 12 images shown here taken from an OR procedure where a guide wire is placed prior to placing an intramedullary rod. It is clear that the resident is struggling to get their guide wire in the proper direction and decision-making analysis would be able to quantify the amount that the resident is struggling.

CHAPTER 5: EVALUATING SURGICAL PERFORMANCE AT A COMPREHENSIVE FRACTURE COURSE

Introduction

The work thus far has demonstrated that training on the wire navigation simulator can improve a resident's ability to minimize their tip-apex distance in a mock operating room environment. However, we have also seen that a surgeon's tip-apex distance is independent of the wire adjustments and decisions they make during the process of placing a guide wire. This suggests that different training may be required to improve performance on surgical decision-making components than is required to decrease the tip-apex distance metric.

Comprehensive fracture courses offer a unique environment for residents to learn basic fracture fixation skills. Organized by groups such as the Orthopaedic Trauma Association (OTA) or the Arbeitsgemeinschaft für Osteosynthesefragen (AO), these courses are often three-day retreats that consist of a series of modules aimed at teaching different orthopaedic trauma skills such as fracture reduction, placing screw and side plates, or putting in compression hip screws for treating hip fractures. Modules are designed to include a didactic portion that covers case discussions as well as a lab portion that involves teaching residents the basic technical skills with materials and surgical instruments often provided by implant companies. One of the challenges of teaching a skill like wire navigation at a fracture course is the lack of fluoroscopy available to the residents. As a stand in, residents drill guide wires into exposed sawbones, similar to that seen in **Figure 45**. They are usually instructed to poke the guide wire through the other side of bone so that they can see if they were on the right trajectory or not. This is obviously a poor representation of what placing a guide wire would actually be like in the operating room. Therefore, without renting a set of C-arm units for the residents, training on wire navigation at these courses is likely very limited.

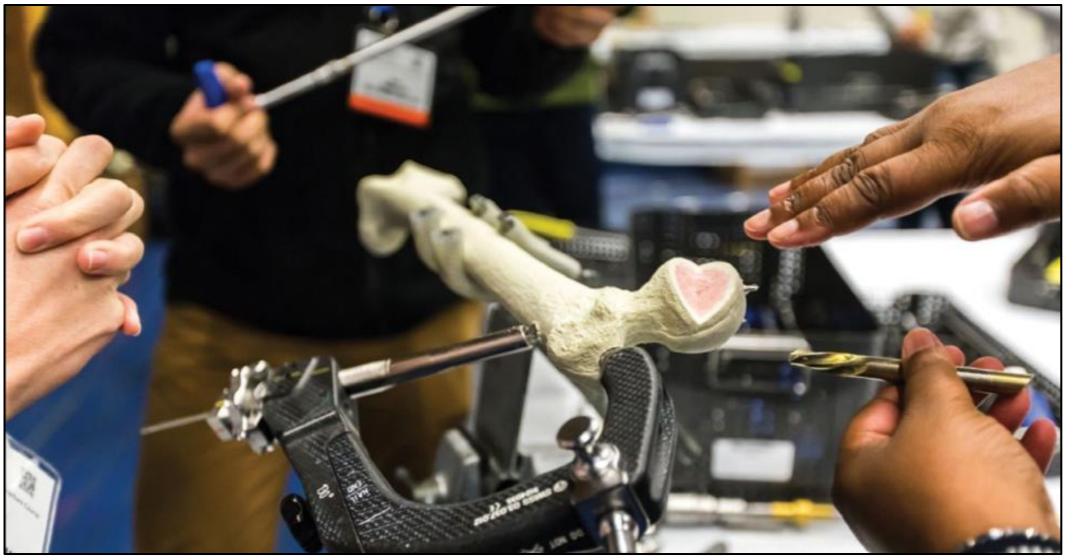


Figure 45: Residents train on naked sawbones at fracture courses. These often can help demonstrate procedures, but likely do not train wire navigation skills given the lack of fluoroscopy.

To bring a higher level of wire navigation training to these fracture courses, our group has developed a partnership with the OTA and has brought a set of hip wire navigation simulators to the Comprehensive Fracture Courses since the spring of 2017. This study is based on the results from the past two years of working with residents at these fracture courses. There are three main objectives that this study aims to achieve. First, this study aims to gather a broad array of resident performance on the wire navigation simulator to better understand the distribution of wire navigation that exists among learners. If a simulation-based test aimed to assess surgical competency, you would first need to understand the different levels of performance to know the difference between competent and not competent performance. The second aim of this study was to explore training based on surgical decision making. Having identified the types of wire navigation decision making errors that residents are prone to making, we believe that providing specific training tasks to residents designed around wire navigation decision making will lead to fewer decision making errors. Lastly, the third aim of this study is to explore the benefit that residents receive from training and learning from others. At the fracture courses, residents are placed into pairs for the lab portion of each module. This study will explore how much benefit residents receive from training with a partner.

Methods

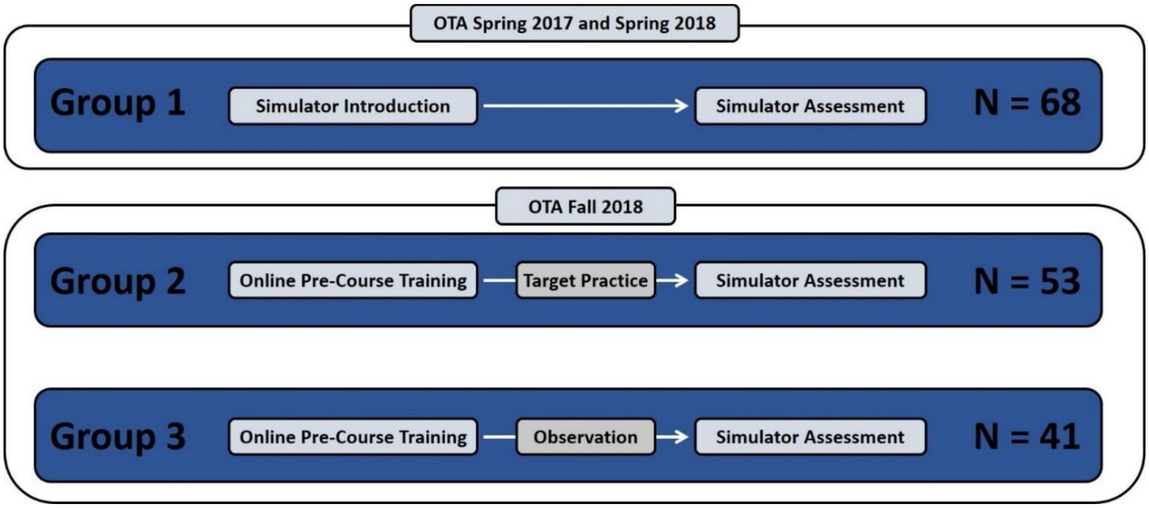


Figure 46: The study design for this experiment. Data gathered from OTA courses between Spring 2017 and Spring 2018 are collected into one group. Data from the fall 2018 course created the training and observing group.

Participants at the fracture courses were placed into one of three groups for this study (**Figure 46**). Group 1 served as a baseline for measuring resident performance without receiving training on the wire navigation simulator. Residents went through a brief introduction on the simulator, allowing them to become familiar with the technology. Residents were then shown an image of an ideal center-center wire position and asked to use the simulator to place a guide wire in the same position. **Figure 47** shows the AP and lateral images that were used during the assessments of all groups in this study. Data for group 1 was collected during the fracture courses in the spring of 2017 and the spring of 2018. A total of 68 resident assessments were collected for group 1.

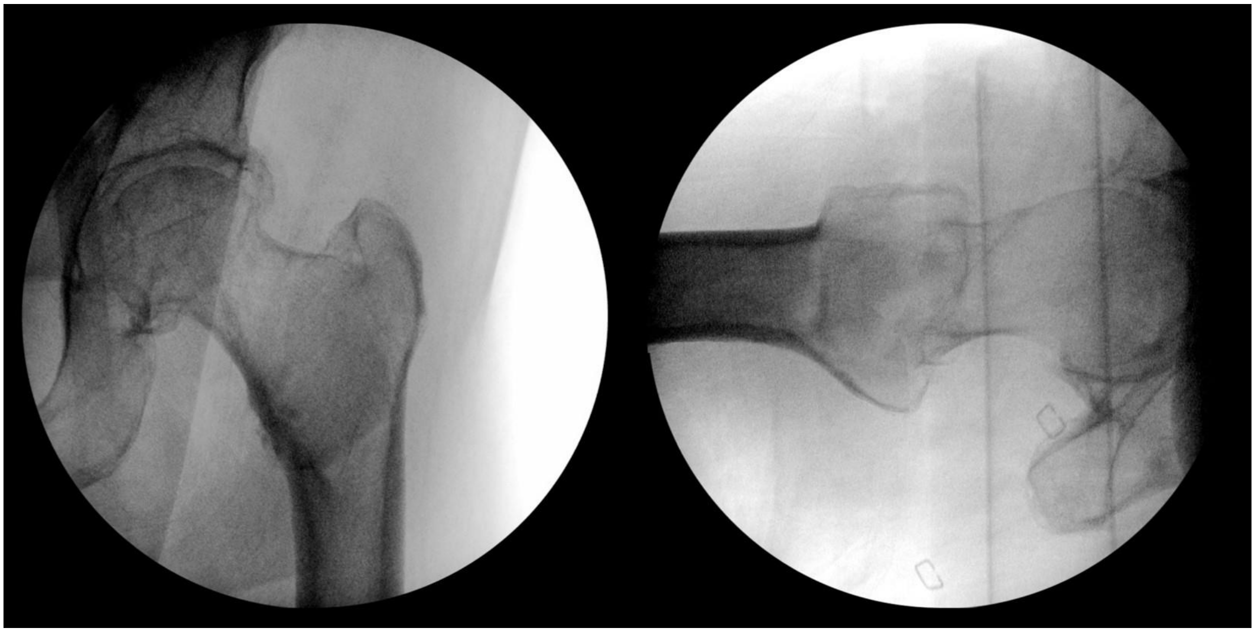


Figure 47: These images were used during the center-center wire placement assessment. These images are taken from a case collected in the OR.

Data for residents in group 2 and group 3 were collected during the fracture course in the fall of 2018. All residents that signed up for the fall 2018 course were asked to complete a pre-course online module. The online module was similar to the previous online module used in the transfer validity study in which residents were presented with a series of AP and lateral images and were asked to properly identify the starting and end points of a center-center guide wire. Residents were given feedback to help illustrate why their starting point was either correct or incorrect (**Figure 48**). The task was intended to get all residents to understand what a center-center wire position should look like on a given fluoroscopic image prior to coming to the OTA fracture course. This allowed the time available during the course to be maximized for wire navigation training.

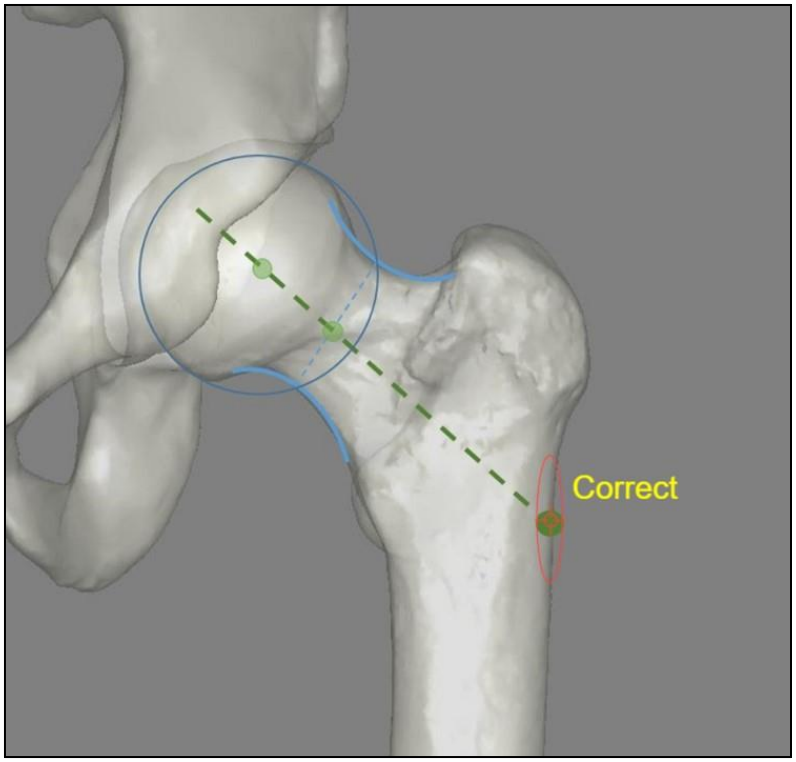


Figure 48: Feedback shown during the online module that residents completed prior to attending the fracture course.

Figure 49: The setup at the OTA fracture course is seen here. Five simulators were set up to train resident pairs on the hip wire navigation task.

Five wire navigation simulators were setup for data collection at the fall 2018 fracture course. Residents worked in sets of pairs as can be seen in **Figure 49**. When residents came to a simulator station they were randomly assigned to be in either group 2 or group 3. Residents in group 2 went through a training exercise on the simulator that involved directing the guide wire into bone towards a series of bubble targets (**Figure 50**). Residents were asked to drill such that their wire would pierce the bubble shown to them. Once a target had been pierced by the guide wire on the fluoroscopic image it would "pop" and a new target would appear. The targets were intentionally placed in positions that did not align with the center-center position that would be eventually tested in the assessment. The goal of the task was to train the residents on the general task of directing a guide wire toward a target, not specifically to have them practice placing a center-center guide wire. The residents were given fifteen minutes to train with the bubble target task. Once a resident had successfully hit all four targets, or the fifteen minutes had expired, they moved on to the assessment task, the same task presented to residents in group 1. During the training, if a resident was struggling to hit a particular target and wanted to move on to the next target, a button could be pressed on the simulator allowing them to skip that target. This allowed residents to get the full benefit of practice at aiming towards multiple targets without getting stuck on one target. Residents that were placed into group 3 did not get to participate in the training exercise. Instead these residents were asked to help "operate" the simulator while their partner went through the training exercise. By asking the residents in group 3 to run the simulator, it kept them engaged in the task and observing their partner. After their partner had gone through the training and the assessment, residents in group 3 were then asked to place a center-center guide wire on a fresh bone model, the same assessment used from group 1 and group 2. In total, 53 residents were placed into group 2 and 41 residents placed into group 3.

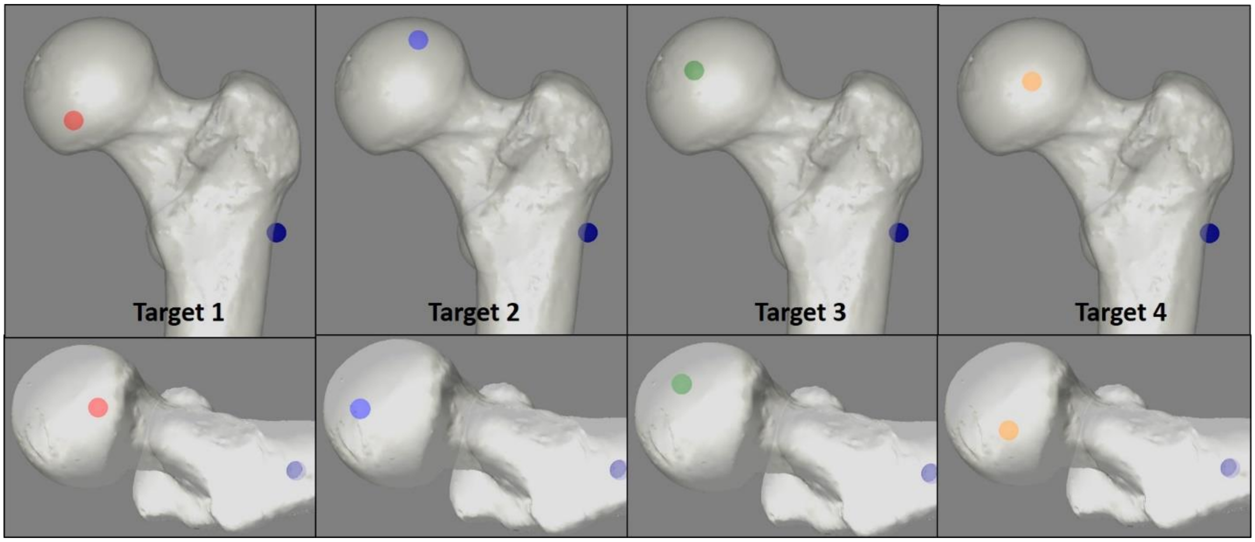


Figure 50: The four bubble targets are shown here. These targets were designed to help residents practice moving in specific planes. They were also designed to intentionally not align with a center-center wire position.

As residents went through the target practice exercise, feedback was given to help emphasize making good wire navigation decisions or point out when they might have made an error in their movement. For example, if a resident made a wire adjustment that had their wire trajectory moving away from the target they were given feedback to show them that they had moved their wire in the wrong direction (**Figure 51**). A dashed line was also temporarily shown to help them project their wire trajectory into bone. Another form of feedback emphasized when to switch between AP and lateral images. If a resident asked for a lateral image when their trajectory in the AP image was not directed on target, then a pop-up window appeared notifying them that they may want to correct their trajectory in that image before switching to the opposite image (**Figure 52**). Residents were given the option to either continue adjusting their trajectory in the AP image or to ignore the suggestion and switch to the lateral image. This allowed them to not feel that their behavior was being overly constrained by the simulator but also helped them understand when they might be making an error.

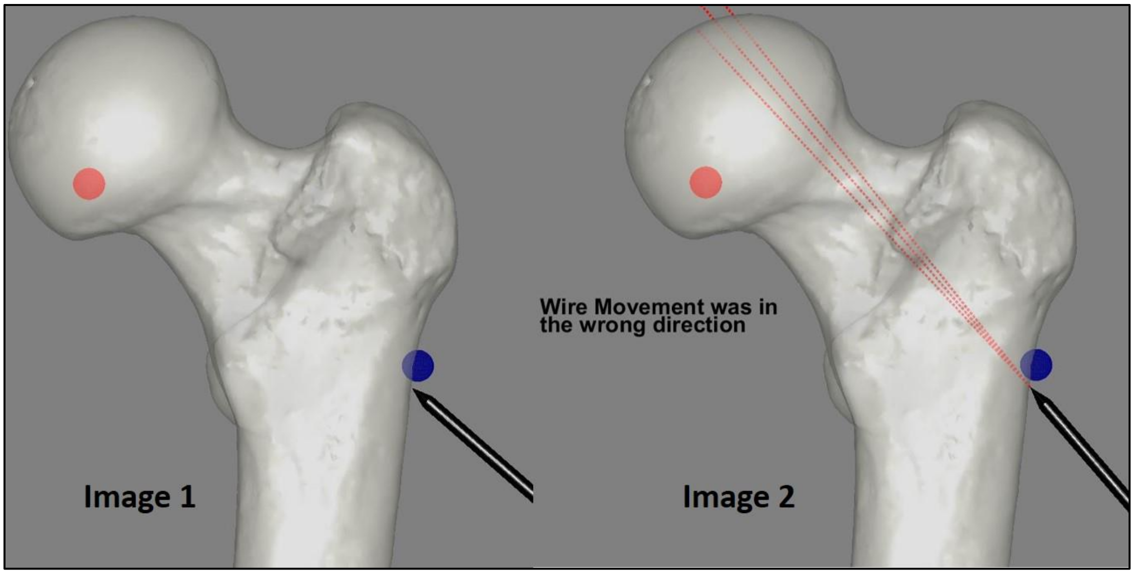


Figure 51: Feedback was provided if a resident moved their wire away from a target.

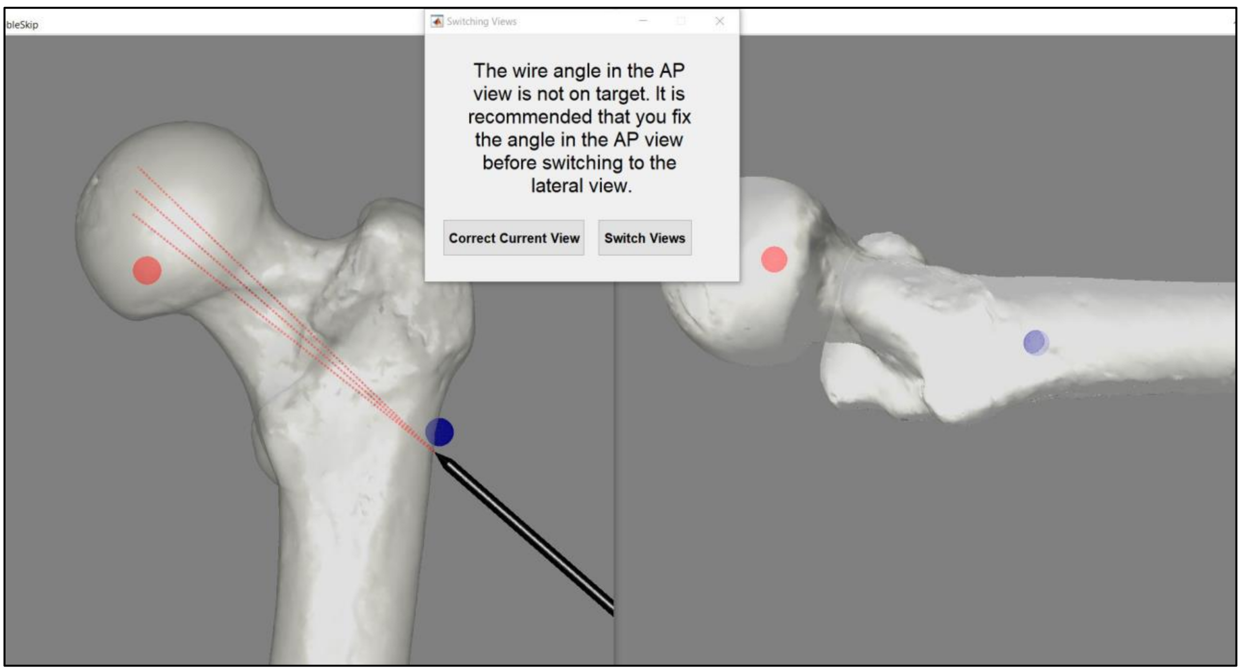


Figure 52: Feedback was provided if a resident requested an image from the alternate plane without first having their wire trajectory properly set.

A fourth group of resident data was used as a comparison for this study. These are data from previous work in which residents were able to go through a more thorough training and practice for a 30-minute time frame. These residents went through the deliberate practice regiment described in the transfer validity study (see CHAPTER 2: SKILL TRANSFER FROM

WIRE NAVIGATION TRAINING ON A SIMULATOR TO PERFORMANCE IN A MOCK OPERATING ROOM ENVIRONMENT for details). At the end of their training, the residents also completed an assessment on the simulator without any additional feedback given. This allows us to compare the training presented to the residents at the OTA fracture course with an extended training that previous residents have received. A total of 36 residents make up group 4.

The metrics used to assess performance on the wire navigation simulator were the tip-apex distance, the number of images used, the overall time, the number of decision errors made, the amount of out of plane movement measured, and the average angle of decision errors. Similar to the composite score calculated when looking at data from the operating room, a composite score of simulator performance was calculated that summed the resident's TAD, out of plane movement, number of decision errors, and the average angle of decision errors. A one-way ANOVA test was performed using SAS statistical software to first see which metrics might have a statistical difference between their means. For the metrics that resulted in a significant ANOVA value they were they run through a Tukey's Standardized Range test at a 95% confidence level. This test allows for a comparison of means between multiple groups.

Results

The residents that participate in the OTA fracture courses range from first year residents to sometimes fourth year residents. For the 68 residents that made up group 1 of this study, the average year in residency was 1.88 years with a standard deviation of 0.96 years. The residents that made up group 2 and group 3 from the fall 2018 course had an average year in residency of 2.1 years with a standard deviation of 0.79 years. In the pre-course online module that residents in group 2 and group 3 were asked to complete, 89 out of 94 residents completed the online module, taking an average of 10 minutes and 3 seconds to complete the exercise. All residents achieved a score of at least 80%, the threshold that was set for competency, in properly identifying the correct starting point and end-point of the wire on the fluoroscopic images.

Data summarizing resident performance on the simulator between the different groups can be seen in **Table 4**. Data is presented as the means and standard deviation for each metric. For a complete look at box and whisker plots of the data, see APPENDIX C. The baseline performance group which did not receive training on the simulator had the highest average values across all performance metrics, or the worst performance values. The group that received training, the target practice group, improved their performance across all metrics in comparison with the no training group. Interestingly, the group that observed the training and the assessment of group 2 also had improved performance metrics in comparison to the no training group. The fourth group which contains data from residents who had trained on the simulator as part of a previous study and received 30 minutes of training appear to have the best performance metrics of all groups, across all metrics. The summative score combining the TAD, number of decision errors, amount of out of plane wire movement, and the average angle of decision errors also shows the same trend in performance between groups. This score is the same summative score that was used to assess performance in the operating room.

Table 4: A Summary of Performance Metrics Between the Different Testing Groups

Group	TAD (mm)	Images	Time (s)	Out of Plane Movement (°)	Decision Errors	Average Angle Error (°)	Standardized Score
No Training (N=68)	19.7 (7.2)	21.7 (10.2)	207 (87.6)	15.5 (20.2)	13.8 (7.6)	3 (2.86)	-0.96 (2.9)
Target Practice (N=53)	15.5 (5.2)	17.3 (8.2)	144.5 (73.8)	6.3 (6.4)	10.7 (7.5)	2.5 (3.2)	0.82 (2.3)
Observation (N= 41)	14.7 (5.6)	18 (7.5)	153.6 (83.2)	7.9 (8.5)	11.4 (7)	2.7 (3.1)	0.65 (2.1)
Full Training (N = 36)	11.3 (5.2)	12.7 (3.3)	129 (61.2)	5 (8.64)	6.9 (4.3)	1.7 (3.1)	2.3 (2.1)

Results from a one-way ANOVA test for each performance metric can be seen in **Table 5**. The TAD, use of time, number of images, number of decision errors, amount of out of plane movement, and the overall score were all found to have significant results for the ANOVA test. This allows for a further investigation into each group to examine if there is a difference between the means for the groups. **Table 6** shows the results from a series of Tukey's Studentized Range test, which is often used to compare the means between multiple groups. In this analysis, there are many results to review. First, across all metrics, there was no difference on the wire navigation assessment between the group that received the target practice training and the group that observed the target practice training. The target practice group had significant levels of improvement on all metrics except for the number of decisions made and the average number of decisions made in comparison with the baseline group. The observation group also had significant levels of improvement in comparison to the baseline group on all metrics except for the overall number of images used, the number of decision errors made, and the average angle of the decision error made. The fourth group which received the most training had significant improvement across all metrics compared to the baseline group except on the average angle of the decision errors made. When comparing the full training group to the target practice group, the only significant differences that existed were in the composite score and the TAD. When comparing the observation group to the full training group, the significant differences were detected between the TAD, number of decision errors made, and the composite score.

Table 5: One-Way ANOVA Test Results

Variable Tested	P-values
TAD	0.001
Time	0.001
Images	0.001
Number of Decision Errors	0.001
Average Angle of Errors	0.181
Out of Plane Movement	0.001
Overall Score	0.001

Table 6: Results from Tukey's Studentized Range Test at 95% Confidence Level. (* indicates a significant result, NS indicates the result was not significant).**

Group Comparison	TAD	Time	Images	Decision Errors	Average Angle of Errors	Out of Plane Movement	Score
No Training vs Target Practice	***	***	***	NS	NS	***	***
No Training vs Observation	***	***	NS	NS	NS	***	***
No Training vs Full Training	***	***	***	***	NS	***	***
Target Practice vs Observation	NS	NS	NS	NS	NS	NS	NS
Target Practice vs Full Training	***	NS	NS	NS	NS	NS	***
Observation vs Full Training	***	NS	NS	***	NS	NS	***

A cumulative distribution plot was created for the composite score of each group (**Figure 53**). When looking at this plot, the x-axis represents the composite score. A score of zero represents the mean for the entire population across all groups. On the x-axis the percentage represents the number of trainees with a worse performance in the group. For example, if the data point is at the 100% value, that means that everyone in the group had a lower score than that participant. The scores at the 50% line represent the mean scores of each group. From this plot we can clearly see the differences between the baseline group, the groups that received target training and observed the target training, and the group that received a full 30 minutes of training, twice as much as the target practice group. Another way to look at this plot would be to examine the number of residents below the composite score of zero, or the population mean. From this chart, it appears that approximately 62% of the baseline group fall beneath the population mean. For the target practice and observation groups, only roughly 30% fall beneath the population mean. And for the group that received the full training, approximately 10% fall beneath the population mean.

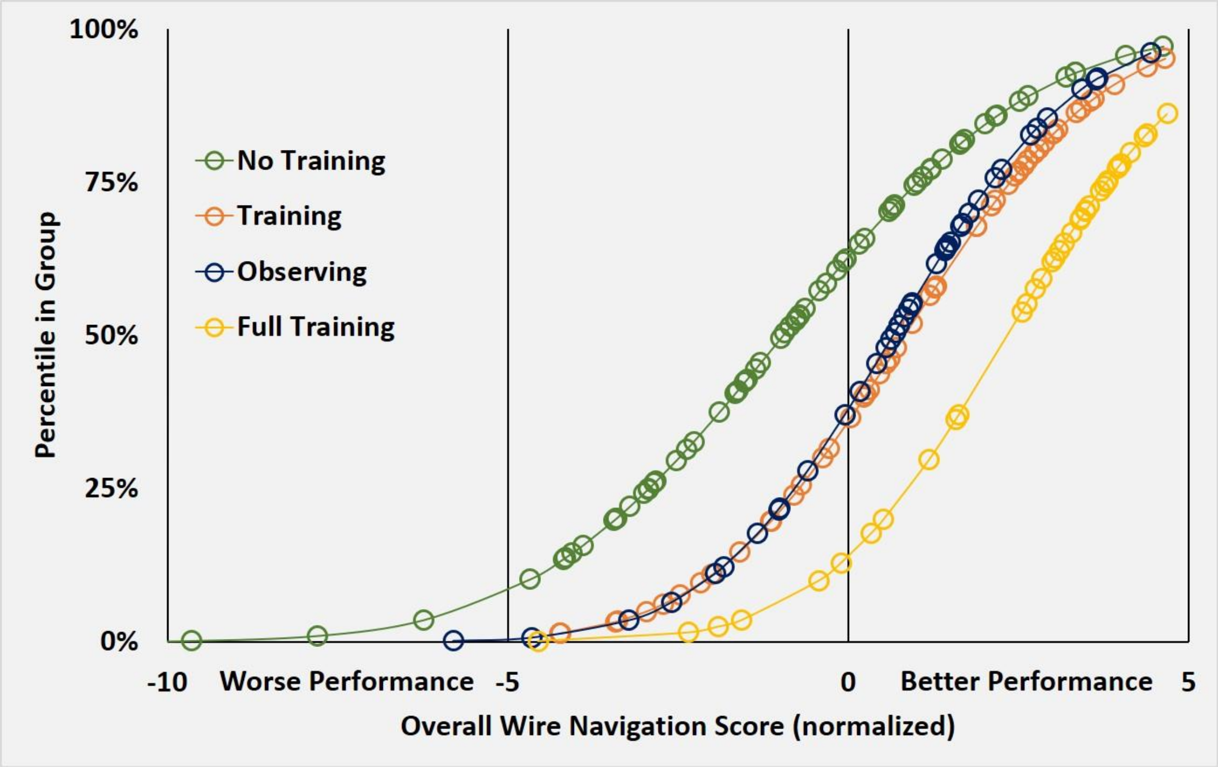


Figure 53: The distribution of the total score on the simulator task between the different groups are shown here. The Full Training group received approximately 30 minutes of practice in comparison to the 15 minutes of practice available to the Target Practice group.

Discussion

A total of 198 residents participated in this study. To date, this appears to be the largest study yet performed examining the effects of simulation-based training on orthopaedic surgical skills. Pedowitz et al. published a study that contained a total of 94 participants from an arthroscopic skills workshop [46]. Schneider et al published a study validating the FAST workstation that contained a total of 73 participants [48]. In orthopaedics research, one of the largest struggles to date has been getting sufficient participant data. In the transfer validity study, our group went to four different residency programs multiple times to get a total of 55 participants. Comprehensive fracture courses put on by organizations like the OTA or AO clearly offer a strong opportunity to gather data on resident performance and training in a short amount of time.

In examining the results of this study, there are several interesting takeaways. First, the results of the target training group and observation group showed that there was no difference in the wire navigation assessment task between these two groups. However, the results also showed that these two groups had a significant improvement in their overall score on the task of placing a center-center wire. One explanation for these results may be in the experimental design that was used during the fracture course. In the experiment, the residents that received the target training then went on to the assessment, with the second group observing during the assessment. Following this assessment, the observation group then performed the same exact assessment, which meant that the wire position they were being asked to place was identical to the one they had just observed. It may be that given the identical cases for the two assessments, the observation group was able to learn what hand position was needed to place the wire properly and was therefore able to achieve the same level of improvement compared to the baseline group as the residents that received the target training. In future exercises this potential confounding factor could be eliminated by having the observation group run through the assessment task first before seeing how the wire should be placed. Alternatively, the simulator could be set up to assess residents on a variety of patient femur anatomies so that residents are not simply learning how to place a wire for a specific femur, but so that they learn the skills of wire navigation that can be applied to a variety of patient anatomies.

Another interesting result was that residents in group 2 and group 3 were able to significantly improve their tip apex distance values compared to the baseline group, however they did not significantly improve the number of decision making errors. Although the training that was provided was designed to target decision errors, this result is not that surprising. The residents at the fracture course were only allotted a maximum of fifteen minutes to practice with the simulator before they moved to the assessment task. Wire navigation is a complex task that requires a strong understanding of the relationship between one's hand position, the fluoroscopic images being taken, and the patient's anatomy. It is unlikely that fifteen minutes is enough for residents to fully grasp all these concepts and make meaningful improvements. The data from the fourth group that received thirty minutes of practice and was able to show a significant improvement across all metrics of performance supports this theory. A question that remains to be answered is how much time and training is needed to sufficiently improve a resident's performance in the operating room? Given the complexity and variety of cases and patient anatomy, it may be that more training will be required to show significant performance improvements in the operating room.

Lastly, the overall distribution of wire navigation performance observed between the different groups is a finding that has many different implications. If we look at that data in the context of a competency assessment we can see how it might be useful. Depending on what score was set as a bar for competency, each group would have different levels of pass rates. If the mean of the population was set as the competency standard, then roughly 60% of the trainees in group one would be deemed not competent. One may argue that this is not a fair judgement to make given they had not yet received the wire navigation training. However, when we think about resident entering the operating room for the first time, likely with little to no training, then having 60% of the residents not reach a level of competency seems rather alarming. Of course, setting the bar for competency may be a political question and will also need to be informed by the performance of expert surgeons to say what competent performance looks like. Once a standard is set by a group of expert surgeons on what competent wire navigation looks like then that can be the level to which trainees work towards during the wire navigation training curriculum.

In conclusion, this study has gathered data on a large array of resident wire navigation performance. From this data, three distinct levels of performance can be observed, based on three different amounts of training. This study is one of the largest orthopaedic surgical skills studies to date and will help the orthopaedic community set the bar for determining what is competent wire navigation performance. The metrics used to assess wire navigation performance in this study have also been shown to be related to surgeon experience in the OR and will be useful when tying simulator performance to performance in the operating room in future studies.

CHAPTER 6: GENERAL DISCUSSION

In orthopaedics, the model for training used to be, “see one, do one, teach one”. It is clear that in the modern world, that model is no longer acceptable. Instead, it is being replaced with a “sim one, do one, teach one” model and potentially even a model of “sim one, prove one, do one, teach one” where residents demonstrate their surgical competency before they enter the operating room. Wire navigation is a fundamental skill in orthopaedics that spans many different surgical procedures. This work has helped move forward the level of evidence needed to have residents and programs subscribe to the “sim one, prove one, do one, teach one” model. Data has shown that with deliberate practice on the wire navigation simulator, residents can improve their tip-apex distance in a mock operating room environment in comparison to residents that received only didactic trainings. The tip-apex measurement is only one component of hip wire navigation surgery. This work examined surgical performance in the operating room and found that by looking at the intra-operative images acquired during a procedure, different types of wire navigation errors could be identified and measured. The number of errors a resident makes in combination with his or her tip-apex distance was then shown to strongly correlate with surgical experience. This is an important component of the evidence needed to understand how simulation training can impact performance in the operating room because it establishes a learning curve of performance that currently exists for residents that do not practice with a simulator. Future studies may examine how a newly trained resident performs in the operating room and how training with the simulator can help advance them further along the learning curve. If 10 cases on average are needed currently to become proficient at wire navigation in the operating room, it would very interesting to see if training on a simulator could make it so that after 1 or 2 cases a resident was performing at the same level of someone who had 10 cases of experience. This would be the ultimate piece of evidence linking simulation training to improved performance in the operating room. Further studies will be done to examine this and explore this idea further.

Determining what level of surgical performance may be required to demonstrate competency on the simulator will also be an ongoing area of study. The data collected thus far shows a broad range of performance by resident surgeons at varying levels of training on the hip wire navigation task. Measuring expert performance on this same task will be necessary to see how the learner performance compares and where a standard of competency may be set. It may be that expert performance is similar to or ahead of the group which received 30 minutes of practice on the wire navigation simulator. In this case, a threshold for competency may end up being the mean of the data presented from the OTA study. However, this also only measures resident performance on one trial of one specific hip wire navigation task. It is likely that to get the most accurate sample of resident performance, multiple measurements will be needed on a variety of hip wire navigation tasks or perhaps even general wire navigation tasks.

One example of another wire navigation task that could be used to assess competency is in treating pediatric elbow fractures. Pediatric elbow fractures are one of the most common types of fractures in children. The treatment of this fracture type often involves driving 3 wires across the fracture after it has properly been reduced to anatomical position. The 3 wires should be in a diverging spread at the fracture line to maximize the amount of stable fixation they provide. Driving 3 wires in diverging paths may be a more challenging task that driving a single wire for a hip fracture. The anatomy of a pediatric elbow is also much smaller than a hip, so the precision required is likely greater. This could potentially be another application that tests resident wire navigation abilities before they go into the operating room. Work has been done to adapt the simulator platform so that it can track multiple wires relative to a fixed pediatric elbow. Like the progression of experiments that were used to evaluate the hip model, studies will need to be done to evaluate training on the pediatric elbow model, to develop measurements of performance in the operating room, and to understand how simulator performance can be tied to performance of treating pediatric elbow fractures in the operating room. It is likely that similar wire navigation decision making errors will be observed in this task to the ones observed in driving a wire for a hip fracture.

In conclusion, the studies presented in this document have furthered our understanding of training and measuring wire navigation performance. More work is needed to be done to definitively link simulator performance to performance in the operating room, but this work has set up a foundation that will make establishing that link possible. This has the potential to lead to better training, more confident surgeons, and better patient outcomes.

REFERENCES

1. Kohn, L.T., J. Corrigan, and M.S. Donaldson, *To err is human : building a safer health system*. 2000, Washington, D.C.: National Academy Press. xxi, 287 p.
2. *ACGME Highlights Its Standards on Resident Duty Hours 2001*. Available from: https://www.acgme.org/About-Us/Publications-and-Resources/Archived-Papers/Position-Papers/ACGME-Highlights-Its-Standards-on-Resident-Duty-Hours-May-2001/Publications/Papers/PositionPapers/HighlightsItsStandardsonResidentDutyHours-.
3. *Working Conditions - Working Time Directive*. 2003; Available from: http://ec.europa.eu/social/main.jsp?catId=706&langId=en&intPageId=205.
4. Bell, R.H., Jr., et al., *Operative experience of residents in US general surgery programs: a gap between expectation and experience.* Ann Surg, 2009. **249**(5): p. 719-24.
5. Blanchard, M.H., S.B. Amini, and T.M. Frank, *Impact of work hour restrictions on resident case experience in an obstetrics and gynecology residency program.* Am J Obstet Gynecol, 2004. **191**(5): p. 1746-51.
6. Bric, J., et al., *Proficiency training on a virtual reality robotic surgical skills curriculum.(Report).* 2014. **28**(12): p. 3343.
7. Carpenter, J.E., et al. *ABOS Surgical Skills Modules for PGY-1 Residents*. 5/20/2018]; Available from: https://www.abos.org/abos-surgical-skills-modules-for-pgy-1-residents.aspx.
8. *ACGME Program Requirements for Graduate Medical Education in Orthopaedic Surgery*. . 5/20/2018]; Available from: https://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/260_orthopaedic_surgery_2017-07-01.pdf.
9. Anderson, D.D., et al., *Objective Structured Assessments of Technical Skills (OSATS) Does Not Assess the Quality of the Surgical Result Effectively.* Clin Orthop Relat Res, 2016. **474**(4): p. 874-81.
10. Cecil, J., A. Gupta, and M. Pirela-Cruz, *An advanced simulator for orthopedic surgical training.* Int J Comput Assist Radiol Surg, 2018. **13**(2): p. 305-319.
11. Khanduja, V., J.E. Lawrence, and E. Audenaert, *Testing the Construct Validity of a Virtual Reality Hip Arthroscopy Simulator.* Arthroscopy, 2017. **33**(3): p. 566-571.
12. Lopez, G., et al., *Construct Validity for a Cost-effective Arthroscopic Surgery Simulator for Resident Education.* Journal of the American Academy of Orthopaedic Surgeons, 2016. **24**(12): p. 886-894.
13. Nousiainen, M.T., et al., *Resident education in orthopaedic trauma: the future role of competency-based medical education.* Bone Joint J, 2016. **98-B**(10): p. 1320-1325.
14. Sonnadara, R.R., et al., *Reflections on competency-based education and training for surgical residents.* J Surg Educ, 2014. **71**(1): p. 151-8.
15. Ferguson, P.C., et al., *Three-year experience with an innovative, modular competency-based curriculum for orthopaedic training.* J Bone Joint Surg Am, 2013. **95**(21): p. e166.
16. Nousiainen, M.T., et al., *Simulation for Teaching Orthopaedic Residents in a Competency-based Curriculum: Do the Benefits Justify the Increased Costs?* Clin Orthop Relat Res, 2016. **474**(4): p. 935-44.
17. Nousiainen, M.T., et al., *Eight-year outcomes of a competency-based residency training program in orthopedic surgery*. Med Teach, 2018: p. 1-13.

18. Pitts, D., D.I. Rowley, and J.L. Sher, *Assessment of performance in orthopaedic training.* J Bone Joint Surg Br, 2005. **87**(9): p. 1187-91.
19. *Specialist Training in Trauma and Orthopaedics*. 2015 [cited 2019 March 1st]; Available from: https://www.iscp.ac.uk/static/public/syllabus/syllabus_to_2017.pdf.
20. Swing, S.R., et al., *Educational milestone development in the first 7 specialties to enter the next accreditation system.* J Grad Med Educ, 2013. **5**(1): p. 98-106.
21. Coughlin, R.P., et al., *A Validated Orthopaedic Surgical Simulation Model for Training and Evaluation of Basic Arthroscopic Skills.* J Bone Joint Surg Am, 2015. **97**(17): p. 1465-71.
22. Chong, A.C., et al., *Validation of Improvement of Basic Competency in Arthroscopic Knot Tying Using a Bench Top Simulator in Orthopaedic Residency Education.* Arthroscopy, 2016. **32**(7): p. 1389-99.
23. Braman, J.P., et al., *Development and validation of a basic arthroscopy skills simulator.* Arthroscopy, 2015. **31**(1): p. 104-12.
24. Dougherty, P.J., et al., *Do scores of the USMLE Step 1 and OITE correlate with the ABOS Part I certifying examination?: a multicenter study.* Clin Orthop Relat Res, 2010. **468**(10): p. 2797-802.
25. Mankin, H.J., *The Orthopaedic In-Training Examination (OITE).* Clin Orthop Relat Res, 1971. **75**: p. 108-16.
26. Cross, M.B., et al., *An analysis of the musculoskeletal trauma section of the Orthopaedic In-Training Examination (OITE).* J Bone Joint Surg Am, 2011. **93**(9): p. e49.
27. Putnam, M.D., et al., *On orthopedic surgical skill prediction--the limited value of traditional testing.* J Surg Educ, 2015. **72**(3): p. 458-70.
28. Koehler, R.J., et al., *The Arthroscopic Surgical Skill Evaluation Tool (ASSET).* Am J Sports Med, 2013. **41**(6): p. 1229-37.
29. Koehler, R.J., et al., *Assessing Diagnostic Arthroscopy Performance in the Operating Room Using the Arthroscopic Surgery Skill Evaluation Tool (ASSET).* Arthroscopy, 2015. **31**(12): p. 2314-9 e2.
30. Martin, J.A., et al., *Objective structured assessment of technical skill (OSATS) for surgical residents.* Br J Surg, 1997. **84**(2): p. 273-8.
31. VanHeest, A., et al., *Objective structured assessment of technical skill in upper extremity surgery.* J Hand Surg Am, 2012. **37**(2): p. 332-7 337 e1-4.
32. Gofton, W.T., et al., *The Ottawa Surgical Competency Operating Room Evaluation (O-SCORE): a tool to assess surgical competence.* Acad Med, 2012. **87**(10): p. 1401-7.
33. Van Heest, A.E., et al., *Resident Surgical Skills Web-Based Evaluation: A Comparison of 2 Assessment Tools.* J Bone Joint Surg Am, 2019. **101**(5): p. e18.
34. Rasmussen, J., *Skills, rules, and knowledge*

signals, signs, and symbols, and other distinctions in human performance models. Systems, Man and Cybernetics, IEEE Transactions on, 1983. **SMC-13**(3): p. 257-266.
35. Bouwman, M.J., *Mind over machine: The power of human intuition and expertise in the era of the computer: Hubert L. DREYFUS and Stuart E. DREYFUS Basil Blackwell, Oxford, 1986, xviii + 231 pages, £15.00 (Book Review).* 1988. p. 353-354.
36. Fitts, P.M. and M.I. Posner, *Human performance*. 1979, Westport, Conn.: Greenwood Press. x, 162 p.
37. Gallagher, A.G., *Metric-based simulation training to proficiency in medical education:-what it is and how to do it.* Ulster Med J, 2012. **81**(3): p. 107-13.

38. Sawyer, T., et al., *Learn, See, Practice, Prove, Do, Maintain: An Evidence-Based Pedagogical Framework for Procedural Skill Training in Medicine.* Academic Medicine, 2015. **90**(8): p. 1025-1033.
39. Ericsson, K.A., *The role of deliberate practice in the acquisition of expert performance*, ed. R.T. Krampe and C. Tesch-Römer. 1993: Berlin-Dahlem, Germany : R. Th. Krampe.
40. Ericsson, K.A., *Acquisition and maintenance of medical expertise: a perspective from the expert-performance approach with deliberate practice.* Acad Med, 2015. **90**(11): p. 1471-86.
41. Derossis, A.M., et al., *The effect of practice on performance in a laparoscopic simulator.* Surg Endosc, 1998. **12**(9): p. 1117-20.
42. *Fundamentals of Laparoscopic Surgery*. [cited 2019 March 10]; Available from: www.flsprogram.com.
43. Korndorffer, J.R., Jr., et al., *Simulator training for laparoscopic suturing using performance goals translates to the operating room.* J Am Coll Surg, 2005. **201**(1): p. 23-9.
44. Mashaud, L.B., et al., *Two-year skill retention and certification exam performance after fundamentals of laparoscopic skills training and proficiency maintenance.* Surgery, 2010. **148**(2): p. 194-201.
45. *Fundamentals of Arthroscopy Surgery Training (FAST) Program*. [cited 2019 March 10th]; Available from: https://www.aana.org/aanaimis/Members/Education/FAST-Program.aspx.
46. Pedowitz, R.A., et al., *Objective Assessment of Knot-Tying Proficiency With the Fundamentals of Arthroscopic Surgery Training Program Workstation and Knot Tester.* Arthroscopy, 2015. **31**(10): p. 1872-9.
47. *FAST Arthroscopy Workstation*. [cited 2019 March 10th]; Available from: https://www.sawbones.com/products/skills-training/fast/fast-arthroscopy-workstation-standard.html.
48. Schneider, A.M., R.A. Pedowitz, and D.A. Evans, *Validation of the FAST Workstation as an Objective Evaluator of Hand-Tied Surgical Knots.* Simul Healthc, 2019. **14**(1): p. 29-34.
49. Tofte, J.N., et al., *Knee, Shoulder, and Fundamentals of Arthroscopic Surgery Training: Validation of a Virtual Arthroscopy Simulator.* Arthroscopy, 2017. **33**(3): p. 641-646 e3.
50. Karam, M.D., et al., *Current and future use of surgical skills training laboratories in orthopaedic resident education: a national survey*. J Bone Joint Surg Am, 2013. **95**(1): p. e4.
51. Molho, D.A., et al., *The Grapefruit: An Alternative Arthroscopic Tool Skill Platform.* Arthroscopy, 2017. **33**(8): p. 1567-1572.
52. Sonnadara, R.R., et al., *Orthopaedic Boot Camp II: examining the retention rates of an intensive surgical skills course.* Surgery, 2012. **151**(6): p. 803-7.
53. Sonnadara, R.R., et al., *Orthopedic boot camp: examining the effectiveness of an intensive surgical skills course.* Surgery, 2011. **149**(6): p. 745-9.
54. Sonnadara, R.R., et al., *Toronto orthopaedic boot camp III: examining the efficacy of student-regulated learning during an intensive, laboratory-based surgical skills course.* Surgery, 2013. **154**(1): p. 29-33.

55. Lopez, G., et al., *A Cost-Effective Junior Resident Training and Assessment Simulator for Orthopaedic Surgical Skills via Fundamentals of Orthopaedic Surgery*. Journal of Bone and Joint Surgery-American Volume, 2015. **97a**(8): p. 659-666.
56. Burns, G.T., et al., *Evaluating Internal Fixation Skills Using Surgical Simulation.* J Bone Joint Surg Am, 2017. **99**(5): p. e21.
57. Hip. *Hip Fractures Among Older Adults*. 9/20/2016 5/20/2018]; Available from: https://www.cdc.gov/homeandrecreationalsafety/falls/adulthipfx.html.
58. Schneider, E.L. and J.M. Guralnik, *The aging of America. Impact on health care costs.* JAMA, 1990. **263**(17): p. 2335-40.
59. Johnson, L.J., et al., *Measuring tip-apex distance using a picture archiving and communication system (PACS).* Injury, 2008. **39**(7): p. 786-90.
60. Baumgaertner, M.R., et al., *The value of the tip-apex distance in predicting failure of fixation of peritrochanteric fractures of the hip.* J Bone Joint Surg Am, 1995. **77**(7): p. 1058-64.
61. Taylor, L.K., et al., *Developing an objective assessment of surgical performance from operating room video and surgical imagery.* IISE Transactions on Healthcare Systems Engineering, 2018. **8**(2): p. 110-116.
62. Sugand, K., et al., *Training effect of a virtual reality haptics-enabled dynamic hip screw simulator.* Acta Orthop, 2015. **86**(6): p. 695-701.
63. Akhtar, K., et al., *Training safer surgeons: How do patients view the role of simulation in orthopaedic training?* Patient Saf Surg, 2015. **9**: p. 11.
64. Christian, M.W., et al., *Construct Validation of a Novel Hip Fracture Fixation Surgical Simulator*. J Am Acad Orthop Surg, 2018. **26**(19): p. 689-697.
65. Vankipuram, M., et al., *A virtual reality simulator for orthopedic basic skills: a design and validation study.* J Biomed Inform, 2010. **43**(5): p. 661-8.
66. Akhtar, K., et al., *Training safer orthopedic surgeons. Construct validation of a virtual-reality simulator for hip fracture surgery.* Acta Orthop, 2015. **86**(5): p. 616-21.
67. Froelich, J.M., et al., *Surgical simulators and hip fractures: a role in residency training?* J Surg Educ, 2011. **68**(4): p. 298-302.
68. Tillander, B., et al., *A virtual reality trauma simulator*. Med Teach, 2004. **26**(2): p. 189-91.
69. Sugand, K., et al., *Teaching basic trauma: validating FluoroSim, a digital fluoroscopic simulator for guide-wire insertion in hip surgery.* Acta Orthop, 2018. **89**(4): p. 380-385.
70. Long, S., G.W. Thomas, and D.D. Anderson, *Designing an Affordable Wire Navigation Surgical Simulator.* J Med Device, 2016. **10**(3).
71. Kho, J.Y., et al., *A Hybrid Reality Radiation-Free Simulator for Teaching Wire Navigation Skills.* J Orthop Trauma, 2015. **29**(10): p. e385-90.
72. Karam, M.D., et al., *Surgical Coaching from Head-Mounted Video in the Training of Fluoroscopically Guided Articular Fracture Surgery.* J Bone Joint Surg Am, 2015. **97**(12): p. 1031-9.
73. Tay, C., A. Khajuria, and C. Gupte, *Simulation training: a systematic review of simulation in arthroscopy and proposal of a new competency-based training framework.* Int J Surg, 2014. **12**(6): p. 626-33.
74. Marsh, J.L. and B.O. Westerlind. *Wire Navigation into the Femoral Head*. 2013; Available from: https://abos.org/media/7102/1._wire_navigation_into_the_femoral_head.mp4.

75. Taylor, L.K., et al., *Assessing Wire Navigation Performance in the Operating Room.* J Surg Educ, 2016. **73**(5): p. 780-7.
76. Ziv, A., et al., *Simulation-based medical education: an ethical imperative.* Acad Med, 2003. **78**(8): p. 783-8.
77. *Technical Specification Sheet for the Edge FaroArm and ScanArm*. 5/20/2018]; Available from: https://knowledge.faro.com/Hardware/FaroArm_and_ScanArm/USB_FaroArm/Technical_Specification_Sheet_for_the_Edge_FaroArm_and_ScanArm?mt-learningpath=faro_edge_downloads.
78. Bydon, M., et al., *Incidence of Sacral Fractures and In-Hospital Postoperative Complications in the United States An Analysis of 2002-2011 Data.* Spine, 2014. **39**(18): p. E1103-E1109.
79. Riehl, J. and J. Widmaier, *A Simulator Model for Sacroiliac Screw Placement.* Journal of Surgical Education, 2012. **69**(3): p. 282-285.
80. Broder, J.M., *In science's name, lucrative trade in body parts.* N Y Times Web, 2004: p. A1, A16.
81. Tonetti, J., et al., *Percutaneous ilio-sacral screw insertion. fluoroscopic techniques.* Orthop Traumatol Surg Res, 2013. **99**(8): p. 965-72.
82. Alvis-Miranda, H.R., et al., *Sacroiliac screw fixation: A mini review of surgical technique.* J Craniovertebr Junction Spine, 2014. **5**(3): p. 110-3.
83. Keating, J.F., et al., *Early fixation of the vertically unstable pelvis: the role of iliosacral screw fixation of the posterior lesion.* J Orthop Trauma, 1999. **13**(2): p. 107-13.
84. Routt, M.L., Jr., P.T. Simonian, and W.J. Mills, *Iliosacral screw fixation: early complications of the percutaneous technique.* J Orthop Trauma, 1997. **11**(8): p. 584-9.
85. Tonetti, J., et al., *Clinical results of percutaneous pelvic surgery. Computer assisted surgery using ultrasound compared to standard fluoroscopy.* Comput Aided Surg, 2001. **6**(4): p. 204-11.
86. Mendel, T., et al., *CT-based 3-D visualisation of secure bone corridors and optimal trajectories for sacroiliac screws.* Injury, 2013. **44**(7): p. 957-63.
87. Francis, D.M., *Surgical decision making.* ANZ J Surg, 2009. **79**(12): p. 886-91.
88. Pugh, C.M., et al., *Faculty evaluation of simulation-based modules for assessment of intraoperative decision making.* Surgery, 2011. **149**(4): p. 534-42.
89. Cristancho, S.M., et al., *When surgeons face intraoperative challenges: a naturalistic model of surgical decision making.* Am J Surg, 2013. **205**(2): p. 156-62.
90. Nathwani, J.N., et al., *Relationship Between Technical Errors and Decision-Making Skills in the Junior Resident.* J Surg Educ, 2016. **73**(6): p. e84-e90.

APPENDIX A

Simulator Decision Analysis Code

```
function [decisionErrors,averageAngleError,outPlaneMvmt] =
simulatorDecAnalysis(foldStr,folderName,f1Model,checkModel,dateCheck)
%this function is designed to read in data from a simulator trial and
%analyze the wire adjusments throughout the trial to measure the number
%and severity of decision errors made.

%this defines the ideal end point of the perfect wire
%placement
idealEnd = [84.66 88.69 -15.03];


    %setting up to read in data from a series of folders
    str1= 'APtip';
    str2= '.mat';
    pathStr= ['C:\\Users\\stelong\\Desktop\\Sim Data\\',folderName,'\\'];

    %this section of code used to get file creation times so that
    %the images can be analyzed based on the order they were taken in
    DOSstr= 'wmic datafile where name=';
    if(dateCheck == 1)
    DOSstr2 = ' get creationdate';
    else
    DOSstr2 = ' get lastmodified';
    end
    apVectStr= 'APvect';
    latVectStr = 'Latvect';
    shotAPStr = [pathStr,foldStr,'AP_count.mat'];
    shotLatStr = [pathStr,foldStr,'Lat_count.mat'];

    skipValAP = 0;
    skipValLat = 0;

    %using a try/catch statement to catch errors that may occur in loading
    %data from a user file
    try
            shotCountAP= importdata(shotAPStr);
            shotCountLat= importdata(shotLatStr);

    catch
            %if there is an error in reading in the summative data
            %then the number of images must be manually calculated by
            %reading through the folder list
```

```
        newPath = [pathStr,foldStr];
        dataList = dir(newPath);
        listLen = length(dataList);
        for ijk = 3:listLen
            dateVar = dataList(ijk).date(13:end);
            date = str2double(dataList(ijk).date(1:2));
            hours = str2double(dateVar(1:2));
            min = str2double(dateVar(4:5));
            sec = str2double(dateVar(7:8));
            timeVal(ijk-2,:) = [date*24*3600 + hours*3600 + min*60 + sec,ijk-2];
        end
        timeVal = sortrows(timeVal,-1);

        shotCountAP = 0;
        shotCountLat = 0;
        for jkl = 1:length(timeVal)
            tipStr = dataList(timeVal(jkl,2)+2).name;
            if(tipStr(1:5)== 'APtip')
                shotCountAP = shotCountAP+1;
            elseif(tipStr(1:6)=='Lattip')
                shotCountLat = shotCountLat+1;
            end
        end
end

%setting up variables for analyzing AP and Lateral image requests
AP_matrix= zeros(shotCountAP, 3);
AP_time= zeros(shotCountAP,1);
Lat_matrix= zeros(shotCountLat, 3);
Lat_time= zeros(shotCountLat, 1);
countAP = 1;
countLat = 1;

%collect data on all the AP images:
for i= 1:shotCountAP

    str3= num2str(i);
    fileString= strcat(pathStr,foldStr,'\',str1,str3,str2);
    vectString = strcat(pathStr,foldStr, apVectStr, str3, str2);
    try
    coords= load(fileString);
    pt2 = coords.tip;
    apVect = load(vectString);
    apVect = apVect.unitVect;
    inPoly = inpolyhedron(checkModel.faces,checkModel.vertices,pt2);
    catch
```

```
        inPoly=10;
    end

    %checking to see if it was a valid point that was in the bone:
    if(inPoly==1)
        if(pt2(2)<=60)
            skipValAP = skipValAP+1;
            %disp('pot error');
        else
            %calculating the angle in the AP plane:
            [minTAD,~] = minTADProj(pt2, apVect, idealEnd);
            curTAD = ceil(sqrt((idealEnd(1)-pt2(1))^2 + (idealEnd(3)-pt2(3))^2)) +
ceil(sqrt((idealEnd(2)-pt2(2))^2 + (idealEnd(3)-pt2(3))^2));

            apAng = atand(apVect(3)/apVect(1));
            latAngAPIm = atand(apVect(2)/apVect(3));

            lineVar = [pt2, apVect];
            [points] = intersectLineMesh3d(lineVar, f1Model.vertices, f1Model.faces);
            if(isempty(points))
                lineVar = [pt2, -1*apVect];
                [points] = intersectLineMesh3d(lineVar, f1Model.vertices, f1Model.faces);
            end

            if(isempty(points))
                skipValAP = skipValAP+1;
            else
                [~,ind] = max(points(:,3));
                wireIntAP((i-skipValAP),:) = points(ind,:);

                %calculating the best posible vector and angle given the
                %point where the wire intersects with bone:
                bestVect = (idealEnd - wireIntAP((i-skipValAP),:))/norm(idealEnd-wireIntAP((i-
skipValAP),:));
                bestAngAP = atand(bestVect(3)/bestVect(1));
                bestAngLat = atand(bestVect(2)/bestVect(3));

                %the difference between the current angle and the best
                %angle for the current point:
                diffAPList((i-skipValAP),:) = abs(apAng-bestAngAP);
                diffLatListAP((i-skipValAP),:) = latAngAPIm;
                diffAng = diffAng + (abs(apAng-bestAngAP));

                AP_matrix (i-skipValAP,:)= coords.tip;
                testString = [DOSstr,'"',fileString,'"',DOSstr2];
                [~,str] = dos(testString);
```

```
            c = textscan(str,'%s');
            createDate = c{1,1}{2,1};
            createDate = createDate(1:14);
            TimeNumber = str2double(createDate);
            AP_time ((i-skipValAP),:)= TimeNumber;
            decisionMatAP(i-skipValAP,:) = [AP_time((i-skipValAP),:),diffAPList((i-
skipValAP),:),diffLatListAP((i-skipValAP),:),pt2,1,minTAD,curTAD];
            countAP = countAP + 1;
          end
        end
      else
          skipValAP = skipValAP+1;
      end
    end

    lstr1= 'Lattip';
    lstr2= '.mat';

    %collect data on all the lateral images:
    for i= 1:shotCountLat

      lstr3= num2str(i);
      fileStringLat= strcat(pathStr,foldStr,'\',lstr1, lstr3, lstr2);
      vectStringLat = strcat(pathStr,foldStr, latVectStr, lstr3, str2);
      try
      Latcoords= load(fileStringLat);
      latVect = load(vectStringLat);
      latVect = latVect.unitVect;
      pt2 = Latcoords.tip;
      inPoly = inpolyhedron(checkModel.faces,checkModel.vertices,pt2);
      catch
        inPoly=10;
      end

      %checking to see if it was a valid point that was in the bone:
      if(inPoly==1)
        if(pt2(2)<=60)
          %disp('pot error');
          skipValLat = skipValLat + 1;
        else
          [minTAD,~] = minTADProj(pt2, latVect, idealEnd);
          curTAD = ceil(sqrt((idealEnd(1)-pt2(1))^2 + (idealEnd(3)-pt2(3))^2)) +
ceil(sqrt((idealEnd(2)-pt2(2))^2 + (idealEnd(3)-pt2(3))^2));
          latAng = atand(latVect(2)/latVect(3));
          apAngLatIm = atand(latVect(3)/latVect(1));
          %diffAng = diffAng + (abs(latAng-latAngIdeal));
```

```
            lineVar = [pt2, latVect];
            [points] = intersectLineMesh3d(lineVar, f1Model.vertices, f1Model.faces);
            if(isempty(points))
               lineVar = [pt2, -1*latVect];
               [points] = intersectLineMesh3d(lineVar, f1Model.vertices, f1Model.faces);
            end

            if(isempty(points))
               skipValLat = skipValLat+1;
            else
               [~,ind] = max(points(:,3));
               wireIntLat((i-skipValLat),:) = points(ind,:);

               bestVect = (idealEnd - wireIntLat((i-skipValLat),:))/norm(idealEnd-wireIntLat((i-
skipValLat),:));
               bestAngLat = atand(bestVect(2)/bestVect(3));

               diffLatList((i-skipValLat),:) = abs(latAng-bestAngLat);
               diffAPListLat((i-skipValLat),:) = apAngLatIm;

               diffAng = diffAng + (abs(latAng-bestAngLat));

               Lat_matrix(i-skipValLat,:)= Latcoords.tip;
               testString = [DOSstr,'"',fileStringLat,'"',DOSstr2];
               [~,str] = dos(testString);
               c = textscan(str,'%s');
               createDate = c{1,1}{2,1};
               createDate = createDate(1:14);
               TimeNumber = str2double(createDate);
               Lat_time ((i-skipValLat),:)= TimeNumber;
               decisionMatLat(i-skipValLat,:) = [Lat_time((i-skipValLat),:),diffLatList((i-
skipValLat),:),diffAPListLat((i-skipValLat),:),pt2,2,minTAD,curTAD];
               countLat = countLat+1;
            end
         end
      else
            skipValLat = skipValLat + 1;
      end
   end

%these contain all the data on the AP and lateral images taken throughout
%a given trial. The data is ordered based on its time values
decisionMatComb = [decisionMatAP; decisionMatLat];
decisionMatSorted = sortrows(decisionMatComb,1);
decisionMat = decisionMatSorted(:,2:end);
```

```
standDev = std(decisionMat(:,4));

%this looks for potential errors where the simulator showed an erratic bone
%position that should not be considered as an actual wire position
[~,locs] = findpeaks(decisionMat(:,4),'Threshold',standDev);

if(isempty(locs))
    locs=0;
end

%start of decision scoring code
weightedAngle = 0;
badAdvanceCount = 0;
badDecCount = 0;
outPlaneCount = 0;
outPlaneMvmt = 0;
badSwitch = 0;
badReverse = 0;
outPlane = 0;
firstTime =1;
overFluoro = 0;
count = 1;

%decision counter
for i=1:(numel(decisionMat(:,1))-1)

    if(i==locs(count))
        count = count+1;
        %disp('pot error');
        if(length(locs)<count)
            locs(count) = 0;
        end
    else

        %out of plane movement:
        if(decisionMat(i,5)>35 && decisionMat(i+1,5)>35) %only look at out of plane movement
when  tip is out of bone
            if(decisionMat(i,6) == decisionMat(i+1,6)) %if they request 2 APs or 2 Lats
                ang1 = decisionMat(i,2);
                ang2 = decisionMat(i+1,2);
                diffAng = abs(ang1 - ang2);
                if(abs(diffAng) > 1.5)
                    outPlaneMvmt = outPlaneMvmt + diffAng;
                    outPlaneCount = outPlaneCount + 1;
                end
```

```
        end
    end

    if(decisionMat(i,6)~=decisionMat(i+1,6))
        if(firstTime ==1)
            prevAngle = decisionMat(i,1);
            firstTime = 0;
        else
            if(prevAngle<decisionMat(i+1,1) && abs(prevAngle-decisionMat(i+1,1))>1)
                outPlane = outPlane + 1;
            end
            prevAngle = decisionMat(i,1);
        end
    end

    if(decisionMat(i,1)<3.5 && (abs(decisionMat(i,5)-decisionMat(i+1,5))>30) &&
decisionMat(i+1,5)>decisionMat(i,5) && decisionMat(i,6)==decisionMat(i+1,6))
        badReverse = badReverse+1;
    end


    if(decisionMat(i,6)== decisionMat(i+1,6)) %if image is either AP or Lat
        ang1 = decisionMat(i,1);
        ang2 = decisionMat(i+1,1);
        firstPt = decisionMat(i,3:5);
        secondPt = decisionMat(i+1,3:5);
        diff = norm(secondPt - firstPt);
        if(ang1 >= 3.5 && ang2 >=3.5 && firstPt(3)>50 && diff>5 && secondPt(3)<65) %if
they are angled wrong but advance the wire
            badAdvanceCount = badAdvanceCount + 1;
        end
    end

    if(decisionMat(i,6) ~= decisionMat(i+1,6)) %if images were not both AP or Lat it means
they swithced
        ang1 = decisionMat(i,1);
        firstPt = decisionMat(i,3:5);
        if(ang1 >= 3.5 && firstPt(3)>0) %if they switched but the angle was bad, they shouldn't
have swithced
            badSwitch= badSwitch + 1;
        end
    else %if images were both AP or Lat
        ang1 = decisionMat(i,1);
        ang2 = decisionMat(i+1,1);
        angDiff = abs(ang2-ang1);
        firstPt = decisionMat(i,3:5);
```

```
            secondPt = decisionMat(i+1,3:5);
            dist = sqrt((firstPt(1)-secondPt(1))^2+(firstPt(2)-secondPt(2))^2+(firstPt(3)-
secondPt(3))^2);
            if((firstPt(3)<63 && firstPt(3)>-10 && dist<30) || (dist<5 && angDiff<2)) %if they are
in a good position but only move 30 mill before checking another shot
                overFluoro = overFluoro + 1;
            end
        end

        if(decisionMat(i,6) == decisionMat(i+1,6)) %if images were both AP or Lat
            ang1 = decisionMat(i,1);
            ang2 = decisionMat(i+1,1);
            diffAng = ang1 - ang2;
            if(abs(diffAng) > 1.5 && (ang2>ang1)) %if they make a good adjustment vs bad
adjustment
                if(ang2>3.5)
                    badDecCount = badDecCount + 1;
                    weightedAngle = weightedAngle + abs(diffAng);
                end
            end
        end


    end

end


decisionErrors = badDecCount + badSwitch + outPlane + badReverse + overFluoro +
badAdvanceCount;
averageAngleError = weightedAngle/badDecCount;


end
```

OR Decision Analysis Code

```
function [decisionErrors,averageAngleError] = orDecAnalysis(xLfile,numSheets)

xlFile = 'C:\Users\stelong\Documents\OR Decisions\DHS Data latest.xlsx';

%data is read in from an excel file
%number of sheets is set to equal the number of cases for analysis
numSheets = 21;

badDecCount = zeros(numSheets,1);
```

```
weightedAngle = zeros(numSheets,1);
badSwitch = zeros(numSheets,1);
badAdvanceCount = zeros(numSheets,1);
caseImages = zeros(numSheets,1);
badDecSum = zeros(numSheets,1);
overFluoro = zeros(numSheets,1);
badReverse = zeros(numSheets,1);
outPlane = zeros(numSheets,1);

%read in data from excel files
for ijk = 1:numSheets

    clear xlData;
    [xlData,~,~] = xlsread(xlFile,ijk);

    caseImages(ijk) = length(xlData(:,1));
    images = length(xlData(:,1));
    shotCountAP = sum(xlData(:,3),'omitnan');
    shotCountLat = images - shotCountAP;

    apMat = zeros(shotCountAP,length(xlData(1,:)));
    latMat = zeros(shotCountLat,length(xlData(1,:)));
    apCount = 1;
    latCount = 1;
    skipVal = 0;
    firstTime =1;

    for counter = 1:images
      if(xlData(counter,3) == 1)
        apMat(apCount,:)= xlData(counter,:);
        apCount = apCount + 1;
      else
        latMat(latCount,:) = xlData(counter,:);
        latCount = latCount + 1;
      end
    end


    AP_matrix= zeros(shotCountAP, 2);
    AP_time= zeros(shotCountAP,1);
    decisionMatAP = zeros(shotCountAP,11);
    Lat_matrix= zeros(shotCountLat, 2);
    Lat_time= zeros(shotCountLat, 1);
    decisionMatLat = zeros(shotCountLat,11);
```

```
%collect data on all the AP images:
for i= 1:shotCountAP

   if(apMat(i,21) == 1)

      %scaleFact is set based on the wire diameter in mm
      scaleFact = 3.2/apMat(i,12);

      pt2 = scaleFact*[apMat(i,6),apMat(i,7)];
      if(isnan(apMat(i,10))== 0)
         pt1 = scaleFact*[apMat(i,10),apMat(i,11)];
      else
         pt1 = scaleFact*[apMat(i,4) apMat(i,5)];
      end

      vectLen = norm(pt2-pt1);
      vectDiff = pt2 - pt1;
      apVect = vectDiff/vectLen;

      idealEnd = scaleFact*[apMat(i,8),apMat(i,9)];

      apAng = atand(apVect(2)/apVect(1));

      wireIntAP = scaleFact*[apMat(i,4) apMat(i,5)];

      %calculating the best posible vector and angle given the
      %point where the wire intersects with bone:
      bestVect = (idealEnd - wireIntAP)/norm(idealEnd-wireIntAP);
      bestAngAP = atand(bestVect(2)/bestVect(1));

      %the difference between the current angle and the best
      %angle for the current point:
      diffAPList = abs(apAng - bestAngAP);
      if(diffAPList>90)
         diffAPList = 180-diffAPList;
      end

      [minTAD,~] = minTADProj([pt2 1], [apVect 0], [idealEnd 1]);
      distIn = norm(pt2-wireIntAP);
      curTAD = norm(pt2-idealEnd);
      if(distIn<50)
         InOut = 0;
      else
         InOut = 1;
      end
```

```
            AP_matrix(i,:)= pt2;

            AP_time((i),:)= apMat(i,2);
            decisionMatAP(i,:) =
[AP_time(i,:),diffAPList,pt2,1,InOut,wireIntAP,minTAD,distIn,curTAD];
        else
            skipVal = skipVal +1;
        end

    end


    %collect data on all the lateral images:
    for i= 1:shotCountLat

        if(latMat(i,21) == 1)
            %scaleFact is set based on the wire diameter in mm
            scaleFact = 3.2/latMat(i,12);

            pt2 = scaleFact*[latMat(i,6),latMat(i,7)];

            if(isnan(latMat(i,10))== 0)
                pt1 = scaleFact*[latMat(i,10),latMat(i,11)];
            else
                pt1 = scaleFact*[latMat(i,4) latMat(i,5)];
            end

            vectLen = norm(pt2-pt1);
            vectDiff = pt2 - pt1;
            latVect = vectDiff/vectLen;
            idealEnd = scaleFact*[latMat(i,8),latMat(i,9)];


            latAng = atand(latVect(2)/latVect(1));

            wireIntLat = scaleFact*[latMat(i,4) latMat(i,5)];

            bestVect = (idealEnd - wireIntLat)/norm(idealEnd-wireIntLat);
            bestAngLat = atand(bestVect(2)/bestVect(1));

            diffLatList = abs(latAng-bestAngLat);
            if(diffLatList>90)
                diffLatList = 180-diffLatList;
            end
            [minTAD,~] = minTADProj([pt2 1], [latVect 0], [idealEnd 1]);
```

```
            Lat_matrix(i,:)= pt2;
            Lat_time(i,:)= latMat(i,2);
            distIn = norm(pt2-wireIntLat);
            curTAD = norm(pt2-idealEnd);

            if(distIn<50)
               InOut = 0;
            else
               InOut = 1;
            end

            decisionMatLat(i,:) =
[Lat_time(i,:),diffLatList,pt2,2,InOut,wireIntLat,minTAD,distIn,curTAD];
         else
            skipVal = skipVal + 1;
         end

      end

      decisionMatComb = [decisionMatAP; decisionMatLat];
      decisionMatSorted = sortrows(decisionMatComb,1);
      if(skipVal == 0)
         decisionMat = decisionMatSorted(:,2:end);
      else
         decisionMat= decisionMatSorted(skipVal+1:end,2:end);
      end

      %decision counter
      for i=1:(numel(decisionMat(:,1))-1)

         if(decisionMat(i,4)~=decisionMat(i+1,4))
            if(firstTime ==1)
               prevAngle = decisionMat(i,1);
               firstTime = 0;
            else
               if(prevAngle<decisionMat(i+1,1) && abs(prevAngle-decisionMat(i+1,1))>1)
                  outPlane(ijk) = outPlane(ijk)+ 1;
               end
               prevAngle = decisionMat(i,1);
            end
         end

         if(decisionMat(i,1)<3.5 && (decisionMat(i,9)-decisionMat(i+1,9)>30) &&
decisionMat(i,4)==decisionMat(i+1,4))
            badReverse(ijk) = badReverse(ijk)+1;
         end
```

```
        if(decisionMat(i,4) == decisionMat(i+1,4))
           if(decisionMat(i,10)>30)
              moveDist = abs(decisionMat(i+1,9)-decisionMat(i,9));
              tadDiff = abs(decisionMat(i+1,10)-decisionMat(i,10));
              angChange = abs(decisionMat(i+1,1)-decisionMat(i,1));
             if((moveDist<10 && decisionMat(i,4)==0 && decisionMat(i,5)==1) ||
(moveDist<30 && decisionMat(i,4) == 1 && decisionMat(i,5)==1) || (moveDist<5 &&
tadDiff<2 && angChange<2) )
                overFluoro(ijk) = overFluoro(ijk)+1;
             end
           end
        end

        if(decisionMat(i,4) == decisionMat(i+1,4)) %if image is either AP or Lat
           ang1 = decisionMat(i,1);
           ang2 = decisionMat(i+1,1);
           pt1 =decisionMat(i,9);
           pt2 =decisionMat(i+1,9);
           diff = pt2-pt1;
           if(ang1 >= 3.5 && ang2 >=3.5 && pt1<50 && pt1>15 && diff>5 ) %if they are
angled wrong but advance the wire
              badAdvanceCount(ijk) = badAdvanceCount(ijk) + 1;
           end
        end

        if(decisionMat(i,4) ~= decisionMat(i+1,4)) %if images were not both AP or Lat it means
they swithced
           ang1 = decisionMat(i,1);
           if(ang1 >= 3.5 && decisionMat(i,5)==0) %if they switched but the angle was bad,
they shouldn't have swithced
              badSwitch(ijk)= badSwitch(ijk) + 1;
           end
        end

        if(decisionMat(i,4) == decisionMat(i+1,4)) %if images were both AP or Lat
           ang1 = abs(decisionMat(i,1));
           ang2 = decisionMat(i+1,1);
           diffAng = ang1 - ang2;

           if(ang2>ang1)
              if(ang2>3.5)
              badDecCount(ijk) = badDecCount(ijk) + 1;
              weightedAngle(ijk) = weightedAngle(ijk) + abs(diffAng);
              end
```

```
            end

         else
         end


         badDecSum(ijk) = badDecCount(ijk) + badSwitch(ijk) + badAdvanceCount(ijk) +
overFluoro(ijk) + badReverse(ijk) + outPlane(ijk);
      end

end

end
```

APPENDIX B

Table B.1: Data from the Operating Room Analysis. For cases 6 and 8 the attending surgeon placed the final wire, so the TAD was omitted from that score.

Test Arena	Case	Final TAD (mm)	Num Images	PGY year	Num Cases Logged	Num Decision Errors	Avg Angle of Errors	Norm Score
OR	1	19	23	3	2	7	6.17	-0.64
OR	2	23	22	3	2	8	1.91	-0.25
OR	3	9	19	3	10	9	2.60	0.64
OR	4	17	51	2	2	24	2.15	-0.69
OR	5	16	51	3	3	16	4.29	-0.56
OR	6	*	33	3	6	11	2.17	0.16
OR	7	17	26	3	4	5	7.31	-0.58
OR	8	*	44	3	3	18	4.68	-1.06
OR	9	19	21	5	13	4	1.11	0.40
OR	10	11	18	3	9	7	1.58	0.77
OR	11	12	20	5	14	8	2.78	0.43
OR	12	12	20	5	13	5	3.85	0.40
OR	13	25	38	5	8	12	2.34	-0.69
OR	14	20	64	3	4	20	1.63	-0.61
OR	15	10	36	7	25	7	1.31	0.89
OR	16	14	24	3	6	5	6.66	-0.24
OR	17	17	35	3	9	7	2.19	0.21
OR	18	14	31	3	6	8	3.38	0.18
OR	19	16	59	3	6	23	1.65	-0.47
OR	20	9	36	5	14	16	0.93	0.56
OR	21	17	17	4	11	2	0.00	0.86

mock OR	**22**	**29**	**18**	**1**	**1**	**12**	**1.81**	**-0.91**
mock OR	**23**	**27**	**18**	**1**	**1**	**15**	**4.41**	**-1.37**
mock OR	**24**	**37**	**12**	**1**	**1**	**9**	**5.72**	**-2.04**
mock OR	**25**	**22**	**36**	**1**	**1**	**16**	**8.88**	**-1.82**
mock OR	**26**	**22**	**22**	**1**	**1**	**12**	**2.76**	**-0.54**
mock OR	**27**	**24**	**19**	**1**	**1**	**12**	**6.21**	**-1.29**
mock OR	**28**	**25**	**29**	**1**	**1**	**12**	**6.31**	**-1.39**
mock OR	**29**	**34**	**17**	**1**	**1**	**13**	**3.28**	**-1.59**

APPENDIX C

Box and Whisker plots from SAS analysis of simulator data from the experiment run at the OTA comprehensive fracture courses are shown below.

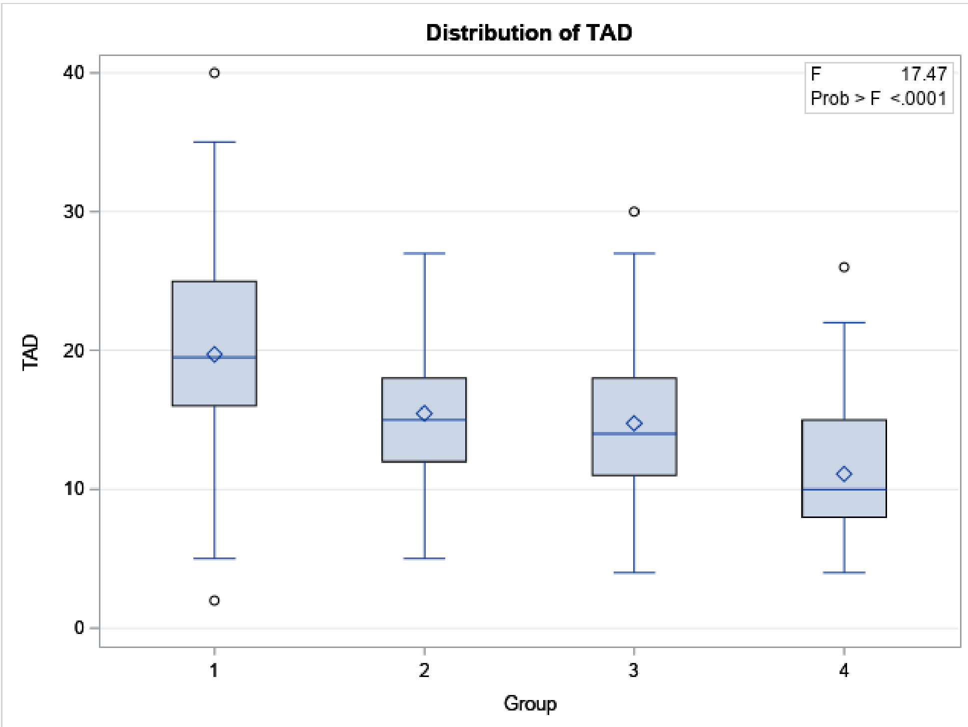


Figure C.1: TAD distributions shown here. Group 1 received no training, group 2 received the target practice training, group 3 observed group 2, and group 4 received a full regiment of training.

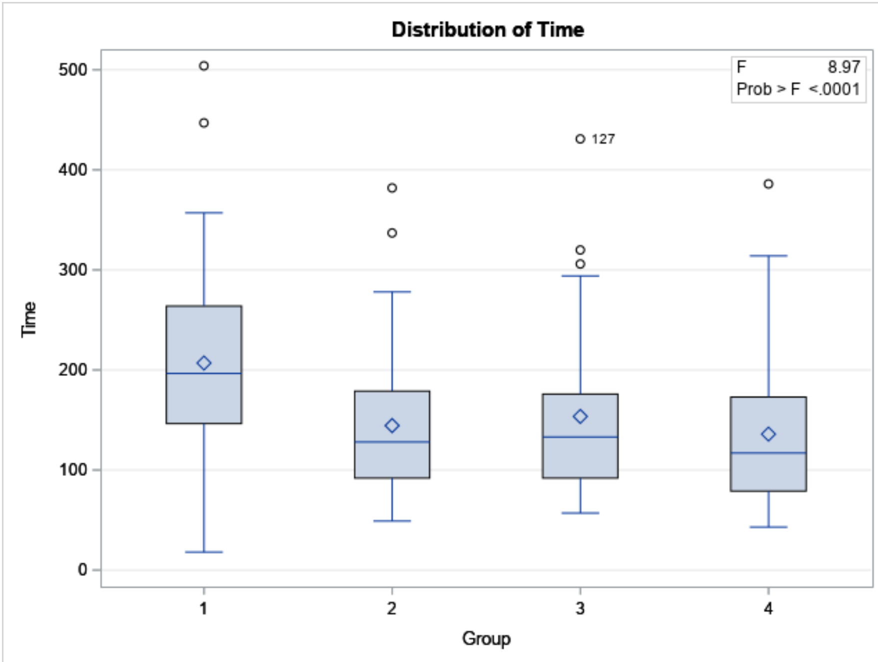


Figure C.2: Distribution of time taken to place the wire is shown here. Group 1 received no training, group 2 received the target practice training, group 3 observed group 2, and group 4 received a full regiment of training.

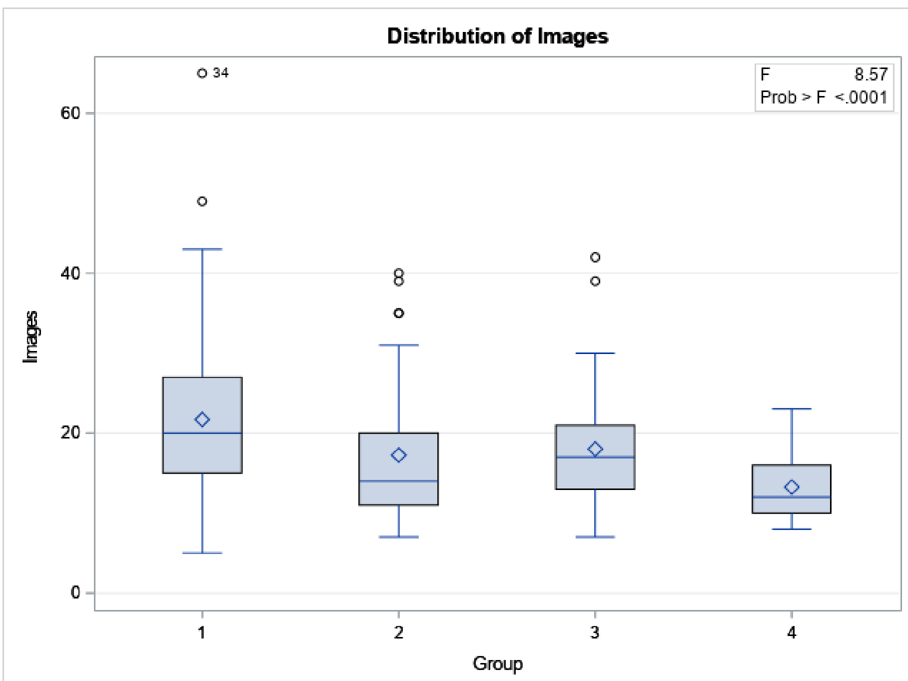


Figure C.3: Distribution of the number of images used is shown here. Group 1 received no training, group 2 received the target practice training, group 3 observed group 2, and group 4 received a full regiment of training.

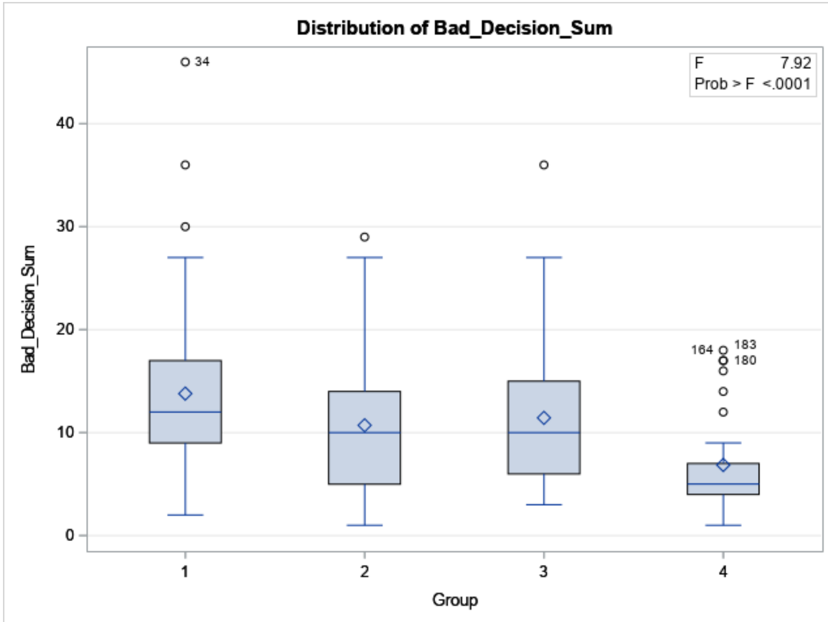


Figure C.4: Distribution of the number of decision errors made is shown here. Group 1 received no training, group 2 received the target practice training, group 3 observed group 2, and group 4 received a full regiment of training.

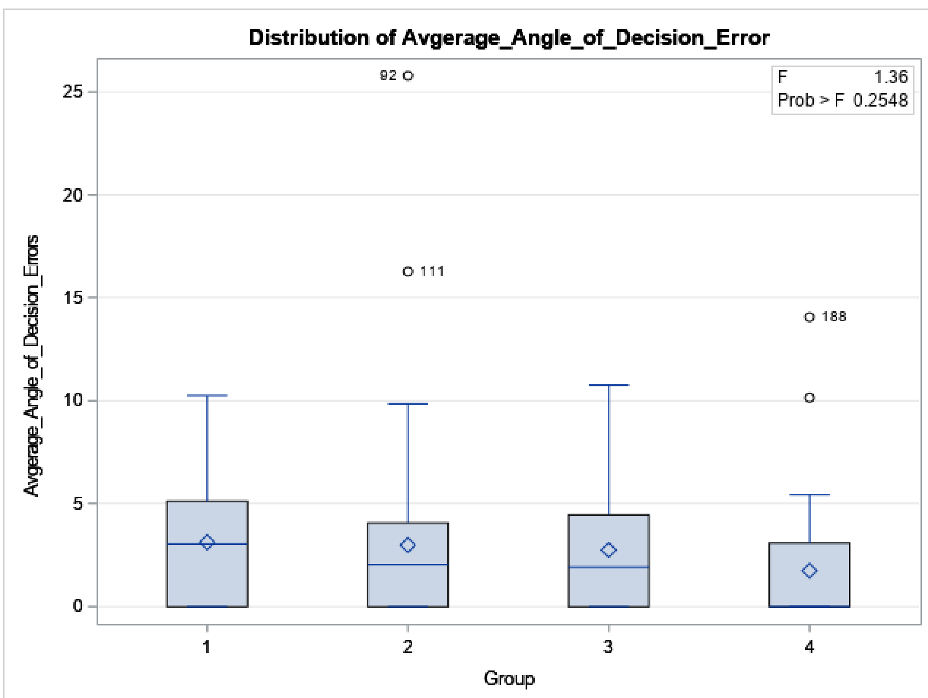


Figure C.5: Distribution of average angle of decision errors is shown here. Group 1 received no training, group 2 received the target practice training, group 3 observed group 2, and group 4 received a full regiment of training.

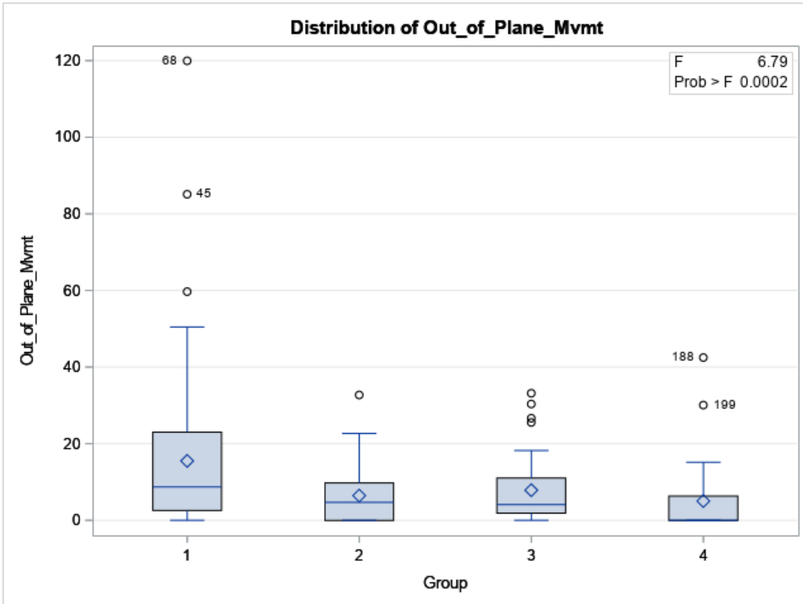


Figure C.6: Distribution of out of plane wire movement shown here. Group 1 received no training, group 2 received the target practice training, group 3 observed group 2, and group 4 received a full regiment of training.

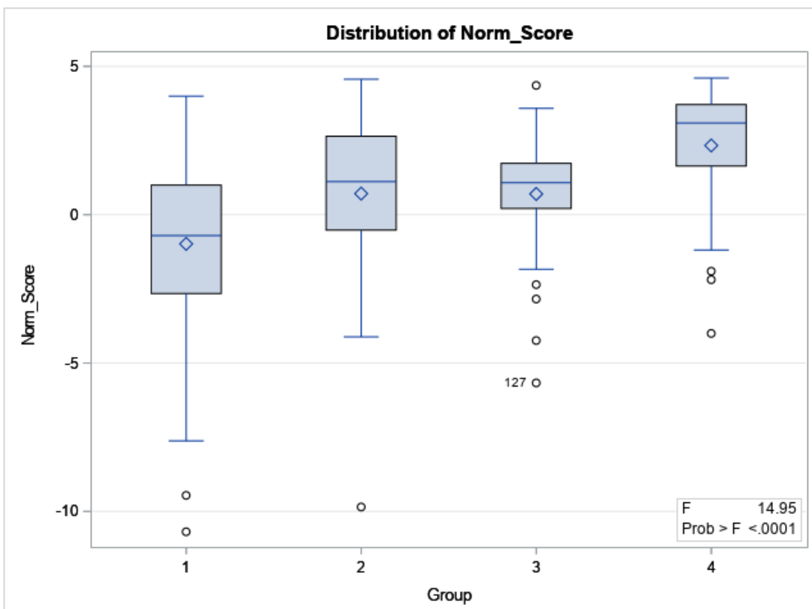


Figure C.7: Distribution of overall scores shown here. Group 1 received no training, group 2 received the target practice training, group 3 observed group 2, and group 4 received a full regiment of training.