1. TITLE PAGE

Title of Project: Simulation to Support Competency-Based Training in Orthopedic Trauma

Principal Investigator and Team Members.

Principal Investigators:

Donald D. Anderson, PhD – Professor, Orthopedics & Rehabilitation University of Iowa
Geb Thomas (contact), PhD – Professor, Industrial & Systems Engineering University of Iowa

Investigators:

Matthew D. Karam, MD – Clinical Professor, Orthopedics & Rehabilitation, University of Iowa
J. Lawrence Marsh, MD – Professor and Chair, Orthopedics & Rehabilitation University of Iowa
Clarence Kreiter, PhD – Professor, Internal Medicine University of Iowa

Organization: University of Iowa

Inclusive Dates of Project: 09/30/2017 - 07/31/2023

Federal Project Officer
David Rodrick, PhD
Agency for Healthcare Research and Quality
david.rodrick@ahrq.hhs.gov
(301) 427-1876

Acknowledgment of Agency Support

This project was funded under grant number R18HS025353 from the Agency for Healthcare Research and Quality (AHRQ), U.S. Department of Health and Human Services (HHS). The authors are solely responsible for this document's contents, findings, and conclusions, which do not necessarily represent the views of AHRQ. Readers should not interpret any statement in this report as an official position of AHRQ or of HHS. None of the authors has any affiliation or financial involvement that conflicts with the material presented in this report.

Grant Award Number: R18HS025353

2. STRUCTURED ABSTRACT

Purpose: Provide scientifically defendable criteria for critical surgical skills at key developmental points in orthopedic residency.

Scope: This research explored the acquisition of fluoroscopically guided wire navigation skill through training and assessment with novel surgical simulators. It developed ground-breaking analysis techniques for assessing wire navigation performance in the operating room from surgical fluoroscopy. The project expanded prior work to more residency programs and to two national organizations: the Orthopaedic Trauma Association, which offers training programs to ~100 residents twice annually, and the American Board of Orthopaedic Surgeons, which reviews the skills of 900 orthopedic surgeons annually.

Methods: Experiments involving orthopedic residents, trauma fellows, and faculty surgeons ranged from simulator training, with associated refinements of both the simulators and training protocols, to evaluation of surgical skills in the operating room. Image analysis methods were developed to objectively evaluate surgical performance based on measuring clinically relevant outcome metrics from intra-operative fluoroscopic images. These methods produced comparable scores whether used by experts or novices.

Results: Simulator training was demonstrated to improve wire navigation skills, both on the simulator AND in the operating room. Performance improvements from simulator training were shown to be comparable to those from experience accrued in the operating room. Simulator performance scores indicative of competence were established, and we demonstrated that training with the simulator allowed residents to attain these scores. Finally, wire navigation skills developed while training on specific simulator tasks were generalizable to other wire navigation tasks.

Key Words: Orthopedic trauma, surgical skills training, operating room performance evaluation

3. PURPOSE

The long-term goal of this research is to develop the scientific evidence needed to bring competency-based evaluation of surgical skills to orthopedics. The objective of this series of studies was to provide the orthopedic training community with scientifically defendable criteria for critical orthopedic surgical skills at three key developmental points in surgical residency: ready to operate, ready to lead a surgery supervised by faculty, and ready for independent practice. The central hypothesis of this work was that orthopedic surgical skill competence can be objectively, quantitatively, and reliably measured from surgical results that impact patient safety. This hypothesis is based on previous multi-institutional simulation studies conducted with residency programs throughout the Midwest region in the Midwest Orthopaedic Surgical Skills (MOSS) Consortium. The research explored training and assessment programs with several novel surgical simulators. It also explored ground-breaking analysis techniques for assessing operating room performance from surgical fluoroscopy and from computational, image-based, 3D model reconstruction. This project further extended previous work to more residency programs and to two national orthopedic organizations: the Orthopaedic Trauma Association (OTA), which offers training programs to over a hundred residents several times a year, and the American Board of Orthopaedic Surgeons (ABOS), which reviews the credentials and skills of 900 orthopedic surgeons each year, certifying many to become official diplomates of ABOS.

The participation of these national organizations, which provide many potential research participants, created the opportunity for achieve exciting new research goals, including three specific aims. **Specific Aim 1** was to measure orthopedic trauma surgery skill through surgical results. The working hypothesis for this aim was that surgical skill in orthopedic trauma can be measured by analyzing radiographic images associated with a surgery. **Specific Aim 2** was to develop simulator-based training programs that improve surgical results. The working hypothesis was that residents who practice key skills until achieving proficiency perform better in the operating room than residents without training. **Specific Aim 3** was to define simulator-based assessments that generalize to clinical performance. The working hypothesis was that simulator assessment measurements can discriminate among levels of surgical skill and can also generalize across surgical tasks and between simulation and clinical practice.

A final objective of this research has been to provide a safe environment for orthopedic surgeons to acquire valuable experience without putting patients at risk. It also helped establish reliable measures of clinical performance competency and their potential use for credentialing and certification. As new simulation capabilities become available over the ensuing years, this work has successfully shown the way from development to changing practice.

4. SCOPE

The time has come for the scientific measurement of surgical skill to be incorporated into the apprenticeship model of surgical training. For surgical training in the United States generally, and for orthopedic trauma surgery in particular, the attainment of competence is certified primarily by the completion of a surgical residency, performance on standardized tests, review of sample cases, and evaluation of surgical logs. Surgical skills are not objectively or systematically assessed in the operating room. Resident skill attainment varies, partly because people learn different skills at different rates and in different ways but also because training opportunities are uneven. Patient census varies so that each resident is presented with different learning opportunities during each fixed-length rotation. It is unlikely that all residents have achieved competence in all skills in the same time interval, yet the march through residency advances: steady, inflexible, and relentless. De-emphasizing this time- and quantity-based paradigm and focusing more on individual resident competency requires a better definition of what skills are needed at each stage in residency and how these skills may be reliably measured. This will ensure that residents are prepared to handle the responsibilities they are given while ensuring patient safety.

Surgical competence is a collection of skill, knowledge, and judgment, incorporating both technical and nontechnical components. Operative competence is broader, including a surgeon's experience and more comprehensive understanding of a number of varied elements involved in leading surgery. A fundamental and pressing challenge preventing the movement towards competency-based education is establishing objective criteria for when a resident is ready to advance to the next level. Historically, the decision is based on time in training rather than on performance. Although there are clear milestones defining competencies expected of orthopedic residents, they are generally framed around subjective rather than objective assessment. Most such definitions inherently imply a minimum standard needed to be safe and/or the ability to perform the operation independently without help.

The use of simulators to train surgical skills has been increasing over the past two decades, but skill attainment on a simulator may not transfer to the operating room. The first step in linking simulator research to technical competence, the skill to perform the required steps of a particular procedure safely and successfully, depends on connecting simulator research to operating room performance. It is essential to define practical, sensitive, objective, and generalizable measures of surgical results that can differentiate experts from novices. Measuring surgical results well will drive evidence-based training, as training then becomes a challenge of producing surgeons who achieve quantifiably excellent results.

5. METHODS

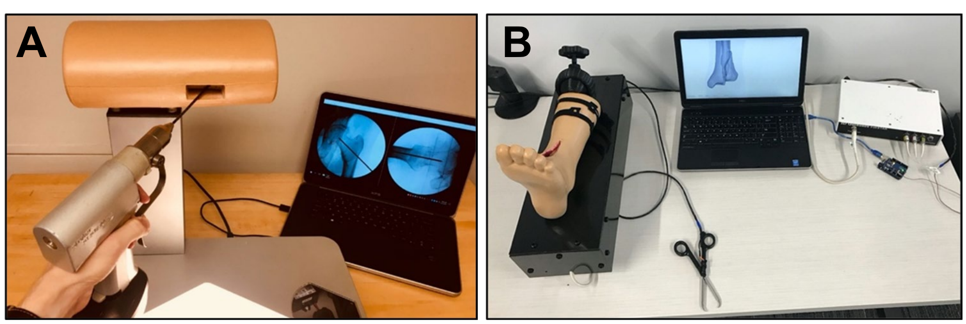


Figure 1. (A) The fluoroscopic wire navigation simulator allows trainees to drive a surgical wire into a plastic bone mounted on the mast and hidden by a soft tissue sleeve. Simulated fluoroscope images allow the trainee to practice guiding the wire precisely into the bone. (B) The articular fracture reduction simulator uses an electromagnetic system to provide real-time tracking of surgical tools and fracture models buried in a soft tissue sleeve.

<u>Fluoroscopic Wire Navigation Simulator Platform</u>. Utilizing prior AHRQ funding, we developed an orthopedic surgical simulation platform that replicates the look and feel of navigating a wire through bone using fluoroscopic guidance (Figure 1). Two differentiating features of the platform are: (1) camera-based tracking of a wire

replaces fluoroscopic radiation exposure, and (2) a foam bone surrogate replicates the feel of drilling through actual bone. We initially developed this platform to simulate the navigation of a wire in the treatment of intertrochanteric hip fractures. We have since shown that the platform is extensible to other orthopedic applications. A small company, led by members of our team, has licensed the technology to inexpensively produce the simulator. After several years of development and testing, this and related wire navigation simulator modules are currently used twice annually at fracture courses for residents.

Specific Aim 1: Measure orthopedic trauma surgery skill through surgical results

Meaningful Measures of Surgical Performance. In other prior AHRQ-funded work, we evaluated the transfer of training on the wire navigation simulator in both a mock OR and in the real OR. This allowed us to differentiate performance improvements caused by simulator training from those accruing with experience. For each specific procedure, we define consistent, clinically important measures common to both the simulator and the actual OR. For intertrochanteric hip fractures, residents were assessed in the mock OR based on their use of fluoroscopy, total time, and tip-apex distance (TAD – Figure 2). A large TAD, a measure of the distance between the tip of the wire and the apex of the femoral head, is a strong predictor of later implant failure. For this reason, a central surgical objective is to minimize the TAD, but it can come at the cost of substantial fluoroscopic radiation and greater operative time when skills are poorly developed. Residents who trained on the simulator had a lower TAD than those who did not ($p=0.001$), and their performance on the simulator (TAD, image use, overall time) correlated with performance in the mock OR.

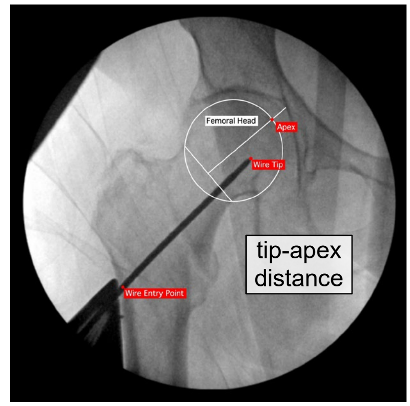


Figure 2. We measure clinically important parameters from fluoroscopic images acquired in surgery to evaluate surgical performance.

Defining Competence Using Expert Performance. Our studies with residents training at various venues have established baseline expectations for resident performance on these skills. To define competence, we assessed the performance of 30 Orthopedic Trauma Fellows (experts near the pinnacle of their surgical training in orthopedic trauma) at the OTA Trauma Fellows course in 2019. Figure 3 plots probability density function fits for normalized composite scores from three cohorts of surgeons: residents with no training, residents with simulator training, and OTA Fellows. These results convincingly suggest that simulator training can improve resident performance to a level approaching that of experts, at least on the wire navigation simulator.

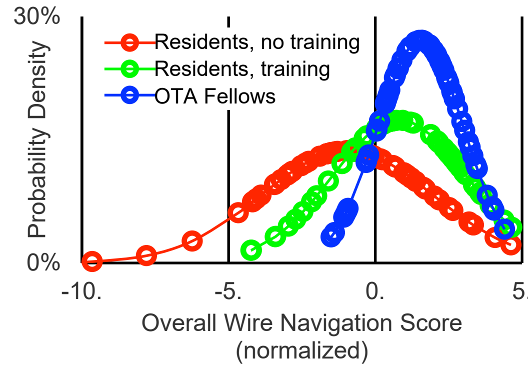


Figure 3. Normalized composite scores of three cohorts of surgeons placing a hip wire on the simulator indicate: performance correlates with experience and training can boost performance.

Image-Based Decision Error Analysis Scoring. In the course of our work defining performance metrics, we found limitations in the use of simple "final result" measures (e.g., the TAD) that do not account for in-the-moment decision making, a hallmark of performance. It is precisely during these moments that there is a unique opportunity to teach. We recently published work showing how surgical skill can be objectively measured from fluoroscopic images using an Image-based Decision Error Analysis (IDEA) Score. We evaluated performance both on our hip fracture wire navigation simulator and in the actual OR during fracture surgery. After examining fluoroscopic image sequences from 176 simulator trials and 21 surgical cases, we identified erroneous skill behaviors that could be measured by comparing wire adjustments made between consecutive images against the goal of targeting the apex of the femoral head. A composite IDEA score integrating decision errors and the final TAD was defined to produce a single overall performance metric, with constituent metrics first normalized to the overall population. A higher IDEA score shows better performance. A score of zero matches perfectly the mean performance across all subjects, and deviations from zero denote performance away from the mean. The IDEA scoring method was used to objectively analyze 37,000 images from the simulator and 688 from the OR (Figure 4). The number of decision errors correlated with fluoroscopic image count on both the simulator and in the OR (R^2=0.76; p<0.001 and R^2=0.71; p<0.001, respectively). Decision error counts did not correlate with the TAD for either the simulator or the OR, indicating that the TAD is independent of decision errors. The IDEA score correlated with surgical case experience in the OR (R^2=0.66; p<0.001 – Figure 5). Overlaid simulator scores of trainee populations from Figure 3 **further demonstrate the value of objectively assessing resident skill development directly from fluoroscopic image sequences**. The IDEA score provides a basis for assessing skill proficiency at key timepoints in residency, such as when rotating onto/off a new surgical service and before performing certain procedures in the OR, and as a tool for debriefing and providing feedback after a procedure is completed.

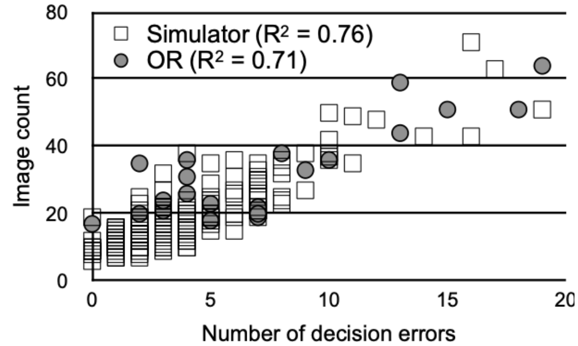


Figure 4. Decision errors highly correlated with image count (p<0.001), whether on the simulator or in the OR (i.e., amounts to increased radiation exposure).

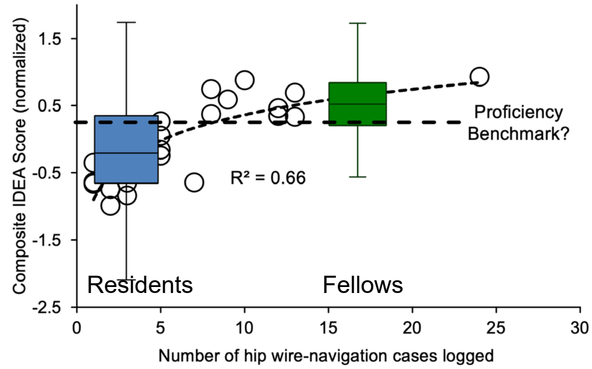


Figure 5. The composite IDEA score was highly correlated (p<0.001) with a resident's case experience.

This work showed how resident skill development can be objectively assessed from fluoroscopic image sequences. The IDEA scoring provides a basis for evaluating the competence of a resident. The score can be used to assess skill at key timepoints throughout residency, such as when rotating onto/off a new surgical service and before performing certain procedures in the OR, or also as a tool for debriefing/providing feedback after a procedure is completed. By setting a threshold to establish competency in wire navigation, an objective quantifiable basis for evaluation is made available.

Using Fluoroscopic Image Sequences to Measure OR Performance. Our prior experience has shown that important insights can be extracted from a surgical fluoroscopic image sequence, including details of surgical strategy, surgical skill, and the quantitative assessment of the final implant placement. One challenge in obtaining these insights involves collecting detailed measurements from as many as 150 images in a sequence for a long or complex surgery. Traditionally, surgeons may spend a half an hour or so analyzing the final one or two images of a sequence with whatever image analysis software they have available to assess the quality of the final implant placement, but relying on surgeons is cost prohibitive for exhaustively analyzing an entire sequence. Conventionally applied to assess the quality of implant placement based upon end-of-surgery images, assessors have refrained from analyzing the scores of images obtained throughout the operation due to prohibitive time and cost constraints.

To address this challenge, we pioneered a semi-autonomous process involving artificial intelligence and customized image analysis software (Figure 6). The software speeds the sequence analysis so that dozens of surgical images can be analyzed by a non-surgeon minimally trained analyst in minutes rather than days. The original software is tailored toward the analysis of fluoroscopic sequences of a dynamic hip screw implant. The dynamic hip screw analysis provides the TAD, the actual versus desired wire trajectory, the duration of the surgical step, the number of images collected, and the wire entry point, some metrics that have been used to assess

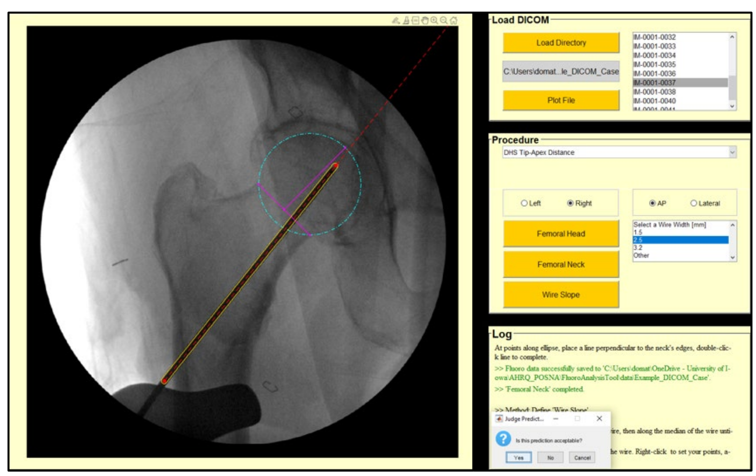


Figure 6. Illustration of the main interaction window of the custom tool developed using MATLAB. Left frame displays an annotated fluoroscopic image; right frame provides the image and annotation controls. A log in the lower right corner provides instructions, confirmations, and warnings.

fixation quality. Four expert and four novice analysts performed an assessment on the same fluoroscopic image sequence by measuring one objective metric on each image using the custom software, while repeating the task using commercial medical imaging analysis software. We found that analysts could measure the objective metric three times faster when using the custom software without any significant difference in the measurement due to experience level (Figure 7). The software reduces the analysis time from 2.75 minutes to 0.66 minutes per image. Details extracted from these images are useful in distinguishing expert from novice performance, especially when collecting a count of decision errors. These results suggest that a well-designed fluoroscopy analysis system can facilitate inexpensive, reliable, and objective assessment of surgical skills.

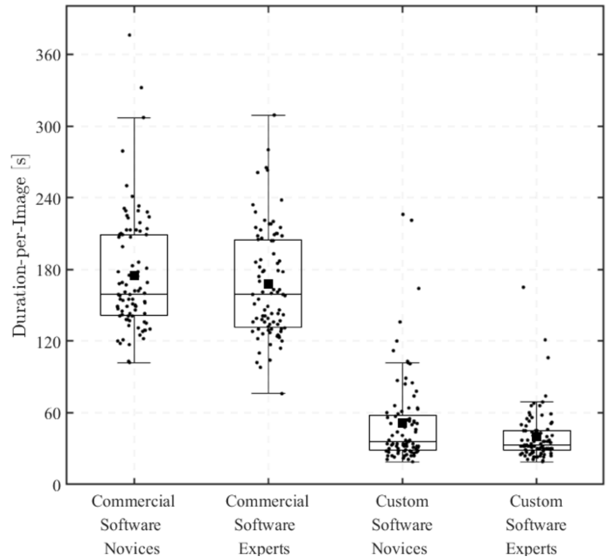


Figure 7. Comparisons of time required to complete evaluations using commercial vs. custom software for novice vs. expert evaluators.

Semi-Automated Objective Measurement Algorithm vs. Expensive Subjective Expert Rating. The current standard for evaluating OR performance, a supervising expert surgeon using subjective rating scales, is prone to bias, exhibits poor repeatability, and scales poorly for high-volume use. Recent advances seek to remedy these concerns, but feedback to performers remains limited by assessor subjectivity. Objective structured analysis of intra-operative fluoroscopic images offers a less expensive, more consistent, and precise alternative assessment strategy. A recent adaptation of the software analyzes the fixation of pediatric supracondylar humerus fractures.[96] We took an algorithmic approach to objectively quantifying fixation quality by measuring clinically relevant metrics from intra-operative fluoroscopic images. Image analysis and geometric relationships between the implants and fracture landmarks were used to quantify success in achieving fixation goals. To compare the algorithmic scoring with expert surgeon ratings, six fellowship-trained orthopedic surgeons independently rank-ordered fixation quality in 23 supracondylar humerus fracture cases based on fluoroscopic images obtained at completion of surgery. The inter-rater reliability of the six experts, measured by Cronbach's alpha, was 0.78 (Figure 8). The maximum difference between expert individual rankings for any case **varied between 5 and 22 ranking positions** (mean=12). The correlation of the objective scoring algorithm to the subjective expert mean rank was 0.78, suggesting an effective performance assessment tool that offers fewer barriers to providing surgeons quality feedback. The root-mean-squared error of the algorithm ranking with respect to the expert mean rank was 4.26, a lesser value than five of the six experts. The experts assessed skill with reasonable reliability. However, individual expert assessments substantially varied in some instances, confirming that the absence of a quantitative metric enabled notable disagreement on performance. A follow-up survey revealed that expert assessments were not very repeatable and that evaluators may welcome quantitative data. The algorithm-derived ranking agreed well with the mean rank of the six experts, suggesting an efficacious performance assessment tool for this surgical procedure that offers fewer barriers to providing surgeons quality feedback.

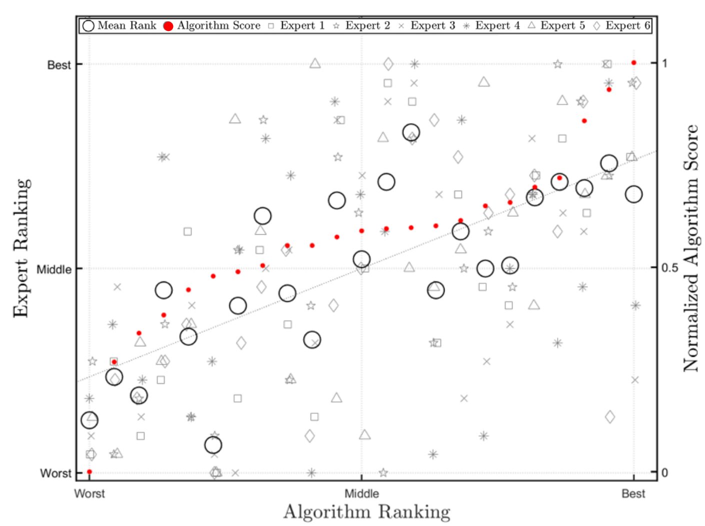


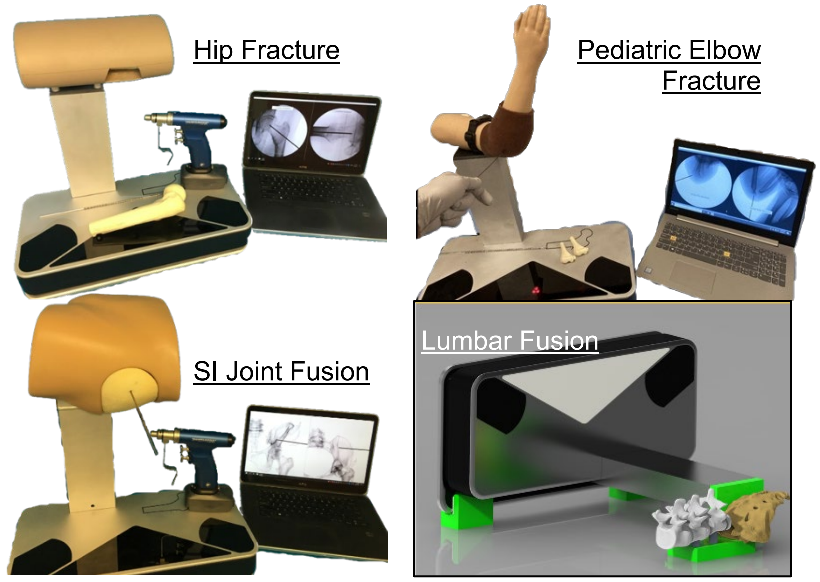
Figure 8. Objective Algorithm vs. Expert Ranking

Specific Aim 2: Develop simulator-based training programs that improve surgical results

Simulator Advancements. Over the duration of this AHRQ funding, we have extended our fluoroscopic wire navigation simulator platform, developed originally for training hip fracture pinning, to simulate a wider range of applications, each with unique skills challenges (Figure 9). These simulators continue to be refined, piloted, and disseminated in the MOSS Consortium, at other U.S. orthopedic residency programs, and at OTA fracture courses.

Figure 9. Fluoroscopic Wire Navigation Simulator Applications

- Hip Fracture (dynamic hip screw, intramedullary nail)
- Pediatric Supracondylar Humerus Fracture (three cross pins)
- Sacroiliac (SI) Joint Stabilization/Fusion (iliosacral screw)
- Lumbar Interbody Fusion (pedicle screws)



Pediatric Elbow Fracture Simulator. To highlight but one new application, the pediatric elbow (supracondylar humerus) fracture simulator was designed to tackle the use of smaller diameter wires, the tracking of multiple wires, and the higher degree of precision required because of the smaller bones and sensitivity to wire position. In the course of this project, we first established that performance on the simulator, which we assess based on objective metrics of the final wire placement and the number of fluoroscopic images used, aligns with surgical case experience, and then we showed that training on the simulator improves skill performance.

We first tested 18 surgeons at our institution, 15 orthopedic residents (4 PGY1, 5 PGY2, 6 PGY3) and three fellowship-trained staff surgeons. Each participant was instructed to place three diverging pins into the distal end of a Sawbones model. Performance was graded based on AP pin spread, lateral pin spread, mid-wire placement, and the total image count. We found that the AP pin spread increased with experience level and that significantly more PGY3 residents, already having logged comparable cases, scored above a clinical threshold of achieving an AP pin spread ratio of >1/3 (Figure 10). This goal is meaningful because the likelihood of loss of fracture reduction is much lower when the AP pin spread ratio is above this threshold. PGY3 residents spread their surgical wires to cover 40% (SD 8%) of the fracture line in the AP plane and used an average of 41 (SD 14) images. Staff surgeons spread their wires to cover an average 61% (standard deviation, SD, 13%) of the fracture line in the AP plane and used an average of 32 (SD 4.5) images to place the wires. The staff surgeons not only achieved a higher AP pin spread ($p<0.01$), well above 1/3, but they did so using fewer fluoroscopic images.

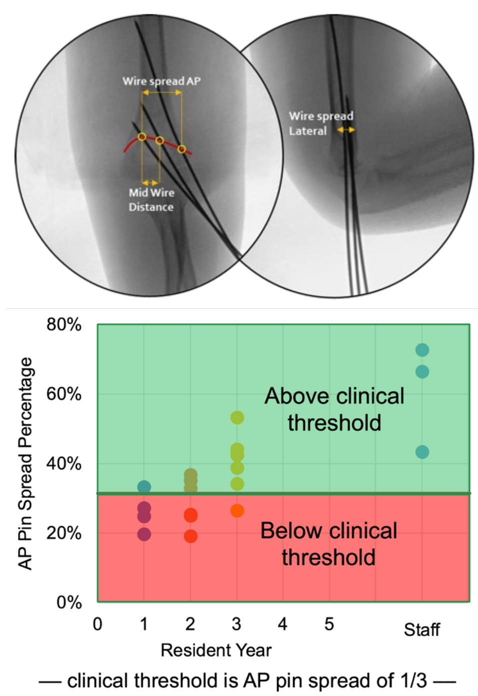


Figure 10. More experienced surgeons achieved greater pin spread on the pediatric elbow fracture simulator.

We next tested six PGY-2 residents, selected for training when they were still "pinning naïve." The pre-simulator training curriculum included an 8-minute video on the basics of supracondylar humerus fractures from epidemiology, classification, surgical indications, and pinning goals. Additionally, residents completed three online modules focused on identifying the fracture line on fluoroscopic images, selecting appropriate pin placement, and assessing pin constructs. Residents were then assessed and trained on the simulator. They first performed a pre-test in which they were asked to place three lateral, divergent, non-planar, bi-cortical wires into a pediatric distal humerus. Each resident then went through two training modules on the simulator, working on wire navigation skills and ideal pin placement with coaching of a senior level resident. A post-test was then performed in the same manner as the pre-test.

Table 1. Performance metrics for six residents in pre-test and post-test assessments.

Resident	AP Spread Ratio		Fluoroscopic Images		Normalized Composite Score	
	Pre-Test	Post-Test	Pre-Test	Post-Test	Pre-Test	Post-Test
1	0.25	0.58	49	41	-0.29	0.28
2	0.19	0.39	76	41	-1.66	0.43
3	0.33	0.37	38	26	-0.18	-0.16
4	0.37	0.71	44	29	-0.50	1.25
5	0.35	0.38	45	24	-0.49	-0.04
6	0.25	0.74	42	23	-0.74	0.72
Mean	**0.29**	**0.53**	**49**	**31**	**-0.64**	**0.41**
p -value	0.01		0.02		0.01	

The AP pin spread ratio, number of fluoro images, and normalized composite score showed statistically significant improvement following training (Table 1). The mean AP pin spread ratio improved from 0.29 to 0.53 (p=0.01). Three of six residents failed to achieve the goal AP spread ratio of >1/3 on the pre-test, whereas all residents successfully achieved this goal on the post-test. The mean number of fluoroscopic images improved from 49 to 31 images (p=0.02). The mean normalized composite score improved from -0.64 to 0.41 (p=0.01).

A Novel Simulator for Fracture Reduction Skills. As a complement to our wire navigation simulator, we recently developed a simulator to train and assess the skill of fracture reduction in pediatric elbow fractures (Figure 11). The standard of care for surgically treating these fractures is closed reduction prior to percutaneous pinning. The reduction maneuvers follow a basic step-by-step process for reducing the fragment. The first step is applying gradual, longitudinal traction to bring the distal fragment out to length. It is important to not be overly aggressive to not disrupt the intact posterior periosteum. The traction should be applied with the elbow flexed about 20°. Keeping the elbow slightly flexed attempts to avoid the possibility of stretching soft tissues including the brachial artery and median nerve across the proximal fragment.

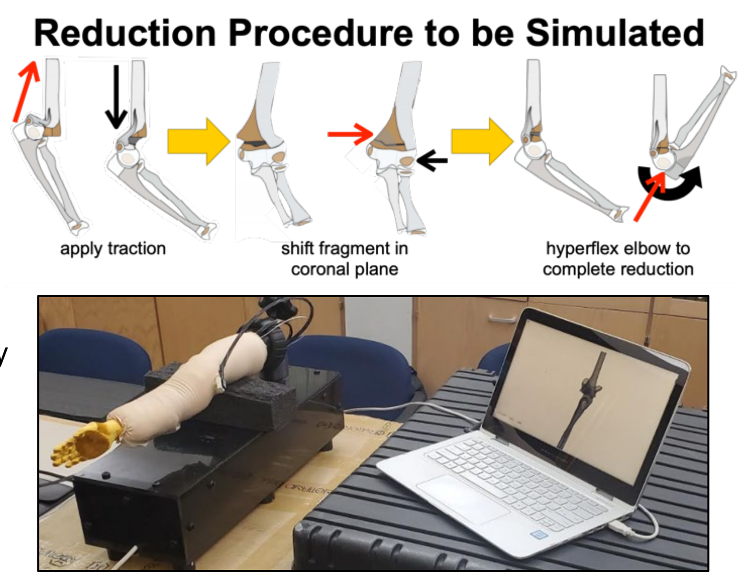


Figure 11. Fracture reduction simulator concept and implementation in a prototype form.

The pediatric supracondylar humerus fracture reduction simulator combines physical models of the tissues involved including bone, bone fragments, and soft tissue with computer-generated images like those obtained in the OR, but without the radiation. This is done using an electromagnetic tracking system that provides real-time tracking of surgical tools and artificial fractures buried in a soft tissue model. The system provides quantitative feedback on a number of performance metrics including novel features, such as a sumative snapshot of fracture displacement over time.

We performed a study to determine if the simulator can differentiate between novice and experienced surgeons, with six subjects in each study group. The experiment included the reduction of three unique fracture patterns performed by each of the participants. Experts significantly outperformed novices in the number of images taken (count 464 vs. 713, p<0.001), the number of resets of the displaced fragment (count 12 vs 44, p = p<0.001), the jerkiness of the forearm's tracking sensor (average 93 m/s^2 vs. 163 m/s^2, p=0.008), a final Euclidean distance error (average 9.83 mm vs. 11.4 mm, p=0.036), and the normalized distanced traveled of the distal fragment (average 4.62 vs. 8.29, p=0.001). These results strongly suggest that the simulator can differentiate between the two skill levels of the participants, providing evidence to include this simulator in the traditional surgical training curriculum.

Specific Aim 3: Define simulator-based assessments that generalize to clinical performance

Generalizability of Fluoroscopic Wire Navigation Skills Learned on the Simulator. As noted above, we previously showed that practice on the simulator leads to skill transfer with a mock OR with real fluoroscopy and life-size models, when the skill being tested is identical to that being trained (center-center wire placement in the femoral head). It remained to be established if this skill can generalize across surgical tasks. We tested if training on a femoral neck fracture simulation improves performance on placing a center-center wire for an intertrochanteric fracture.

This study took place at the 2019 Spring and Fall OTA Resident Comprehensive Fracture Courses. Across the two courses, 160 residents participated. Residents began by performing a baseline assessment, using the simulator to place a guide wire in a center-center position, trying to minimize their TAD, use of fluoroscopy, total time, and number of decision-making errors. Following the baseline assessment, the simulator software provided training related to wire placement in a femoral neck fracture surgery. In this task, residents were shown bubble targets for the specific three-wire configuration used to stabilize femoral neck fractures. Residents were given feedback as they performed the task; a projected line on the simulated fluoroscopic image helped show the residents their wire trajectory in bone. If they successfully hit their target, the animated bubble would pop, and the next wire target would be displayed. Following the training exercise, residents were asked to place another center-center wire, this time on a patient with a different anatomy than the earlier baseline assessment. The composite IDEA score was calculated for baseline and post-training assessments.

Of 160 residents that participated in the study, 129 completed the entire assessment and training protocol in the allotted time. At baseline, residents had an average TAD of 16 mm (SD 6.6mm), made 14.3 decision errors (SD 8.4), and had a composite IDEA score of -0.04 (SD 0.7). Following the training, residents had an average TAD of 14 mm (SD 5.5mm), made 9.2 decision errors (SD 5.4), and had a cumulative score of 0.41 (SD 0.6). Residents improved significantly after training on all performance measures ($p<0.05$). When looking at residents with a baseline score less than zero (arguably below acceptable performance levels), 91% of residents improved their IDEA score on average 1 SD of performance improvement.

This study shows that when residents receive training on a three-wire configuration to treat femoral neck fractures, they can apply those skills to placing a center-center guide wire. Because the fundamental skill of targeting a trajectory in bone applies to both procedures, residents were able to generalize and transfer their skills, reducing the number of decision errors, their TAD, fluoroscopy use, time, and composite IDEA score. Furthermore, the greatest improvement was seen by the residents who needed the training the most. The results suggest that wire navigation is a skill in which basic training can be applied to other core orthopedic procedures.

Rigor and Reproducibility: Throughout our work, we make every effort to ensure that our research approach is scientifically rigorous and that our results are robust and unbiased. Simulators that are developed undergo extensive testing to ensure that their performance characteristics meet high standards for accuracy and reproducibility. We report these characterizations of simulator performance in peer-reviewed publications. Whenever possible, evaluations of surgical performances are blinded as to the identities, institutions, and level of training of the resident/surgeon. Only later in the process when statistical comparisons are made do these data get re-introduced into our analytical pipeline. The reproducibility of measurement methodologies are routinely tested by repeated measures analyses and studies of inter-observer variability in the measurements.

6. RESULTS (please see above for details)

Simulator training was demonstrated to improve wire navigation skills, both on the simulator AND in the operating room. Performance improvements from simulator training were shown to be comparable to those from experience accrued in the operating room. Simulator performance scores indicative of competence were established, and we demonstrated that training with the simulator allowed residents to attain these scores. Finally, wire navigation skills developed while training on specific simulator tasks were generalizable to other wire navigation tasks.

7. LIST OF PUBLICATIONS AND PRODUCTS (Inclusive Over Project Duration)

Peer-Reviewed Papers

1. Taylor LK, Thomas GW, Karam MD, Kreiter CD, Anderson DD. Developing an objective assessment of surgical performance from operating room video and surgical imagery. IISE Transactions on Healthcare Systems Engineering. 2018;8(2):110–116. doi: 10.1080/24725579.2017.1418767. PMC6020852.
2. Taylor L, Thomas GW, Anderson DD, Marsh JL, Karam MD. Measures of hip fracture wire navigation performance in the operating room reflect surgical experience. Included in: American Orthopaedic Association's (AOA) Council of Orthopaedic Residency Directors (CORD) Summer 2017 Conference: Top Abstracts. J Bone Joint Surg Am. 2018;100(10):e68(1–6). doi: 10.2106/JBJS.18.00006.
3. Tatum M, Thomas GW, Anderson DD. Design of a percutaneous articular fracture reduction simulator. ASME Frontiers in Biomedical Devices, 2018 Design of Medical Devices Conference: V001T03A008. doi:10.1115/DMD2018-6852
4. Long SA, Thomas GW, Anderson DD. An extensible orthopaedic wire navigation simulation platform. J Med Device. 2019 Sep;13(3):031001-310017. doi: 10.1115/1.4043461. PMC6677394.
5. Atesok K, Hurwitz S, Anderson DD, Satava R, Thomas GW, Tufescu T, Heffernan MJ, Papavassiliou E, Theiss S, Marsh JL. Advancing simulation-based orthopaedic surgical skills training: An analysis of the challenges to implementation. Advanc Orthop. 2019;2586034;2019:7. doi: 10.1155/2019/2586034. PMC6745149.
6. Long SA, Thomas G, Karam MD, Anderson DD. Do skills acquired from training with a wire navigation simulator transfer to a mock operating room environment? Clin Orthop Relat Res. 2019;477(10):2189–2198. doi: 10.1097/CORR.0000000000000799. [**highlighted in Editor's Spotlight**.] PMC6999951.
7. Karam MD, Long S, Marsh JL, Anderson DD, Thomas GW. Looking beyond the tip apex distance to assess wire navigation performance. Included in: American Orthopaedic Association's (AOA) Council of Orthopaedic Residency Directors (CORD) 2018 Summer Conference: Top Abstracts. J Bone Joint Surg Am. 2020;102(2):e5(1– 6). http://dx.doi.org/10.2106/JBJS.19.00959.
8. Johnson Z, Thomas GW, Long S, Anderson DD. A pediatric supracondylar humerus fracture wire navigation simulator. ASME Frontiers in Biomedical Devices, Proc 2020 Design of Medical Devices Conference. Minneapolis, MN, USA, April 6–9, 2020. V001T07A002. doi.org/10.1115/DMD2020-9031.
9. Thomas GW, Long S, Tatum M, Kowalewski T, Mattioli D, Marsh JL, Kowalski HR, Karam MD, Bechtold JE, Anderson DD. A vision for using simulation and virtual coaching to improve the community practice of orthopedic trauma surgery. Iowa Orthop J. 2020. 40(1):25-34. PMID: 32742205; PMC7368513.
10. Long S, Thomas GW, Karam MD, Marsh JL, Anderson DD. Surgical skill can be objectively measured from fluoroscopic images using a novel Image-based Decision Error Analysis (IDEA) Score. Clin Orthop Relat Res. 2021 Jun 1;479(6):1386–1394. doi.org/10.1097/CORR.0000000000001623. PMC8133282.
11. Williams E, Long S, Tatum M, Anderson DD, Thomas GW. Designing a 3D printed bone simulant for wire navigation training. ASME Frontiers in Biomedical Devices, Proc 2021 Design of Medical Devices Conference. 2021 Design of Medical Devices Conference. Minneapolis, MN, USA, April 12–15, 2021. V001T01A002. ASME. doi.org/10.1115/DMD2021-1034.
12. Hill J, Long S, Williams E, Thomas GW, Bechtold J, Anderson DD. Developing a wire navigation simulator for pedicle screw placement in minimally invasive transforaminal lumbar interbody fusion. ASME Frontiers in Biomedical Devices, Proc 2021 Design of Medical Devices

Conference. Minneapolis, MN, USA, April 12–15, 2021. V001T10A004. ASME. doi.org/10.1115/DMD2021-1052.

13. Karam MD, Anderson DD, Thomas GW, Long SA. Trauma fellows set wire navigation competency benchmarks. Included in: American Orthopaedic Association's (AOA) Council of Orthopaedic Residency Directors (CORD) 2020 Summer Conference: Top Abstracts. J Bone Joint Surg Am. 2022;104(1):e1.
14. Karam MD, Long S, Connor E, Anderson DD, Thomas GW, Kowalski H. Testing the validity of a novel simulator for training pediatric supracondylar elbow fractures. Included in: American Orthopaedic Association's (AOA) Council of Orthopaedic Residency Directors (CORD) 2021 Summer Conference: Top 10 Abstracts. JBJS Open Access. 2022;7(3):e22.00049.
15. Hill JJ, Gulbrandsen TR, Wynn MS, Anderson DD, Thomas GW, Marsh JL, Karam M. Failure of orthopedic residents to voluntarily participate in laboratory skills training. J Am Acad Orthop Surg. 2022 Feb 15;30(4):161-167. doi.org/10.5435/JAAOS-D-21-00680. PMID: 34910714.
16. Mattioli DD, Thomas GW, Long SA, Tatum M, Anderson DD. Minimally trained analysts can perform fast, objective assessment of orthopedic surgical skill using fluoroscopic images. IISE Trans Healthc Syst Eng. 2022. 12(3):212-20. doi.org/10.1080/24725579.2022.2035022. PMC9488091.
17. Holt JB, Long SA, Tatum M, Thomas GW, Anderson DD, Karam MD, Kowalski HR. Integrating simulation for developing pediatric supracondylar humeral fracture reduction and fixation skills into an orthopaedic surgery residency program. J POSNA. 2022;4(S1). doi.org/10.55275/JPOSNA-2022-0070
18. Holt JB, Long SA, Tatum M, Thomas GW, Anderson DD, Karam MD, Kowalski HR. Integrating simulation for developing slipped capital femoral epiphysis fixation skills into an orthopaedic surgery residency program. J POSNA. 2022;4(S1). doi.org/10.55275/JPOSNA-2022-525
19. Tatum M, Long S, Anderson DD, Thomas G. Creating a fracture reduction and wire navigation simulator for orthopaedic skills training and assessment. ASME Frontiers in Biomedical Devices, Proc 2023 Design Med Devices Conf. Minneapolis, MN. April 17–19, 2023. V001T06A004. doi.org/10.1115/DMD2023-4204
20. Tatum M, Thomas GW, Anderson DD. Design and evaluation of a system for CT-free volume reconstruction from intra-operative fluoroscopy for navigation in orthopedic surgery. ASME Frontiers in Biomedical Devices, Proc 2023 Design Med Devices Conf. Minneapolis, MN. April 17–19, 2023. V001T10A007. doi.org/10.1115/DMD2023-9640
21. Williams E, Thomas G, Long S, Anderson DD, Karam MD. Discovering patterns in orthopedic surgical resident behavior during a cephalomedullary nail procedure with a wire navigation simulator. ASME Frontiers in Biomedical Devices, Proc 2023 Design Med Devices Conf. Minneapolis, MN. April 17–19, 2023. V001T06A002. doi.org/10.1115/DMD2023-4016
22. Rölfing JD, Salling LB, Long SA, Vogt B, Anderson DD, Thomas GW, Jensen RD. Establishing construct validity of a novel simulator for guide wire navigation in antegrade femoral intramedullary nailing. Iowa Orthop J. 2023. 43(1):31-5. PMC10296486.
23. Thomas GW, Long SA, Kurtzhals T, Connor E, Anderson DD, Karam M, Kowalski H. A dedicated simulator training curriculum improves resident performance in surgical management of pediatric supracondylar humerus fractures. JBJS Open Access. 2023. 8(3):e23.00031.PMC10489481
24. Mattioli DD, Thomas GW, Long SA, Rölfing JD, Anderson DD. Fluoroscopic image-based behavior analysis can objectively explain subjective expert assessment of wire navigation skill. J Orthop Res. 2023 Aug 31. doi: 10.1002/jor.25685. [Epub ahead of print] PMID: 37652571.

Theses/Dissertations

1. Long SA. Orthopaedic surgical skills: Examining how we train and measure performance in wire navigation tasks. Ph.D. Thesis, 2019, Biomedical Engineering, The University of Iowa.

2. Tatum M. Simulating articular fracture reduction for surgical training. M.S. Thesis, 2019, Biomedical Engineering, The University of Iowa.
3. Johnson Z. The design and verification of a pediatric supracondylar humerus fracture reduction simulator. M.S. Thesis, 2021, Industrial & Systems Engineering, The University of Iowa.
4. Mattioli DD. Orthopedic data science: Fluoroscopic image analysis for the objective assessment of technical skill. Ph.D. Thesis, 2021, Industrial & Systems Engineering, The University of Iowa.
5. Kurtzhals T. The effect of training on surgical resident skill acquisition. M.S. Thesis, 2022, Industrial & Systems Engineering, The University of Iowa.
6. Williams E. Fluoroscopy use analysis of cephalomedullary nail procedures performed by surgical residents. M.S. Thesis, 2022, Industrial & Systems Engineering, The University of Iowa.