Summer 2013

Surgical simulation training models for orthopaedic fracture surgery

Gary Thomas Ohrt
University of Iowa

SURGICAL SIMULATION TRAINING MODELS FOR ORTHOPAEDIC FRACTURE SURGERY

by

Gary Thomas Ohrt

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

August 2013

Thesis Supervisor: Associate Professor Donald D. Anderson

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

_____

MASTER'S THESIS

_____

This is to certify that the Master's thesis of

Gary Thomas Ohrt

has been approved by the Examining Committee
for the thesis requirement for the Master of Science
degree in Biomedical Engineering at the August 2013 graduation.

Thesis Committee: _____
Donald D. Anderson, Thesis Supervisor

_____
Thomas D. Brown

_____
Geb W. Thomas

_____
Matthew D. Karam

To Keri, my loving fiancée, who has put up with me through the late nights and crankiness

ACKNOWLEDGMENTS

I would first like to thank my advisor, Dr. Don Anderson. Without your guidance and support, none of the work I accomplished would have been possible. I appreciate you taking a chance on me as an undergrad, and believing that I would be a worthy student to continue the success that graduate students have had in the Lab. Thank you for all the hours spent teaching me, and the patience when it took multiple tries for me to get it right.

To the rest of my committee, Dr. Thomas Brown, Dr. Matthew Karam, and Dr. Geb Thomas, thank you for taking the time out of your busy lives to help me along the way. Dr. Brown, thank you for allowing me to be a part of your lab, and learn from you. Dr. Karam, thank you for providing the funding that allowed me to do my research, and always being willing to provide a surgeon's perspective. Dr. Thomas, thank you for allowing me to use your equipment to make the simulator a success, and for all the ideas and methodologies from a human factors & learning perspective.

I would also like to thank Andrew Kern, for taking the time to help me with MATLAB and the DEA algorithm. I would also like to thank the rest of the Biomechanics Lab for all the help along the way. I would also like to thank my family for their support during my academic journey.

ABSTRACT

Articular fracture reduction is a complex surgical task that requires surgeons to be competent at multiple surgical skills to successfully complete. The list of skills needed includes the ability to use fluoroscopic images to build a 3D mental model of the fracture during reconstruction, the proper handling and use of surgical instruments, how to manipulate the fracture fragment into a reduced configuration with minimal hand motion, proper k-wire placement, and the preservation of surrounding soft tissues. Current training methodology is based on an apprenticeship model. The resident learns by watching a senior surgeon, and then preforms the procedure on live patients under the guidance of the senior surgeon to gain competence. This endangers the patient and does not provide the best outcome for either patient or resident.

The work presented in this thesis is the early development of an articular fracture reduction simulator, the subsequent use of the simulator in the training of orthopaedic residents, and assessment of the improvement of residents after practice on the simulator. To date, the simulator has been tested on four different groups of residents,3 different groups from the University of Iowa and one group from the University of Minnesota. Considerable effort has been made to validate the improvement seen in resident performance through objective means. The Objective Structured Assessment of Technical Skills (OSATS) is a global rating score and procedural checklist that has been previously validated to objectively measure surgical skill. Other assessment metrics include hand motion capture to count the number of discrete actions and measure distance traveled during the surgical procedure, fluoroscopic usage and radiation exposure, articular ‘step-off’, the surface deviation from an intact or ideal reconstruction, and contact stress exposure.

The results indicate that the goals for the simulator have been met, that the simulator provides a means of training orthopaedic residents, assessing improvement,

decreased the cost of training, and improved patient safety. The simulator is not without limitations including sample size, and radiation exposure. The task being trained is complex and can be broken down into basic subtasks that could be trained individually. Even with flaws, the simulator is an improvement over current training methods and is an excellent first step toward creating a surgical skills curriculum to comply with new mandates from orthopaedic surgery's governing bodies.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

CHAPTER 1- INTRODUCTION

1.1 Project Background

This thesis describes work that builds upon prior development effort in the University of Iowa Orthopaedic Biomechanics Laboratory, led by Thaddeus Thomas, PhD and Andrew Kern, MS during their graduate studies. Dr. Thomas led the team that developed an articular fracture (tibial plafond) reduction simulator in early 2010 to establish the value of pre-operative planning based upon a computational three-dimensional (3D) puzzle solving approach. Meanwhile, Mr. Kern adapted Discrete Element Analysis (DEA) software routines for the purpose of assessing contact stress exposures in ankle fractures following surgical reconstruction of the tibial plafond, based upon CT-derived 3D models of each patient's actual reconstructed ankle.

My first introduction to the fracture reduction simulator was in September 2010, at which time I assisted Dr. Thomas in running five senior (4^{th} year and 5^{th} year) residents through the simulation. I performed low level data collection and analysis for Dr. Thomas to include in his PhD dissertation. In late November 2010, Dr. Thomas and I ran seven junior (1^{st} year and 2^{nd} year) residents through the first generation simulator. I performed all data analysis on the junior residents and began comparing the results with the previously collected senior resident data.

The value of the simulator as a standalone teaching tool became apparent as data analysis progressed. A second generation of the fracture reduction simulator was developed primarily as a platform for teaching surgical skills, and the simulation was run multiple times in November and December of 2011. In the course of that work, it became clear that simple whole-joint-surface summary measures of the degree of residual surface incongruity were inadequate outcome indicators of a resident's performance in precisely reducing the fracture. Building upon Mr. Kern's work, a decision was made to shift from incongruity measures to estimates of contact stress exposure provided by DEA. With Mr.

Kern's assistance, I modified the software to be able to work using the reconstructed fracture constellations obtained from simulation.

1.2 The Need for Surgical Skills Training

Many fractures can be managed using well-established treatments such as external fixators, casts, orthoses, surgical plates and screws, and intramedullary nails. The treatment choice depends primarily on the anatomic location of the fracture, the number of fragments, and the geometry and displacement of fragments. Re-establishing alignment, restoring the joint surface, and ensuring stable fixation during the period needed to achieve bony union between the fragments are the primary goals of articular fracture management.

In low-energy fractures with minimal comminution and fragment displacement, the well-established treatments listed above generally provide satisfactory patient outcomes. In high-energy fracture scenarios such as a fall from a height or a high-speed car accident, however, the resulting degree of comminution and fragment displacement present unique challenges. In these situations a highly precise surgical restoration of the articular surface is generally required to provide these patients the most favorable outcome. Obtaining that precise joint restoration and being able to maintain it are two important related surgical challenges faced in treating orthopaedic trauma.

Due to the high rate of long-term morbidity and cost associated with comminuted fractures that extend into articular joints, educating surgical residents on the proper reduction techniques is extremely important [1]. Clinical evidence indicates that correct fracture reduction is vital, not only for the bone healing process, but it also plays a critical role in avoiding cartilage contact stress elevations [2, 3]. Intra-articular fractures fare better when the original anatomy is recreated with generous inter-fragmentary contact [4]. It is also important for the original anatomy to be restored with minimal surgical insult to the soft tissue [5-7]. Although the tolerance of the ankle joint (the focus of this

simulation work) to residual displacement is not well defined, in other joints even as little as 1 mm of residual displacement can cause poor outcomes [5, 8].

The trauma surgeon reconstructing a highly comminuted fracture faces a problem that is similar to solving a 3D puzzle, albeit with additional complexities. Traditional surgical techniques have involved using an open reduction in which the surgeon exposes the fragments by surgical approach through a large incision in the soft tissue. This approach introduces increased risk for failure of wound healing, infection, joint stiffness, delayed fracture healing, and further articular surface damage. To mitigate these risks, an alternative is to use a limited exposure approach, where the fragments are manipulated percutaneously through a smaller incision (≈5 cm) under fluoroscopic guidance. However, the obstructed visibility and poor accessibility can lead to imperfect reductions, a likely cause of post-traumatic osteoarthritis [9, 10].

Residents in orthopaedics (trainees) must acquire the basic requisite skills to perform such surgeries during a five-year training period. The mastering of these skills comes only after repeated practice or execution of the surgeries. Recent work hour restrictions mandated by the America Council on Graduate Medical Education (ACGME) reduce the number of surgical case exposures that residents can use to acquire the skills. The recognition of the need for surgical skills training outside the operating room has led to the mandating of new regulations for the education of first year residents [11, 12]. The simulator described in this thesis provides a means to teach the surgical skills needed in the limited approach technique (the ability to visualize the location of a 3D fragment from 2D fluoroscopic images, to properly use surgical tools, and to efficiently manipulate fragments while minimizing insult to the surrounding soft tissues) outside of the operating room. In this respect the simulator has great potential to improve patient safety and outcomes, as the surgical residents are able to acquire the skills without making trial-and-error mistakes on live patients.

By the same token, there is a legitimate risk that simulation can teach improper techniques and inappropriate strategies, in which case use of the simulation technologies can actually be a detriment to education. Therefore considerable effort is warranted in establishing the validity of a newly developed simulator prior to its use for teaching orthopaedic residents.

While simulation will never wholly replace the operating room as the definitive venue for honing surgical skill, it does open the door for resident trainees to more tightly focus their scarce learning time in the operating room. The work described in this thesis includes early development efforts in articular fracture reduction simulation, subsequent use of the simulation technologies in teaching orthopaedic residents, assessment of the simulation and the residents, refinements of the simulator, and the best efforts to date at establishing the validity of the simulation.

CHAPTER 2 – LITERATURE REVIEW

2.1 History of Surgical Simulation

2.1.1. Origins of Simulation in Orthopaedics

The use of simulation in skills training is not a new or novel concept. The approach has been perhaps most notably successful in the training of aircraft pilots and nuclear reactor technicians. The idea of surgical skills simulation has been around since the early 1950's [13]. However, not much was done in simulating surgical conditions and procedures until the 1980's [14, 15]. Interest in surgical simulation greatly increased in the early 1990's with the advent of more powerful computers and a rise in the use of limited incision surgeries such as arthroscopy and endoscopy [16-23]. In 1999, the role of simulation in surgical education became a hot topic in medical education [24-26]. As of 2004, the FDA requires training and certification on a validated simulator before physicians can place carotid stents [27].

Orthopaedic surgery has been slower to embrace simulation as a tool for surgical skills acquisition. In 1995, Ziegler et al reported on the development of an early arthroscopic simulator [28]. The American Board of Orthopaedic Surgery provided funding for the development of a knee arthroscopic surgery simulator in 1997. The simulator included the following elements: (1) haptics, (2) a knee mannequin that could be manipulated by the trainee, (3) realistic imaging, and (4) realistic instrumentation [29]. Advances in the realm of craniofacial surgical reconstruction simulation also served as a catalyst in orthopaedic surgery simulation [30-39]. The majority of these simulations involved creating a physical model of the fracture from digital medical images and using the model in pre-surgical planning or practice [30-34, 37, 38]. The use of simulators has recently been mandated as part of training and assessment methods at the national level for laparoscopy and endoscopy [40-44].

2.1.2 The Rationale for Surgical Simulation

The Flexner report of 1910 describes the medical education process as a mixture of didactic lessons coupled with apprentice-based experiences [45]. Despite the substantial passage of time, this same medical education process is still practiced today. This type of a learning environment has several disadvantages: (1) experience and training is not standard among different learners, (2) teaching objectives are dependent on the educator and tend to be ad hoc, and (3) the amount and variety of apprenticeship based learning opportunities is limited [46]. Michelson outlined the usefulness that education by simulation could provide in orthopaedic surgery for addressing these shortcomings.[46]. The use of actors simulating patients with structured interactions during step 2 of the National Board examinations for medical students has established standardized teaching, learning, and assessment goals within medical education [46, 47].

The use of simulators allows (1) for the trainee to learn in a purely educational environment where they are not worried about clinical concerns for mistakes, (2) for the criteria for success to be clearly defined and objectively measured, (3) for the ability to progress up the learning curve without affecting patient safety, and (4) for an optimized clinical experience in interactions between the resident and the patient [46, 48].

Michelson described the different levels of simulation education as at the lowest level teaching a specific task (such as performing an intubation), at the next level introducing decision making into the simulation (e.g., choosing the correct instrument to perform a specific function during a surgical procedure), next adding in multiple levels of decision making, and at the highest level including outside influences on the simulation (such as another person making a mistake during the training exercise) [46]. Three levels of validity with sublevels are described by Michelson [46]. The first level of validity ‘Content’ is where the knowledge given in the simulation is compared against the knowledge needed to perform the real activity. The second level is ‘Criterion’ validity, how well the simulation compares to independent measures of performance, and can be

predictive or concurrent. The third level is 'Construct' validity or how well the simulation measures against the concepts being taught or tested; meaning does it place experts and novices in different groups and do most novices place in the same group[46]?

The first rigorous assessment of simulator validity in orthopaedics was reported in 2002[49]. Pedowitz et al. used an arthroscopic simulator of the shoulder, in which the subjects demonstrated construct validity as completion time and economy of motion correlated to experience levels between medical students, residents, and attending physicians [49]. Strom et al were able to demonstrate a significant training improvement with medical students after only an hour of practice [50]. In a different study, Strom issued a precautionary note that the skills taught by a simulator may be too specific, in a comparison of performance on a knee arthroscopic simulator between an untrained group of medical students versus a group that had trained using laparoscopic and shoulder simulators showed no difference in skill[51].

The findings of these earlier studies underscore the need for simulators to not only train for the specific task simulated, but also to teach skills that can be generalized across different procedures. Two specific simulators have factored prominently within surgical skills training; the McGill Inanimate System for Training and Evaluation of Laparoscopic Skills (MISTELS) and the Imperial College Surgical Assessment Device (ICSAD)[52, 53], both of which were developed to assess laparoscopic skills[41, 54-57]. Both have been shown to have excellent face, content and construct validity [41, 53-64].

While the MISTELS was able to show competence differences between novices and experts, Fraser et al. found that choosing a scoring cutoff would lead to passing incompetent surgeons and failing competent surgeons, and that a compromise between the two was necessary[57]. Fraser et al. also found that not all of the original subtasks of the MISTELS system were necessary, and they were able to still differentiate competence while not testing on the two most expensive tasks[57]. The MISTELS system was also able to show that with an objective scoring metric, inter-rater reliability was 0.998 at a

95% confidence interval (CI), and intra-rater reliability was 0.892 at a 95% CI [55]. During testing of anesthesiologists of different skills levels, the ISCAD was able to differentiate between novice residents and experienced residents by motion distance, and novice residents and staff anesthesiologists were differentiated by number of movements and time [52]. Scores on a global rating scale and procedural checklists also differentiated between skill levels and correlated with the ISCAD outcomes [52].

2.2 Assessment of Clinical Performance

Current methodology for evaluating surgical skill and clinical performance of residents involves case logs and review, observation by board certified surgeons, and cadaveric procedures which involve no objective metrics, only a surgeon subjective review at completion. There is a need to have a direct assessment of surgical skills. In non-surgical fields, use of a checklist and rank scoring system has been adopted and standardized everywhere from saber metrics in baseball to pilot certification.

The Objective Structured Assessment of Technical Skills, or OSATS, is a method for evaluating surgical performance using two parts: a global rating scale and a task specific checklist [65-67]. An expert surgeon evaluates the participant either directly or under video review [65-67]. The global rating includes between five and ten surgical behaviors such as respect of tissue, economy of motion, and correct use of surgical instruments [65-67]. These surgical behaviors represent the collective knowledge from experts in this task regarding what constitutes good surgical technique.

The OSATS scoring metric has been shown to match the validity and reliability of the Objective Structured Clinical Examination (OSCE) [67-70]. The OSCE is a standard of assessment in clinical competence[71], used in many fields in health sciences, such as medicine, dentistry, pharmacy, nursing, and radiography. The OSATS scores have also shown good concordance with scores from the motion capture system ICSAD, and the MISTELS laparoscopic evaluation and training system [61, 67, 72, 73].

2.3 Motion Tracking Systems and Validation

The use of a system to track motion is not a novel or a recent development. The rotoscope was utilized around 1915 to animate motion. Live action film was projected on to a light table, thereby allowing an artist to trace the image onto paper for realistic motion for cartoon characters, such as Betty Boop and Snow White [74, 75]. The technology has advanced considerably since those early days, to where now highly sophisticated motion capture systems are routinely used to generate advanced computer visualization for movies and television [75].

Two technologies for passive live motion capture currently predominate: optical marker tracking and electromagnetic tracking [76, 77]. Passive motion capture systems are more robust in an uncontrolled environment[76]. Optical tracking systems use light to measure and record movement through reflection and triangulation [76]. Markers may or may not be used to enhance reflectivity. Optical systems have the advantages of not placing heavy equipment onto the object or person being tracked and have relatively accurate sensing[78]. Optical systems need to maintain line of sight with the object being tracked and can suffer from noise and distortion in environments with high intensity of the type of light being used for tracking. Electromagnetic tracking systems use sensors in a generated magnetic field to track motion. Distortions of the field and field fluctuations can cause noise and introduce error in the measurements[78]. However, there is no requirement for line of sight between the electromagnetic sensors and the magnetic field source for tracking purposes [79].

Two motion capture systems have been used in the research presented in this thesis: a Qualisys Oqus camera system and a Polhemus G4 electromagnetic tracking system. The Qualisys Oqus system uses infrared light reflected off of passive markers to track motion[80]. The cameras are placed into position and calibrated using a known orientation of four markers on a plate, and a free floating T shaped wand with a known length of the cross of the T, with markers placed on each end[80]. Using the Qualisys

Track Manager (QTM) software, a calibration (mapping) of the spatial volume of interest is created by waving the wand around the area[80]. The Oqus camera system used has a maximum frame rate of 500 fps with full field of view,[80] with an accuracy error of <0.5mm for tests performed in the Buckwalter Surgical Skills Laboratory where the simulations were mostly performed. Motion capture is started within QTM and is recorded for a set time or until a stop command is given. Relevant markers are then identified and extraneous markers are discarded, and then post processing can be done to produce measures of the position, velocity or acceleration of the markers[80].

The Polhemus G4 system generates an electromagnetic field. Magnetic calibrated sensors are placed in the field; the voltages generated in the sensors by the field are then converted into reading of position and orientation within the Polhemus tracking software[79]. The G4 system has a maximum frame rate of 120 fps [79]. Preliminary experiments in the simulation area within the Buckwalter Surgical Skills Laboratory showed an accuracy error of as low as 1mm when in Polhemus's recommended field range of 6 inches to 6 feet from the source. Data for location were collected using inches as the unit of measure, which is the default unit for the G4 system. The G4 system has no extra markers that need to be identified before post processing produces measures of position, velocity, and acceleration of the markers, but the data were converted into millimeters before processing.

Qualisys motion capture systems have been used in kinematic evaluations since 1993[81]. The University of Salford uses the Qualisys Oqus camera system to track human gait[82]. The Oqus system has been used to track motion during weight lifting[83]. Qualisys Oqus camera systems have also been used to track hand motion during function tasks [84, 85]. Liebermann was able to use hand motion data to show a difference in the motion of stroke vs. non-stroke patients based on distance traveled and speed[85]. Qualisys systems have also been used to track arm and hand motion during

badminton smashes[86], hand gestures for building a gesture recognition database[87], and tracking dart throwers motions during daily living [88].

Polhemus developed electromagnetic sensors in the late 1970's to track the rocking motions of the blind[89-91]. This led to the development of the Polhemus company and their line of motion tracking systems. One of Polhemus's modern tracking systems is the G4 motion tracking system used in the research presented in this thesis. The patent for using coordinate transforms in order to track and determine orientation of an object was issued in 1976 [92]. The patent for the electromagnetic field generator that is used in the Polhemus systems was issued in 1977 [93]. The Polhemus motion capture systems have been used in a wide variety of motion capture applications, from avionics, virtual reality, art creativity and cultural preservation, to medical applications such as image guided surgery, biomechanics and kinematics motion tracking, and training simulations[79]. The ICSAD uses a Polhemus IsotrakII motion capture system to track hand motion during hand movements [53, 61, 63, 64].

In 2001, researchers were able to show (using the ICSAD) that the number of hand movements decreased with skill during a small bowel anastomosis procedure and during a vein patch insertion procedure [61]. Hand motion tracking has also been used to differentiate between novice and experts during an ultrasound-guided supraclavicular block [93]. Howells was able to show that hand motion analysis is a valid method for measuring skill in arthroscopy [94]. Yamaguchi found that experienced surgeons had shorter path lengths for hand motion during laparoscopic skills exercises [95].

2.4 A Functional Outcome Metric of Surgical Repair Precision in Articular Fractures

2.4.1 Contact Stress Evaluation and Clinical Relevance

Post-traumatic osteoarthritis (PTOA) is a manifestation of osteoarthritis (OA) following a joint injury. PTOA onset and development depend upon mechanical factors

such as the acute fracture severity, residual joint incongruity, and joint instability [5, 94-96]. Studies have shown that the mechanical factors of fracture severity and surgical reduction quality are closely related [97]. Joints that are subjected to higher energy impacts during fracture events have a larger amount of articular comminution, which in turn makes it more difficult to successfully reduce the fracture[97].

For present purposes, the skills training simulation work is focused on restoring articular congruity to reduce chronic aberrant contact stress. Contact stress can increase significantly with incongruities in the articular surface [9]. A simple step-off of 3mm was shown to cause a 75% increase in the contact stress in cadaveric human tibia [9]. Brown et al also observed an inverse relationship between the thickness of cartilage and measured contact stress [9]. This would indicate that a single step-off measurement may not be an effective measure of successful reduction, as fractures with similar levels of incongruity may have different contact stress and have different risk for joint degeneration[5].

By contrast, direct assessment of the resultant contact stress associated with a given articular reduction could provide considerable insight regarding the clinical outcome. Computational stress analysis methods offer a means whereby contact stress at the articular surface may be accurately estimated. Hadley et al. showed that a contact stress-time metric of >10MPa-years correlated with poor long term outcome of congenital dislocation of the hip (CDH) following a closed reduction, introducing the idea of a contact stress-time exposure metric [98]. The outcome of hips with less than 10 MPa-years was satisfactory in 81% of the patients studied[98]. Maxian et al. used two different computer modeling techniques for calculating contact stress on the Hadley 1990 patient database to compute contact stress overdose [99]. Reasonable correlation between the patient outcome and the overdose was found for both models, albeit with each model having a different stress overdose threshold due to modeling assumptions[99]. Anderson et al. demonstrated a relationship between increased contact stress time exposure during

gait and PTOA in fractured distal tibias [2]. By accounting for the differences between the Anderson study and the Hadley and Maxian studies, it was shown that the contact stress thresholds for damage were relatively consistent [2].

2.4.2 Finite Element Analysis

Finite Element analysis (FEA) is the standard of practice for computational stress analysis in musculoskeletal biomechanics, as well as in other fields. FEA was introduced to the field of biomechanics in 1972 [100]. It was first used to determine the internal stresses present in skeletal models[100]. This was a revolution as the methodology was able to compute stress in the complex and unique structures of human bone[101].

FEA follows from the subdividing of 2D or 3D geometry into many smaller blocks (finite elements). This subdivision, or discretization, is called meshing. Each finite element is then assigned parameters such as material properties and boundary conditions (forces or displacements) that are specific to the item being meshed. A computer then assembles and solves the system of partial differential equations governing force/moment equilibrium at nodal interfaces[101]. With the dividing of geometry in FEA, complex structural analysis problems are converted from being intractable into a system of fairly straightforward and solvable equations. In many situations the accuracy of the FEA solution can readily match real world phenomena. Because of advances in computer technology, larger, more complex models can now be computed with relative ease.

Yet, one particularly difficult task in FEA remains the computation of contact between different bodies [102]. Contact presents a computational challenge because of the inherent discontinuity and nonlinearity of contact between two bodies [103]. Due to the contact forces, boundary conditions, external forces, and resultant deformations, the interaction of contact and deformations of the models cause increased run times, numerical convergence issues, and increased time investment by the analyst[102]. Because of the irregular nature of human structures, specifically bone, the requirement of

relative uniformity in element size necessitates extra knowledge and time input from the analyst, because the size and shape of elements can dramatically affect the quality of the final contact solution [104]. In addition, irregularities in the mesh or sharp incongruities within the model can cause major numerical convergence issues [102, 105]. Because of these issues in FEA, the methodology has been deemed by some to be unsuited for use in larger population-based studies (involving >20 subjects) [106].

The original precedent work in this area involved a finite element analysis study where 11 patient-specific post-op models were created and evaluated for contact stress over 13 instants from the gait cycle[2, 10, 107, 108]. A chronic contact stress-time exposure metric was used as a direct functional outcome measure [2], and it was compared against minimum two year follow-up Kellgren-Lawrence (KL) scores. The KL score is a clinical radiographic measurement of OA indicators with score ranging from 0 (no OA) to 4 (severe OA) [109]. The contact stress-time exposure metric demonstrated a 100% concordance with OA onset (KL≥2) and a 95% concordance with KL score [2].

2.4.3 Discrete Element Analysis

Given some of the disadvantages of FEA in solving articular joint contact problems, researchers have turned to an alternative computational stress analysis methodology known as Discrete Element Analysis (DEA) [94]. DEA holds attraction for solving a specific type of elastic contact problems [110]. DEA involves a simplified set of computations compared with FEA, but nonetheless it has been shown to reliably provide accurate contact stress estimates for articular joints [111]. A number of 2D and 3D DEA algorithms have been utilized in orthopaedic biomechanics for contact between both natural and artificial surfaces [106, 112-117]. DEA has primarily been used for solving joint contact forces in larger multi-body dynamic or quasi-static simulations. To a smaller degree, the algorithms have also been used for direct quantification of contact stress in joints [118].

Because of the simplicity of DEA when compared to FEA, many groups have created purpose-built implementations to model a specific joint or task[114]. DEA algorithms treat articular cartilage as an array of compressive-only springs connected to the underlying bone surfaces, which are treated as rigid [119]. Load or displacement is then applied to one or more of the bodies in the model and the resulting deformation of the springs is used to compute the contact stress between the bodies [116]. This is done iteratively, using a matrix of spring stiffnesses that are computed to solve using minimum energy principle and find the contact stress and displacements [119].

Blankevoort et al. 1991 created an implementation of deformable contact DEA with the assumption that the contacting surfaces are isotropic, linear-elastic, and bonded to a rigid surface [113]. The theory states that springs are spread over the contact surface, with the springs treated as the elastic deformable surface. Each spring is considered to be independent so that pressure on one spring does not affect pressure on any other spring, simplifying the contact and reducing the computation time and complexity[118].

Three key assumptions are made in DEA theory: (1) the contact area is large compared to cartilage thickness, (2) the cartilage is an isotropic linear elastic material, and (3) the subchondral bone is rigid [113]. Computed spring deformations are converted to contact stress using a linear elastic spring equation. The spring deformations are computed using a penetration analysis, in which overlap between the two contact surfaces are indicative of surface deformations [114, 120]. Different implementations have incorporated a proximity-based analysis in which the contact surface is implicit and the distance between underlying (rigid) subchondral bone surfaces are calculated [116, 117, 119, 121]. The theory has been expanded to use other spring models for large strain, small strain, and use of nonlinear surface material models to represent native cartilage [106, 114, 120]. The DEA algorithm requires the following as inputs: (1) two aligned and triangulated cartilage surfaces with corresponding thickness values for each facet, (2) desired loads, (3) displacements, (4) termination constraints, and (5) desired kinematics

[94]. More information on the exact procedure for this DEA algorithm can be found in the methods section, as it was adapted for use in computing the stress values of the simulation results.

2.5 Summary

Although simulation has started to gain traction within the realm of orthopaedic training, it is a relatively new practice with considerable opportunity for growth of new simulators and improvement of the currently available simulations. Compared with other medical disciplines, orthopaedic surgery in general, and orthopaedic trauma specifically, have lagged behind. General surgery now relies upon laparoscopic simulators not only for training residents, but also in assessment and board certification. Although hand surgery [122] and arthroscopy [123] have implemented surgical skills training simulations within orthopaedics, there has been no validated orthopaedic trauma surgical simulator, beyond the work presented in this thesis. This is unfortunate, since within orthopaedic trauma patients usually do not have the ability to pick their surgeon; the patients depend on whoever is on call and expect the surgeon to have the necessary skills to successfully care for the patient's injuries. This is a troubling notion, because the ability to perform orthopaedic trauma surgery is learned by resident trainees in programs and practiced throughout the country without public assurance of even minimal procedural competence.

Orthopaedic residents are not the only surgeons whose competency needs to be assessed within orthopaedic trauma. Currently, trauma call is shared by practicing orthopaedic surgeons for whom trauma surgery is performed infrequently. The most important and widely accepted method to assess competence is through board certification. The American Board of Orthopaedic Surgery (ABOS) mandates recertification every ten years and is continuously evolving the maintenance program for certification. In the process of board certification and recertification, the orthopaedic

surgeons' competency in medical knowledge, professionalism, communication, and judgment are assessed, yet surgical skill can only be assessed by radiographic review and case discussion; procedural skill is not directly assessed to even guarantee minimal levels of competency. Recently the ABOS, the orthopaedic residency review committee, and the American Academy of Orthopaedic Surgeons released new guidelines for the graduate education of orthopaedic residents in skills training[12]. The new guidelines mandate formal requirements for 1st year residents to undergo basic surgical skills training within orthopaedics outside the operating room [12]. This mandate has spurred the development and use of surgical simulators such as the ones detailed within this thesis.

CHAPTER 3 – METHODS AND MATERIALS

The skills training simulation technologies at the heart of this work have evolved through multiple design generations in the course of performing the work detailed in this thesis. Several different test subject groups were run through the simulator at different points in its evolution to aid in design improvements. As experience with the simulation was gained, improvements were made in the simulator itself, as well as in associated assessment technologies. Overarching design goals were (1) to train surgical residents on techniques that improve surgical articular fracture reduction and transfer to different orthopaedic treatments, (2) to decrease the costs associated with training surgical residents, (3) to discover and refine metrics for surgical performance assessment, and (4) to indirectly improve patient safety by improving surgical performance.

3.1 First Generation Tibial Plafond Fracture Simulator

The first generation of the tibial plafond fracture reduction simulator was developed as a way to establish the value of a virtual fracture reconstruction process[48, 124]. The main component was a lower extremity soft tissue-mimicking housing manufactured by Sawbones (Model #1518 Sawbones Inc., Vashon WA, USA) with a fractured tibia surrogate insert consisting of three fragments made of a polyurethane foam doped with barium sulfate ($BaSO_4$), which adequately replicates the appearance on radiography of a native tibia[125]. The housing had a longitudinal opening on the posterior aspect that extended from the heel cord to its most proximal extension, and the opening could be secured with a Velcro strap. An anterior surgical opening was cut at the tibio-talar joint, by a senior traumatologist.

The fracture models were created using distal tibia replicates fabricated from polyurethane foam. In order to create the fracture pattern that was used in the simulation, a human distal tibia replica was machined [126, 127] from high-density (640 kg/m^3) polyurethane foam (Grade FR7140 Last-A-Foam, General Plastics, Tacoma, WA) and

impacted by an artificial talus cast from Poly (methyl methacrylate) [PMMA]. Two different fracture configurations were produced. The specific foam was selected because its structural and fragmentation behavior were roughly comparable to that of human bone [125]. The impact was delivered in a drop tower with a 7.5 kg mass released from 0.5 m height [124]. The tibial fracture surrogate was produced as three fragments molded separately. The surrogate fragment pieces were cast in a mold made of silicone rubber (TinSil® 70 Series RTV, U.S. Composites.). Each surrogate was cast using 16lb/ft^3 density polyurethane foam doped with $BaSO_4$ at 15% by weight. To better replicate the radiographic appearance of the subchondral plate, the fragments' articular surfaces were painted using glue doped with 40% by weight $BaSO_4$.

The original fracture fragments were aligned by a senior orthopaedic traumatologist to represent a clinically realistic configuration (figure 1), and a template for placing the fragments reproducibly in this configuration was made from Poly (methyl methacrylate) PMMA. Figure 2 illustrates the two fracture constellations. Each surrogate tibia had the three fragments placed into the PMMA template, and the fragments were then glued together. The deviations between the fragment configurations produced this way were sub millimeter [124], as confirmed using a laser scanner (NextEngine Model 2020i; NextEngine, Inc., Santa Monica, CA). The low strength glue was placed only at a select few interfragmentary points, to ensure ready separation when manipulated by surgical techniques. Between simulation trials, new tibia fragment surrogates were exchanged into the soft tissue housing for the old from the posterior in order to not affect the anterior surgical opening. The access was approximately 5 cm in length and located anterior to follow a limited approach surgical technique. In order to prevent lengthening of the opening, colored tape was wrapped around the foot and ankle housing. The assembly was placed into an acrylic sleeve at the proximal end with a sand bag placed on top of the sleeve to prevent the foot and lower leg from moving freely (figure 3), while

still allowing for motion comparable to that which would be naturally associated with a leg attached to a patient in surgery.

Performance on the surgical simulation was assessed using several technologies. Video was recorded using a digital camera capturing at 29.97 frames per second from a single wide-angle view. Hand motions were tracked using an infrared optical tracking system designed and built by Qualisys AB using 4 Oqus 300 series camera and a laptop PC to record the data (Qualisys AB, Gothenburg Sweden)(figure 3). Motion data were captured at 100Hz. The video and hand motions were synchronized using timestamps, and the video was used to identify and discard any irrelevant gestures, through human removal of superfluous data. The motion was tracked using a custom apparatus that attached to the back of each hand by double sided tape. Each apparatus had four reflective fiducial markers locked in a known arrangement on extensions that moved the markers away from the back of the hand by 1 inch.

The primary objective in the simulation was to stabilize and reduce the fracture so that all fragments were rigidly attached to the base fragment. Participants were given up to 30 minutes to successfully reduce the fracture configuration, with full credit given if completed with 15 minutes and decreasing credit until 30 minutes was reached. Timing was from a start command until the participant verbally declared reduction completed or the time limit was reached. The reduction was performed with a limited approach using standard reduction tools (tenaculums, hemostats, self-retaining retractors, etc.), and the fragments were stabilized with Kirschner wires (k-wires). Secondary objectives were to minimize the time to completion, hand motion distance, the number of discrete hand motions, and the number of fluoroscopic images that were used.

After completion the reduced and fixated surrogates were removed from the soft tissue housing and visually graded with a Pass/Fail for whether the tibia fragments were stabilized, with failure being one or more fragments not secured to the base fragment (the large fragment that contains the proximal tibia). The reconstructed tibias were later laser

scanned; the reduction accuracy was quantified by comparing the reduced fragments to a native ideal reconstruction, which was mapped by using the intact geometry and registration using an Iterative Closest Point algorithm (ICP) in Geomagic Qualify (3D Systems Inc., Rock Hill, SC). Due to the importance of restoring articular surface congruity, the articular surface deviation ('step-off') after reduction was assessed. OSATS [65, 128] scoring was performed by an expert grader reviewing the video (figures 4 and 5). Motion data were processed in MATLAB using custom scripts.

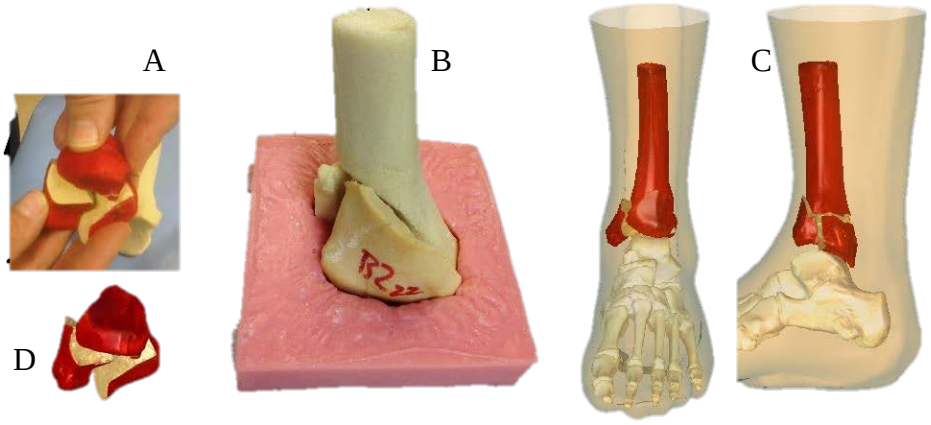

Figure 1: Fracture constellation created using drop tower
A: Surgeon holding fragment in realistic displaced condition waiting for glue to be applied
B: Fracture constellation in PMMA template used to hold cast fragment in identical position for gluing.
C. Digital placement of fracture into soft tissue housing AP and LM view
D: Digital view of articular surface in displaced configuration and the digital recreation of fracture constellation

Source: Yehyawi 2012 [48]. Modified and used with permission.

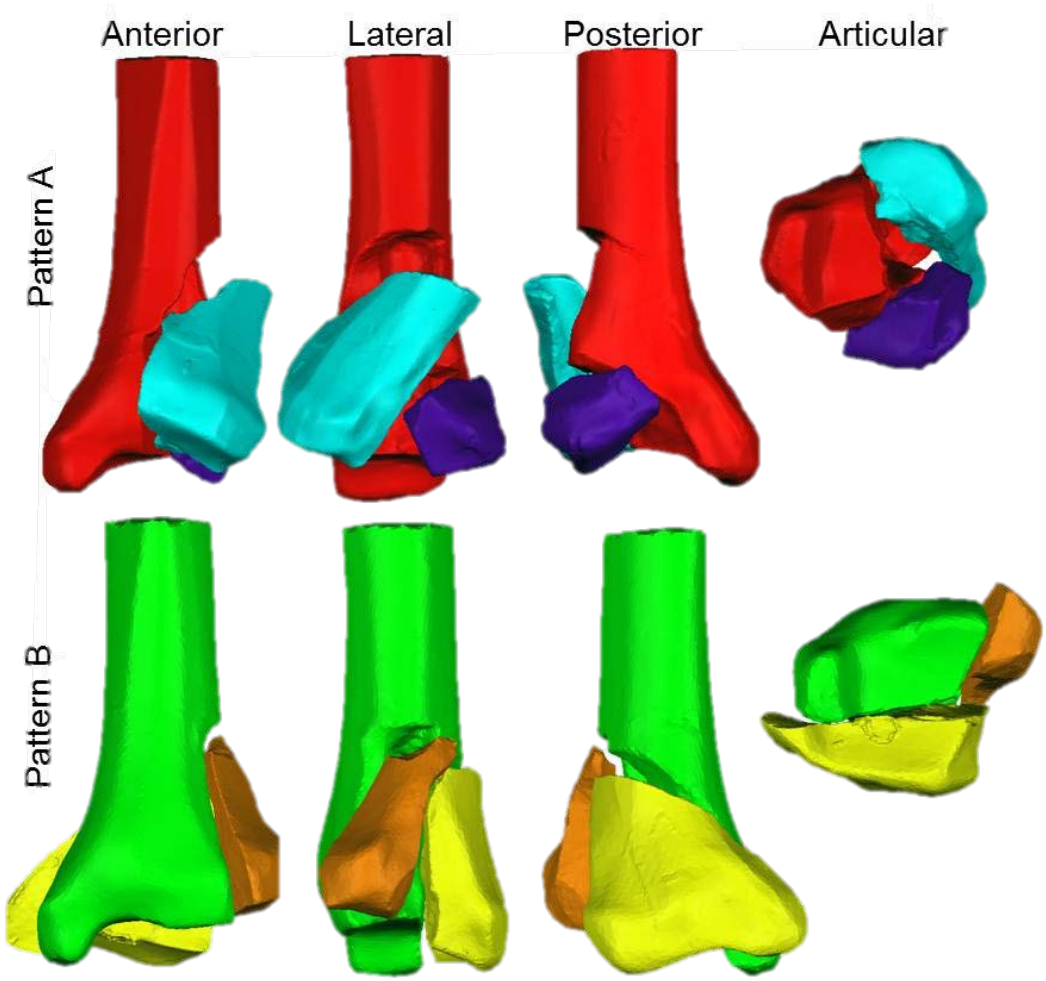

Figure 2: The two fracture constellations with views detailing the exact positioning of each fragment (different colors)

Note: Pattern A fragments are numbered as follows: Fragment 1-red base fragment, Fragment 2-purple posterior fragment, Fragment 3 is teal anterior-lateral fragment

Note: Pattern B fragments are numbered as follows: Fragment 1-green base fragment, Fragment 2-yeallow posterior lateral fragment, Fragment 3-brown anterior-lateral fragment

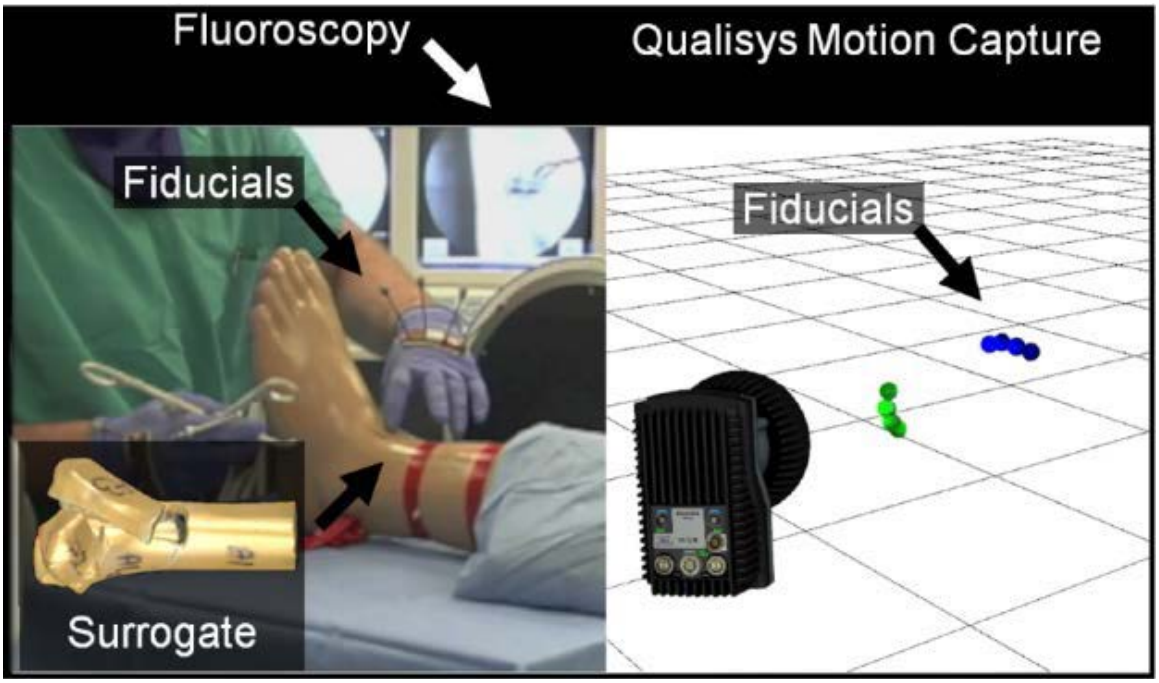

Figure 3: The tibial plafond simulation in the earliest form. Single camera feed shown on left with view of hand motion and fluoroscopic monitor. Qualisys motion capture software displaying the synced motion capture data with each hand's fiducial markers. In lower left, a fracture constellation is shown from the lateral view of the lower left leg.

Source: Yehyawi et al 2012 [48]. Used with permission.

Expert Assessment: Procedural Checklist

Please mark the box corresponding to the candidate's performance during the procedure

Reviewer: _____ Candidate Video Number: _____

Procedural Step	Performed	Not Performed
Appropriate identification of landmarks and incision		
Appropriate selection of surgical instruments		
Appropriate handling of surgical instruments		
Appropriate preparation/manipulation of fracture fragments		
Appropriate placement of reduction devices/aids		
Appropriate reduction achieved		
Appropriate use of fluoroscopy		
Appropriate placement of fixation devices		

Figure 4: OSATS procedure checklist used during simulations

The group of participants that used the original fracture reduction simulator included five senior (4^{th} or 5^{th} Post graduate year (PGY)) and seven junior (1^{st} or 2^{nd} PGY) orthopaedic surgery residents from the University of Iowa. This population was chosen due to the ease of access and their willingness to participate in this educational exercise. In order to more closely replicate the clinical situation, the residents were provided standard pre- and intra-operative imaging: plain radiographs, axial CT images, and fluoroscopy.

Expert Assessment: Global Rating Scale

Please circle the number corresponding to the candidates performance in each category

Reviewer: _____ Candidate Video Number: _____

	1	2	3	4	5
Preparation for Procedure	Did not organize equipment well. Has to stop procedure frequently to prepare equipment.		Equipment generally organized. Occasionally has to stop and prepare items.		All equipment neatly prepared and ready for use.
Respect for Tissue	Frequently used unnecessary force on tissue or caused damage.		Careful handling of tissue but occasionally caused inadvertent damage.		Consistently handled tissues appropriately with minimal damage.
Time and Motion	Many unnecessary moves.		Efficient time/motion, but some unnecessary moves.		Clear economy of movement and maximum efficiency.
Instrument Handling	Repeatedly makes tentative or awkward moves with instruments.		Competent use of instruments, but occasionally appeared stiff or awkward.		Fluid moves with instruments and no awkwardness.
Use of Fluoroscopy	Frequently requested excess images. Requested images with repeated improper orientation.		Used images effectively, but occasionally had to ask for extra images due to error.		Successfully completed procedure with minimal images. Used proper orientations during image capture.
Use of Kirschner Wires	Consistently placed wires in poorly placed positions or used in excess.		Good use of wires with occasional repositioning.		Efficient use of wires with appropriate placement.
Flow of Procedure	Frequently stopped procedure and seemed unsure of next moves.		Demonstrated some forward planning with reasonable progression of procedure.		Obviously planned course of procedure with effortless flow from one move to the next.
Knowledge of Procedure	Deficient knowledge.		Knew all important steps.		Demonstrated familiarity with all aspects of the procedure.
Overall Performance	Very Poor		Competent		Clearly Superior
Overall, should the candidate:	Pass / Fail				

Figure 5: OSATS rating score sheet used during all simulations. Contains nine categories on which performance is rated as well as a Pass/Fail criterion

There were a number of limitations in the first generation of the simulator. A single camera did not provide an adequate understanding of what was occurring within

the soft tissue housing, as no direct view was obtained. This made the OSATS scoring difficult. We also did not save or measure fluoroscopic data and usage. The posterior opening in the soft tissue housing prompted some participants to try to use that window in the course of the simulation. The sandbag and sleeve which constrained the ankle model proximally allowed the participants to move the lower extremity around the table in manners that would not be possible in the operating room (OR). However this did not present an issue once participants were verbally reminded during the simulation to only move the model as they would a real leg. The attachment method for the markers was also somewhat problematic, as the double sided tape occasionally failed mid procedure. When this happened the participant had to stop so that the markers could be reattached. Motion capture data were also lost when the markers traveled outside of the line of sight for at least two Oqus cameras. Data were also corrupted when the Qualisys Oqus camera failed to transfer its on-board memory to the laptop hard drive, resulting in no motion data for trials in which this error occurred.

3.2 Second Generation Tibial Plafond Fracture Simulator

The second generation of the tibial plafond fracture simulator retained the use of the $BaSO_4$ doped polyurethane tibia surrogates in a Sawbones housing[129]. Motion capture continued to involve using a Qualisys Oqus four camera system[129]. Video data were also captured from multiple angles including a head mounted camera. The video was consolidated into a single composite split-screen view that included head mounted view, wide angle view, overhead view, and a feed from the fluoroscopy machine shown in figure 6.

The attachment method for the fiducial markers was changed from hot glue on the gloves to double sided tape and rubber banding between the fingers of the participant and around the extensions for the markers. We found that the tape did not provide greater attachment than the rubber bands alone and so we removed that after a few participants.

One participant who had both rubber bands and tape on one hand and only rubber bands on the other remarked that he preferred only the rubber bands as it was less invasive and more comfortable.

Figure 6: 2011 Simulation Video Capture. Top Left: Head camera video feed, Top Right: Overhead mounted camera feed, Bottom left: wide angle camera feed, Bottom Right: C-Arm fluoroscopic image feed.

Source: Karam et al 2012 [129]. Used with permission.

Participants were randomized into a control or intervention group with 3 PGY1 and 3 PGY2 residents in each group, for a total of 12 participants. Each participant performed the simulation twice, one month apart, with the intervention group receiving

training between simulations. The participants were assessed on time-to-completion, objective reduction of the articular surface ('step-off') using a post-hoc 3d laser scan compared to the idealized reconstruction used in the 1st generation simulator, discrete hand motions, cumulative hand motion distance, radiation dose and fluoroscopy time. Faster time, less step-off, fewer discrete hand motions and less distance, and lower dose and fluoroscopy time were all judged to reflect better performance. OSATS [65, 128] scoring was performed live by an unblinded senior traumatologist.

The intervention group received two types of training the week of the second simulation: a cognitive module, which was implemented through the University of Iowa's online course system, and a motor skills training session. The cognitive module consisted of a pretest of basic knowledge, a presentation of general knowledge related to plafond fractures, fluoroscopy, and online video performance reviews. The motor skills training consisted of a brief didactic introduction into the training objectives. Surgical instrumentation, reduction, and fixation techniques were taught in a group setting. A review of the participants' first simulation video with one-on-one coaching by a senior traumatologist in a private, non-adversarial manner was given. The traumatologist had previously reviewed and documented the performance and custom tailored the coaching based on the participants strengths and weaknesses[129]. Immediately prior to the second simulation, the participants of the intervention group were given a guided practice similar to the real simulation except that a different fracture pattern was used and no fixation was performed. The control group was given unguided practice. After completing the practice, the tools were reset and the second simulation was set up with the test fracture configuration placed in the soft tissue housing. The second trial occurred exactly as the first did with respect to room set-up, personnel, metrics, and live OSATS scoring.

Some limitations of this system were that the camera system was bulky and low resolution, fluoroscopy was captured using an analog video signal and a video splitter for

the visual outputs, radiation data were hand collected, and the sandbag system was used again. The Qualisys motion capture system limitations were still present.

3.3 Third Generation Tibial Plafond Fracture Simulator

The third and most recent generation of the simulator incorporated several additional design changes. There were three major changes, the first being the complete production of the fracture pattern (to specifications) by Sawbones and accompanying changes that were required. The change to the fracture model now produced by Sawbones is that it uses a zinc paint coating to replicate radiographic properties rather than the $BaSO_4$ doping. Sawbones was provided with digital models of our fracture fragments individually and in the fractured configuration, which allowed them to create molds for mass manufacturing of the fragments. Sawbones molds and then uses a spray painting technique to cover the entire fragment in the zinc coating. In order to better replicate the subchondral plate, the articular surface of the fragments are sprayed with multiple layers of the zinc coating, which then better matches the radiographic appearance in the human ankle according to multiple senior orthopaedic traumatologists. The fragments are then placed into a mold and glued together using low strength glue. The fragment constellation is then wrapped in a thin clear flexible plastic sheet, which mimics the periosteum. This sheet is also used to prevent the fragments from sticking to the foam that is subsequently poured to replicate the surrounding soft tissue.

The plastic-wrapped constellation is loaded into a fixture along with the other bone replicas that make up the foot and lower leg, and the soft tissue foam is poured around the bone replicas. The manufacturing process then enters Sawbones' standard method for producing intact bone models. This process ensures that simulation participants cannot enter through the posterior aspect of the ankle and are forced to use the given surgical opening, which is marked out by a senior traumatologist immediately prior to simulation, using a template to ensure a constant size opening of 6cm-7cm,

depending on how the participant performed the cut. Due to difficulties in extracting the bone surrogates from the model because of the lack of a rear opening in the soft tissue mimicking housing, the reduced fractures are no longer laser scanned; instead they are scanned in an experimental Weight Bearing CT (WBCT) machine with voxel resolution of 0.37mm x 0.37mm x 0.37mm. The bones are then segmented from the CT scans using a combination of automated methods with manual clean-up and accuracy checking. These segmentations are then used to register the individual fragment digital models with the WBCT data, at which point the models enter the workflow for measuring articular surface deviation and/or DEA.

The second change is that a new digital camera system was implemented using 2 HD cameras to capture 1080P 30FPS footage from overhead and wide angle. A GoPro (GoPro Hero 3 Silver Edition & Head Mount, Woodman Labs Inc., USA) camera was mounted in a commercially available head mount and used to record what the participants were viewing during the simulation. Fluoroscopic data and images were captured digitally using the export mechanism on the different fluoroscopy equipment and a standard USB flash drive. Video between the devices was synchronized by instructing the participants to perform a double hand clap at the beginning and ending of the simulation. Videos were compiled and fluoroscopic images were spliced post hoc using commercially available software (iMovie; Apple Inc., Cupertino, CA, USA). The video was then used for the coaching and was made available for participants to view at their leisure once coaching had been performed. Video from the head camera was primarily used for this purpose, with aspects of the wide and overhead shot used when the head camera was obscured.

The third change involved a switch in the motion capture technology. The Qualisys motion capture system was replaced by a Polhemus G4 motion tracking system (Polhemus, Colchester, VT, USA). Instead of using optics the G4 uses an electromagnetic field source and sensors that respond to the field. Three sensors were used: one on the

back of each hand secured inside of a latex/nitrile glove and one that was secured using Velcro onto the back of the head mount of the GoPro camera. The head motion was not analyzed, as the sensor could not be adequately secured.

To date, the third generation tibial plafond fracture reduction simulator has been used to study skills acquisition in 6 University of Iowa PGY1 orthopaedic residents, 7 University of Minnesota PGY1 orthopaedic residents, and 8 University of Minnesota PGY2 orthopaedic residents. The University of Iowa residents performed the simulation 3 times during the month of January 2013, and the University of Minnesota residents performed the simulation twice, first during March 2013 and again in April 2013.

Iowa residents' first two trials occurred on the same day with approximately three hours between attempts, and a third trial was done two weeks later. In between the first and second trial, training was performed on a visual skills and mental model-building box trainer, that involved the use of two camera located orthogonally, and required the residents to perform tasks using the views that the cameras provided live or by still frames. Tasks included placing pins into holes drilled in an acrylic box, aligning a shaft with marking on the still images, and locating points in space. The dedicated training between that day and the third trial was one-on-one coaching and didactic information similar to the previous iteration's training. Also during the month, skills training in other areas was performed as a pilot program for the new mandate required by the orthopaedic board for the education of first year residents, which may have affected the results of the third attempt at the simulation[130].

The Minnesota residents were split into two groups, one of which received the coaching before the second simulation, while the other did not receive it until after the second simulation. Both groups were trained with the box trainer immediately following the first trial of the simulation in March. During the Minnesota trials, the sand bag model stabilization was abandoned for use of bone foam and surgical positioning gel packs. This better secured the ankle model into anatomically possible positions. The gel packs held

the leg so that it could be easily rotated around the long axis of the tibia, but could not be inverted so that the foot could be placed posterior side up. It likewise prevented motions that would over-torque the knee and hip.

Limitations in this generation of the simulator mostly involved battery issues. The Go Pro camera battery life was cut short by enabling the Wi-Fi live streaming option, which resulted in lost video data for one subject during which the battery died mid simulation. This was rectified by purchasing multiple batteries and only using the live stream to properly position the camera to match the subject's field of view and then turning off the live stream. The G4 system's wireless sensor hub has an internal battery that is non-removable, and the battery indicator is obscured when placed onto a participant. When the battery died with no warning, it resulted in three of the eighteen trials' motion data being lost from the University of Iowa simulations.

3.4 Motion Analysis

3.4.1 Qualisys Optical Motion Capture

Motion data captured using the Qualisys Oqus camera system were post-processed using Qualisys Track Manager, a program for use with Qualisys camera systems, to identify and register motion data to the fiducial system. This removed markers that were noise and identified a continuous path for the markers representing each hand. The marker data were then exported into a MATLAB-friendly format. Data from the video sync was then used to remove data capture pre and post simulation as well as motions not directly related to reduction such as direction for the fluoroscopy technician or reaching for a tool. The motion data were used to calculate cumulative motion (Euclidean Distance). Since surgical instrument motion was related to hand motion, iatrogenic surgical soft tissue trauma could be estimated to scale with motion distance. Hand velocities were calculated from the position data, and smoothed using an averaging filter with a window width of 11 frames. Time-points during which

acceleration was zero ± a threshold, picked by comparing the raw motion data to synchronized video and determining new motion points, were flagged as discrete hand motions. Initially this resulted in an overestimation of hand motions due to signal noise, so local minima in the velocity trace were identified with MATLAB's 'imextendedmin' function. With a threshold of 10 mm/s, this method proved less susceptible to noise than the acceleration-based method. The correspondence between the local minima and discrete actions were subjectively validated using the synchronized video.

This system has two weaknesses that contributed to its being replaced: line of sight of the markers must be maintained or no data are collected, and the camera on-board memory can fail to transfer data to the computer hard drive (1 case in 2010 and 3 cases in 2011). Both of these limitations result in a failure to accurately capture hand motion. The system also had to accommodate noise caused from reflective surfaces (such as a trim ring on the fluoroscope) being in the field of view that are not markers. Finally, a substantial time investment is required to identify the markers in each trial due to noise issues; the Qualisys system was picking up reflections that were not the markers on the back of the hand causing the system to not be able to identify the eight markers that were supposed to exist automatically.

3.4.2 Polhemus G4 Electromagnetic Motion Capture

The Polhemus G4 system uses the device-specific calibrations done in the Polhemus for use with the basic tracking software PiMgr, provided by Polhemus. The calibration must be configured by the end user for the environment in which it is to be used, and is tailored using the Polhemus software. Once the system is configured and turned on, using PiMgr, an export pipe is created for communication. Using a custom written Python script executed in the command window, the data pipe is streamed to a text file in real time. The G4 system has a theoretical maximum data collection rate of 120 Hz. In our experience (we achieved between 40-100 Hz collection rate), the Python

script was able to record all data. Post hoc the data were imported into Excel and converted into a MATLAB-friendly format.

Once in MATLAB, the data were truncated using the double clap hand motions at the beginning and ending of the simulation. In MATLAB, the data were then parsed into Cartesian axis positions; roll, pitch, and yaw data were not used. The time difference between each data collection instance was then computed. Position data were then converted into velocity data using the local time interval between two points. The number of data points collected per second was calculated. Euclidean cumulative hand motion distance was calculated from the position data. The procedure for calculating the discrete hand actions is identical to the methodology used for calculation of hand actions with the Qualisys system, except for only having one marker per hand instead of four.

This system has a main limitation, which is the susceptibility to noise caused by ferro-magnetic objects moving within the electromagnetic field. Once the field is clear of ferro-magnetic objects, or the objects have been accounted for in the system configuration, noise is minimal. The non-changeable battery in the sensor hub is also an issue, as you must either wait for the hub to recharge or reconfigure the system to a different hub when the battery dies.

3.4.3. Comparison of the Motion Capture Systems

A limited comparison test was done between the Qualisys and G4 systems. The markers for the Qualisys system and the sensors from the G4 system were placed on the back of the hand exactly as during the simulations, and a variety of hand motions were tracked. The difference between the systems was calculated and compared. Six tests were performed consisting of different motions and time lengths. The first was a completely static test (i.e., there was no hand movement) to determine the noise that the two systems registered. The second and third tests consisted of linear movements in vertical and horizontal directions, respectively. The fourth test involved moving both hands in

continuous circular motions. The fifth and sixth tests consisted of random movements limited to the recording space of the Qualisys system. The number of hand motions was counted during the motion, and counted again by watching the Qualisys tracker data played in a 3D viewer, to compare against the hand motion data measured from the motion capture systems. A portable X-ray machine was placed in proximity to the test area to replicate the effects that a fluoroscopy unit would have during the surgical simulation. The G4 source was placed 1 foot away from the motion window origin of the Qualisys to mimic the set up used during the Minnesota trials.

3.5 Articular Surface Deviation (Step-off) Measurement

The residual incongruity of the articular surface is an important primary performance measure for the surgery that was simulated. Three different methods of surface deviation analysis have been performed on the simulated fracture reductions. The first involved using the commercial software package Geomagic Qualify, the second used a combination of Geomagic Qualify and Microsoft Excel, and the third method used a custom MATLAB script.

The Geomagic process started with a laser scan performed using the Next Engine laser scanner. The scan was then imported into Qualify and post-processing was performed. The post-processing included using built-in capabilities for filling in holes, removing non-manifold surfaces, and noise reduction. The cleaned digital surface model was then duplicated, and the duplicate was split into 3 surface models, one model for each of the individual fragments. The base fragment from the reduction was then manually brought into a rough alignment with the base fragment from the ideal digital reconstruction. After rough alignment, the ‘Best Fit Alignment’ tool was used to better align the reconstructed base fragment to the ideal base fragment. This tool makes use of a two stage Iterative Closest Point (ICP) algorithm to accomplish the alignment; the first stage does a rough global alignment, and the second stage uses a high fidelity local ICP

to complete the alignment process. Thresholds of 0.5mm for the global rough alignment and 0.25mm for the second stage local alignment were used; this was done in order to allow for inaccuracies in the molding and laser scanning processes.

The spatial transformation associated with this alignment was then saved, and applied to the other fragments to bring them into the correct orientation for surface comparison. After the two non-base fragments were aligned, the articular surfaces were copied to create a surface mesh for each of the reconstructed fragments and their ideal counterparts. A 3D deviation analysis was then performed using the built-in compare tool in Geomagic Qualify. This tool performs a deviation mapping using a nearest surface method, in which a point on the test surface (reconstructed articular surface) is compared to the reference surface (puzzle solved articular surface, 'ideal reconstruction') to find the smallest Euclidean distance between the two, regardless of whether the point on the test surface is the same point as on the reference surface. The measurement provides both a positive and negative average of the distances, as well as a min and max positive and negative distance. The positive and negative result from the surfaces have an inner and outer face. When these data are exported, a Geometric Dimensioning and Tolerancing (GD&T) operation was applied in which the two averages are converted to absolute values and added together, creating an artificially larger deviation mean.

The second surface comparison using a combination of Geomagic and Excel follows the same workflow as the Geomagic alone methodology, save for the last step. Instead of reporting a single value from Geomagic, the points from the test surface and the points from the reference surface against which Geomagic compared the test points were exported into Excel. The Euclidean distance was then found for each set of points, and converted to the absolute value. The mean and standard deviation of the distances were then calculated using built-in Excel functions. The values were then aggregated across the simulation and compared to the values reported by Geomagic.

The third method of surface deviation analysis was performed in MATLAB. The reconstructed surfaces (obtained from either laser scanning or CT scanning and segmentation) were aligned to the ideal template as outlined in the Geomagic method. Next the ideal digital models of the individual fragments were copied, and the copies were aligned to the reduced fragment's position. This was done to ensure that the same points would be compared between the ideal positioning and surgically reduced position. The transform obtained through that alignment was exported into a file that could be read by MATLAB. The articular and whole-fragment surfaces from the ideal digital model were also exported as .STL files so that the files could be imported into MATLAB. Each simulation trial has a transform corresponding to the difference between ideal and reconstructed surface for the two non-base fragments, and all the cases were compared using the same articular surface models generated from the ideal fragment model.

The transforms and articular surface .STL files were imported into MATLAB. Once imported, the surface models were copied and had the respective transform applied. Due to .STL surfaces being a list of points and their position in 3D space, the difference in each of the three axial directions between points results in a vector relating the distance between each set of points. After finding each difference vector between the set of points, the vectors were converted using the 3D Pythagorean Theorem into a list of Euclidian distances between the points. The mean and standard deviation of each list of points was calculated using the 'mean' and 'std' functions built into MATLAB. Histogram data were then compiled in 0.5 mm wide bins. This was done for both fragments for each case. The averages and standard deviation for each fragment were then added together using a weight based on the number of points on each surface. The two histogram counts were also added together to create a histogram for the combination of both fragments. A colored surface mapping of the two fragments and their ideal counterparts was then created and displayed in MATLAB using the VTK plugin (www.vtk.org, open-source toolkit BSD license, copyright 1993-2008 Ken Martin, Will Schroeder, Bill Lorensen).

Because the alignment used the base fragment to position the reduced tibia with respect to the ideal tibia, no surface comparisons were done using this fragment, as the deviation would be identically zero (or very nearly so). This would contribute to produce an artificially lower deviation as the articular surface on the base fragment consists of approximately fifty percent of the native articular surface depending on which fracture model was used.

3.6 DEA of Simulated Ankle Fracture Reduction

The DEA algorithm is implemented in MATLAB. The model is created from the segmented CT scan or the refined laser scan. The model is aligned to a functional apposition and then a uniform cartilage thickness is offset form the subchondral bone surfaces in Geomagic. The cartilage models are then exported as .STL files.

The contact surfaces are aligned so that they are unloaded or barely touching, for proper contact force calculation as they are run exclusively in load control DEA. One surface is classified by the analyst as the master, and the other as the slave. The master surface has one contact spring for each facet on the surface, where the slave does not. This accommodates for mismatches in the number of facets of the surface meshes. The algorithm implemented consists of a pair of rigid contact surfaces lined with a system of linear springs as contact points. The algorithm iterates as a series of static solutions. Each iteration involves the computation of contact stress at that joint position, the resulting reaction forces are computed and the solution is checked against completion of the termination constraints. Load balancing adjusts the displacement of the joint incrementally until termination constraints have been met. Every iteration starts with the assumption that no contact exists and no springs connect the surfaces. Overlap is then determined between the contacting surfaces, and a spring system is created between the contacting regions.

The deformed length of each spring then relates to the engendered contact stress in each facet of the contact surface. Each iteration of the contact solution is solved by (1) selecting the nearest neighbor between polygon centroids on contacting surfaces, (2) creating a system of springs between the nearest neighbors, (3) correcting spring displacements to normal projections, (4) computing contact stress between remaining springs, and (5) determining forces engendered by contact stress[94] (Figure 7). Kern found the most efficient methodology to find the nearest neighbor(NN) to be the Voronoi method [94].

After the nearest neighbors have been identified, determination of which polygon to include in the DEA calculations is made. Springs are generated for the polygons of which the contact surfaces have undergone penetration, which is determined by the sign of the dot product of each polygon's outward facing normal and the vector connecting the two polygons paired in contact. If penetration has occurred the sign will be negative[94]. Because the contact between the distal tibia and talus can be roughly modeled as two cylinders, the distance between the center of mass for each can be computed, and the distance between the talar center of mass and the tibial center of mass can be found. The distance between the talar center of mass and talar surface will be less than the distance between the talar center of mass and the tibial surface points for each set of nearest neighbors which do not have penetration. This allows for culling of candidates which would not be ideal for contact without setting an arbitrary minimum distance threshold using the dot product.

After the spring system has been created, adjustment is made to correct as the algorithm assumes that the stresses computed on each polygon are exclusively normal to the face of the polygon [113, 114]. The correction is done by projection of the spring displacement vector along the normal vector of the polygon. This is accomplished using the dot product between a unit normal vector and the spring vector to calculate the new contact spring displacement.

$$S' = dot(norm(N), S)$$

Where S is the original spring vector, N is the polygon normal and S' is the new projected spring direction. By projecting all of the springs in the normal direction of the contact surface, most aberrant contact stress patches are corrected. After spring creation and correction, contact stress can be computed. A force-displacement spring model is used. This uses the spring deformation (d), effective cartilage thickness (h), Young's Modulus (E), and Possion's ratio (v) to determine contact stress[113].

$$p = \frac{(1-v)E}{(1+v)(1-2v)}\frac{d}{h}$$

The contact stress is computed for each facet on the contact surface. In the use of this algorithm, cartilage thickness is assumed to be uniform at 1.7 mm [131]. The deformation is represented by the corrected distance between each NN pair, because contacting surfaces are used and the bone surfaces are implicit[94].

The contact pressures and areas for each contacting polygon are used to compute the discrete contact forces acting on each facet in contact. The contributions are summed vectorially to find total contact force and moment acting over the whole contact surface. The vector is found with respect to each surface's centroid and is output for each iteration of the DEA algorithm. It is used with the iterative load control algorithm aspect of the DEA until termination conditions are achieved [94].

The load control is used as an iterative cycle of quasi-static positioning in which an optimization algorithm is used to compute force and moment equilibrium subject to appropriate tolerances. Before starting the iterative process, the user inputs tolerances and loads are defined for all six degrees of freedom for each surface model, and step sizes for each translation and rotation are also input (Figure 8). For each iteration of the contact stress analysis, the algorithm checks the forces and moments against boundary and loading conditions. If the force or moment meets the requirements for equilibrium, no transformation is made for the corresponding degree of freedom. If equilibrium is not

achieved a rigid affine spatial transformation is applied. The magnitude of the transformations is tied to the associated step size for the given deformation. Due to the step size, after an arbitrary number of iterations, the algorithm may be oscillating around the desired loading due to too large of a step size. If the loading value oscillates for four consecutive iterations, then the step size is decreased 50%; this results in the optimizer being able to start with large gross adjustments and refine as needed to find a suitable kinematic position[94].

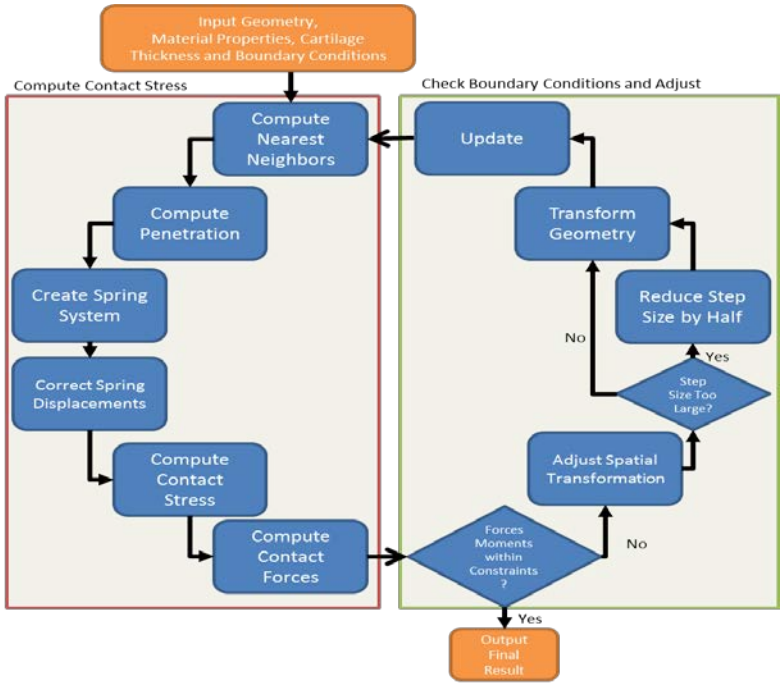

Figure 7: Flowchart of DEA algorithm where subject-specific geometries, material properties, cartilage thicknesses and boundary conditions are input into the algorithm.

Source: Kern 2011 [94]. Used with permission

Note: During an iterative process, contact stress is computed using a penetration computation and balanced using a simple load control algorithm until final termination criteria are completed.

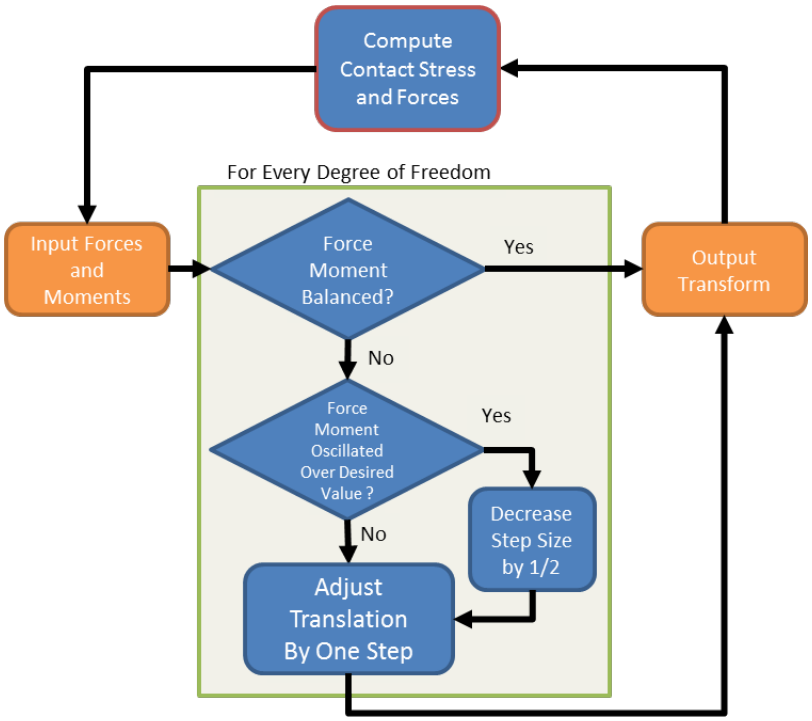

Figure 8: Flowchart of iterative load control system which manipulates translations and rotations to push the DEA algorithm to equilibrium.

Source: Kern 2011 [94]. Used with permission.

Note: Step sizes for translation are reduced once the system begins to oscillate about the termination criteria.

The DEA algorithm was run through a 13-step flexion-extension motion path based on the stance phase of the gait cycle [94, 105]. The subject was loaded for each of the 13 step gait cycle as a percent of body weight for each step as determined by motion capture and force plate data [132]. The body weight used with the surrogate ankle was 600N, corresponding to 135lbs. This is an assumption based on the size of the tibia. A Young's modulus of 12 MPa and a Poisson's ratio of 0.42 was used to correspond to previous DEA and FEA studies for the cartilage, which was considered to be a linear elastic frictionless material[94, 105]. Using the DEA results, the area of contact stress

above a given threshold at the step time with the largest force applied and area with a peak contact stress above a threshold over the entire stance phase were plotted. The contact stress metric was also evaluated.

CHAPTER 4 – RESULTS AND DISCUSSION

4.1 Results of Early University of Iowa Simulation Trials

4.1.1 2010 Results

In the 2010 trials, motion capture was done using the Qualisys Oqus camera system, measurement of the articular step-off was done using the Geomagic method and OSATS scoring was performed by blinded video. During the trials, there was a wide variation in performance by the residents, with more variation among the junior residents. Two junior residents and one senior resident failed to complete the fixation of all fragments. In discussions following the simulation, residents all commented that the exercise recreated the operating room experience, and felt that it improved their ability to successfully reduce articular fractures using a limited surgical approach.

For the most part, no significant differences were found in the objective metrics except for the cumulative hand distance measurement in which junior residents had a mean of 390 meters compared to 79 meters for the senior residents (table 1). This was a significant result ($p<0.01$) even with a very small sample size, suggesting that as skill level increases hand motion will decrease. The research group has coined this unskilled extra motion 'flailing', and hypothesized that as experience increases the amount of flailing will decrease. With no difference in time to completion or number of discrete hand motions, it appears as though the speed and number of motions required to complete the procedure do not need to be improved, unlike the need to improve the economy of motion. With figure 9**,** it is obvious that the senior residents do a better job at working close to the surgical opening and limiting unnecessary motion.

The large difference in standard deviation for the articular error indicates that testing a larger population of residents could prove significance for articular step-off between junior and senior residents. Since the articular error was measured using the

Geomagic method, all of the associated disadvantages of the point-to-surface comparison, are present.

The OSATS global rating metric almost showed a significant difference (p=0.088) between junior and senior residents; senior residents scored higher, as expected based on their greater level of experience (table 2). With a small increase in sample size, a significant result might well be found. A weak negative correlation was found between the OSATS global rating score and cumulative motion distance (r=-0.31). No significant difference was observed between junior and senior residents on the procedural checklist portion of OSATS. No correlation was observed between the OSATS global rating score and articular error, duration, or number of discrete motions.

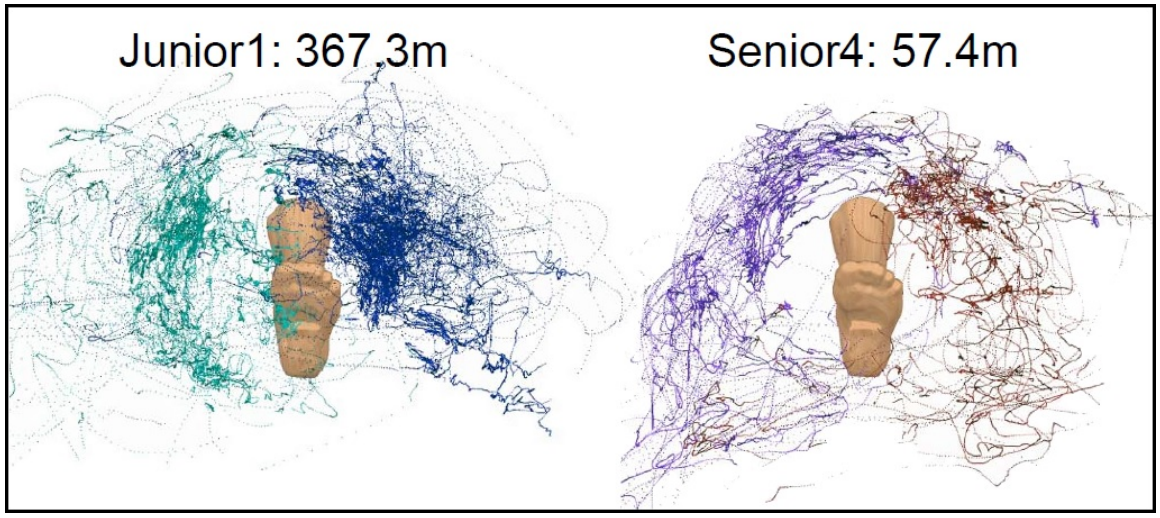

Figure 9: Hand motion traces for each of the junior and senior the resident with the closest to mean cumulative hand distance, illustrating the difference between the two groups where the junior not only has more distance, but more motion at the edge of the operating environment

Source: Yehyawi 2012 [48] Used with permission

Table 1: Mean results for 2010 University of Iowa simulation for senior and junior residents with standard deviation and p-value comparing the two groups

	Senior	Junior	P-Value
Mean Articular Error (mm)	3.00	3.09	0.856
Standard Deviation	0.43	1.25	
Time To Completion(s)	805.75	885.00	0.734
Standard Deviation	281.00	467.00	
Cumulative Hand Motion (m)	78.71	390.42	0.006
Standard Deviation	48.48	175.78	
Discrete Hand Motions(#)	539.50	511.16	0.879
Standard Deviation	303.03	227.39	

Note: Significance test was a two tailed t-test

Table 2: Results for OSATS Global Rating Metric for the 2010 University of Iowa simulation for senior and junior residents

	Senior	Junior
Mean	28.60	23.00
Standard Deviation	3.98	5.37
Max	34.00	32.00
Min	22.00	13.00
P-Value	0.088	

Note: Significance test was a two tailed t-test

Note: Score out of a possible 45 points

4.1.2 2011 Results

In the 2011 simulation trials, motion capture was performed using the Qualisys Oqus camera system, articular step-off was calculated primarily using the Geomagic method (6 cases were tested using the MATLAB method, and all 12 were tested using the Excel method), and OSATS scoring was performed live. The residents were randomly assigned to a control group or intervention group immediately following completion of

the pretest, without the senior resident making the assignments having knowledge of any of the outcome metrics' results from the pretest.

The OSATS scores for the pretests indicate that the two randomized groups were not identical and that the control group performed better. However, this difference was most likely due to chance and the use of a small sample population (table 3 and figure 10). The intervention group did show a significant improvement in OSATS score after the intervention, and the scores also became more consistent; whereas there was no improvement in the control group and the variation increased.

These data suggest that the guided practice and video watching provided a significant increase in the skills that expert surgeons consider important to successfully complete an articular fracture reduction using a limited approach. The results also indicate that limited practice with the simulator alone does not improve the skills and qualities that an expert surgeon would find important.

In this round of simulation, the motion capture data did not show any significant differences between the study groups (figure 11). There was however an issue where the data from two control group residents and one intervention group resident were lost when transferring the data from the camera to the laptop PC. Data trends did indicate that as the residents became more familiar with the procedure and surgical instruments, the distance their hands moved was decreased. The intervention group also had a slight increase in the number of discrete hand motions and a small decrease in total hand distance, which would also indicate an increase in skill when compared to the results of the five senior surgeons from the 2010 simulation trials. Once again a larger sample would prove to be beneficial as a significance threshold might well have been reached even with small differences between the two groups and two simulation trials.

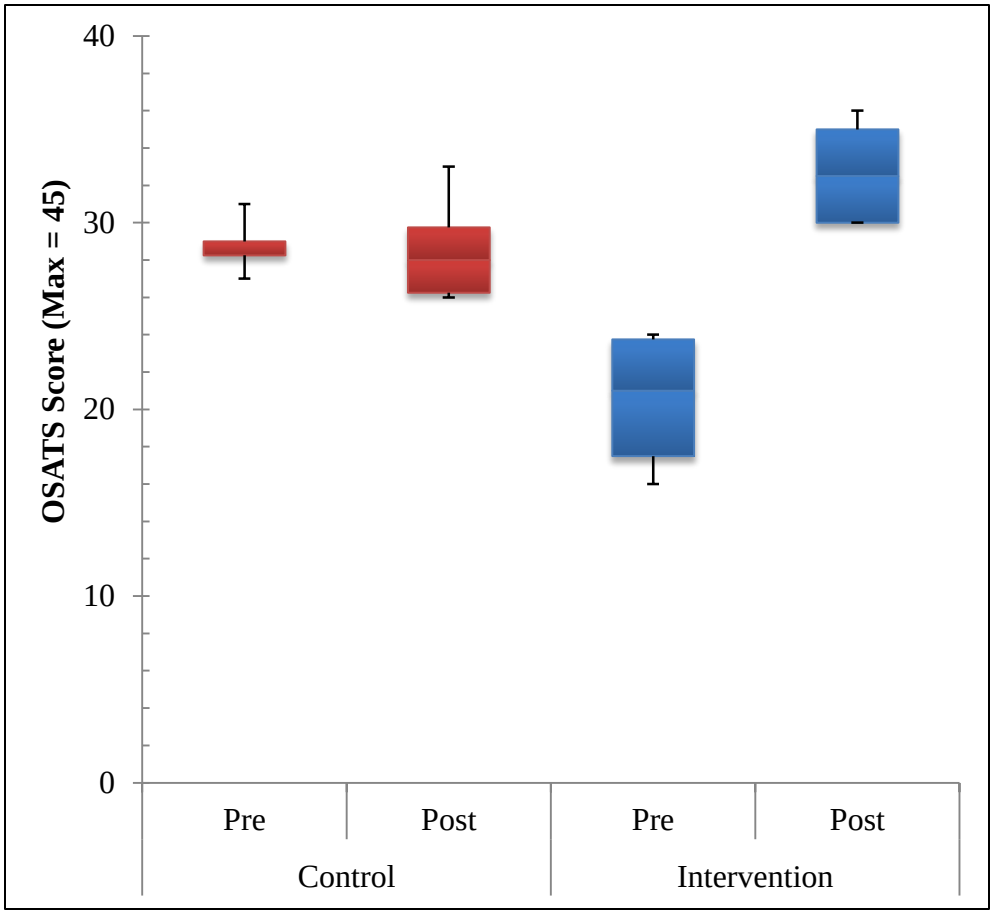

Figure 10: Box and Whisker Plot of OSATS score for 2011 University of Iowa simulations for both the pre and posttest simulations, illustrating the improvement in the intervention group in both overall score and consistency between resident performances

Note: Control group in red and intervention in blue

Note: For the Control Pretest, the median score and 75th percentile are the same value, due to small sample size.

Table 3: Student T-test P-values for 2011 University of Iowa simulations with Mean score for each group for each simulation instance

	OSATS Overall Score		
	Control	Intervention	P Value
Pre-Training	28.8	20.5	0.0004*
Post-Training	28.5	32.7	0.029*
P Value	0.788	0.001*	
Improvement	-0.3	12.2	0.003*

Note: Right P value column compares the Control and Intervention scores for the pre or posttest simulation along with the improvement between each group as well.

Note: Bottom P value compares the resident group between the two simulations

Note: T-Test are two tailed, and homoscedastic for intergroup comparison and paired for intragroup comparison

Note: Due to small sample, even with random distribution of residents; the significant difference between the control and intervention group occurred most likely due to different native skill level

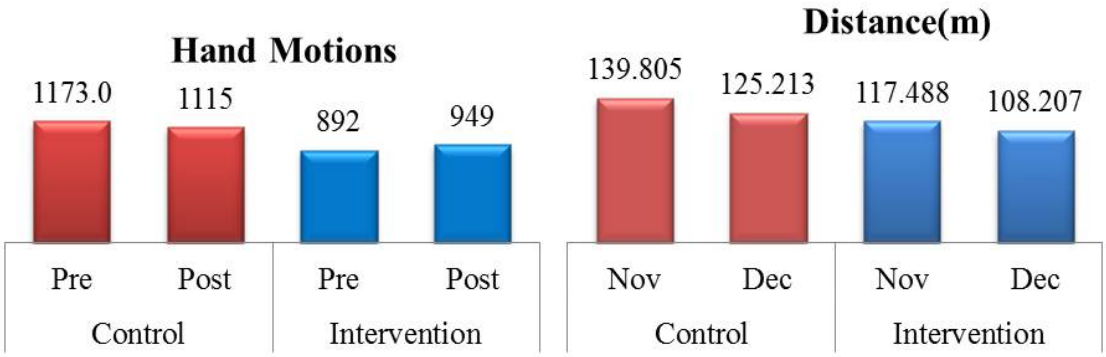

Figure 11: Hand motion data from 2011 University of Iowa simulation for control and intervention groups.

Note: Motion capture data acquired using Qualisys Oqus camera system

Fluoroscopy usage decreased for both groups (figure 12), but the decrease was not significant for the control group. The control group's median usage went from 27.20 seconds of use to 23.20, and the intervention group's median decreased from 27.70 seconds to 16.70 seconds. The variation in use between members of the intervention group decreased as well, with the standard deviation of fluoroscopic time decreasing from 6.26 seconds to 6.00, and the max decreased from 40.40 seconds to 23.90 seconds. The control group's standard deviation also decreased from 11.65 seconds to 9.14, but the max decreased from 42.30 seconds to 35.00 seconds. The mean score from the control group only dropped from 24.63 seconds to 21.98 whereas the intervention group dropped form 29.05 seconds to 16.38.

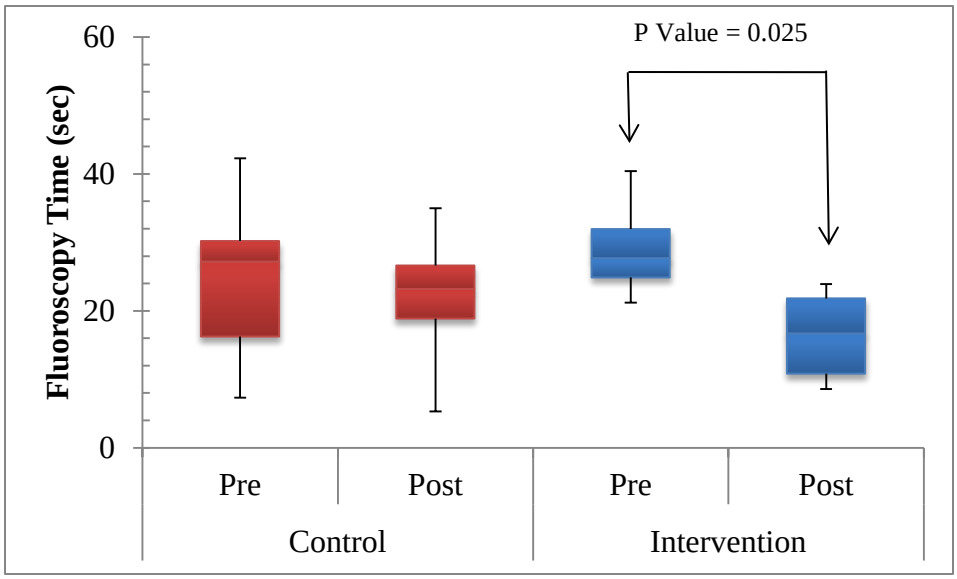

Figure 12: Fluoroscopy time usage box and whisker plot for 2011 University of Iowa simulation.

Note: Fluoroscopy time was used instead of fluoroscopy images as multiple residents chose to utilize the video feature of the fluoroscopic rather than static images for a part of the simulation

Note: P value obtained from Student's T-Test as a paired two-tailed comparison

The reason that fluoroscopy was recorded as a time usage rather than simply the number of images was because multiple residents used the ability of the fluoroscopy equipment to continuously capture imagery in a low frames per second video, rather than taking single static shots. The machine used in the simulation did not have the ability to record the number of individual frames taken in the video method and did not save them all to be counted post hoc. Regardless of the method used to acquire the fluoroscopy, any decrease in the radiation is considered desirable, all other factors being equal.

Table 4: 2011 University of Iowa simulation articular step-off values

	Articular Step-Off(mm)					
	November Session (Pre)		December Session (Post)		Improvement(mm)	
	Control	Intervention	Control	Intervention	Control	Intervention
	3.25	6.70	3.39	3.66	-0.14	3.04
	6.08	3.96	3.54	3.02	2.55	0.93
	6.02	6.24	2.89	3.93	3.13	2.30
	3.86	5.30	4.52	6.63	-0.66	-1.33
	5.92	3.69	1.81	3.48	4.11	0.21
	3.24	3.71	5.64	3.28	-2.41	0.43
Mean	4.73	4.93	3.63	4.00	1.10	0.93
T-Test Compare	P=	0.827	P=	0.603	P=	0.886
	Control Comparison T-Test		Intervention Comparison T-Test			
	P=	0.338	P=	0.205		

Note: Improvement is shown for each of the 12 residents along with a mean for both the control and intervention groups.

Note: P-values are given for student t-test results

Note: T-tests are all two tailed. For intragroup comparison between testing session the tests are paired, otherwise the tests are treated as homoscedastic

Note: Articular step-off calculated using the Geomagic method

The articular step-off measurements using the Geomagic method did not show any significant results (table 4), yet both the control and intervention groups show a trend of improvement, which might prove significant with a larger sample size. The control group had three residents that got worse while the intervention group only had a single resident do worse, but the control group residents who did improve, improved by a larger margin. Limitations in the Geomagic calculations and non-significant results preclude any definite conclusions, but the data do indicate that the training creates a more consistent outcome with regards to articular step-off. The data also indicate that the control group was very inconsistent, and the improved reductions may be due to chance rather than surgical skill. A larger sample population would provide clarity on the ability of guided practice to improve the minimization of articular step-off, or if practice is all that is needed to improve the resident's abilities.

Table 5: Pearson's r correlation values for different outcome metric for 2011 University of Iowa Simulation with different strength of correlations indicated by highlight color

Overall	Pearson's r
Action/distance	0.946
Action/Step-Off	-0.045
Action/OSATS	0.087
Distance/Step-Off	0.090
Distance/OSATS	-0.042
Step-Off/OSATS	-0.185

Control	Pearson's r
Action/distance	0.948
Action/Step-Off	0.171
Action/OSATS	-0.717
Distance/Step-Off	0.246
Distance/OSATS	-0.612
Step-Off/OSATS	-0.446

Correlation
None
Small
Medium
Strong

Intervention	Pearson's r
Action/distance	0.950
Action/Step-Off	-0.117
Action/OSATS	0.262
Distance/Step-Off	0.045
Distance/OSATS	0.064
Step-Off/OSATS	-0.117

Table 5 provides the correlation of outcome metrics for the 2011 University of Iowa simulation trials. Overall, most metrics had no discernible correlation, except for the understandably strong correlation between number of hand motions and hand distance, and a weak negative correlation between the articular step-off and OSATS score. The negative correlation is a positive indicator, because as articular step-off decreases the OSATS score (or what an expert surgeon finds important) increases.

When the control and intervention groups were considered separately, most metrics exhibited stronger correlations. The biggest positive indicator is that the step-off/OSATS correlation is of medium strength for the control group and does not significantly change for the intervention group. The control group data also indicate that the number of hand actions (motions) and distance data are negatively correlated with the OSATS score, and the hand actions and distance weakly correlate with step-off. The control group had almost ideal correlations between the outcome metrics. The intervention group did not show as strong or as many correlations, part of which may be due to the poor initial OSATS performance by the intervention group.

Three of six intervention residents increased their number of hand motions by several hundred, whereas two slightly decreased their motions, and one resident had no preliminary hand motion data. Subjectively, the three that increased the number of motions seemed to be very inexperienced and frustrated by the fragments not reducing as the residents wanted. When separating the pretest and post test data for the intervention group, the correlations are more evident as the pretest correlation of a Pearson's r of 0.785 for actions to OSATS score, which makes sense, since the three residents that increased the number of actions between the pretest and posttest, had few actions and relatively poor OSATS scores compared to the two with more motions. For the posttest correlations a strong correlation of -0.671 exists between the number of hand motions and articular step-off, and for hand motions compared to OSATS a weak positive correlation exists (r = -0.178). The posttest correlations make sense as the resident with the fewest

hand motions had the highest step-off, but that resident did have several hundred fewer motions that the next closest resident. The resident with the third fewest motions had the third highest step-off as well, and the second worse setup resident had the most hand motions. Once again, a larger population would improve clarity.

4.2 2013 University of Iowa Simulation Results

For the University of Iowa simulations that took place in 2013, the results include fluoroscopic data (number of images and radiation exposure), hand motion data (recorded by Polhemus G4 system), OSATS score (from live scoring), articular step-off (measured using the MATLAB analysis method), and contact stress analysis (DEA). The simulation trials were all conducted on fracture pattern A.

Fluoroscopic results indicated that practice alone did not help the residents minimize their use of fluoroscopic imagery (i.e., student T-test for the difference in number of images taken between the pretest and posttest show no significant difference) (table 6). There is however a significant improvement from before the video coaching by a senior orthopaedic traumatologist and after the coaching in which the traumatologist outlined and explained what needed to be seen in the images and how to obtain the best possible fluoroscopic images (between posttest and retention test). Between the posttest and retention test, there was a decrease in the variability in the number of fluoroscopic images (figure 13). As radiation exposure scales with number of images, similar results occur in exposure. The slight differences between the number of images and exposure is due to the ability of the resident to ask the radiology technician to turn up the power of the fluoroscope for sharper images, which some residents chose to do.

With the motion capture data from the G4 system, no data were lost. However, after visual comparison of the head mounted video and motion capture data, some inconsistences did appear. The magnetic field source was placed at one end of the surgical table as to not impede the ability of the resident to place tools near and work

with the ankle model as they saw fit. This led to source placement at the opposite end of the table from where some residents chose to work, which corresponded with being at the edge of the acceptable range of the G4 system. As such, the accuracy of the results must be interpreted with some caution. This issue was addressed prior to the University of Minnesota trials. There were no significant differences between the simulations for the number of discrete motions or total distance traveled (table 6). There was a trend of the number of motions becoming more consistent across the different residents as the residents became more experienced (figure 14). There was also an effect where the mean distance decreased from pretest to posttest accompanied by a large increase in variability in distance traveled (Figure 14). During the retention test the residents became more consistent with distance, but the mean increased to between the pretest and posttest. Sample size once again caused issues, as the changes are too small to be significant with such a small sample size.

Table 6: 2013 University of Iowa Simulation T-Test P values for fluoroscopic exposure and images, hand motion capture data, OSATS scoring, and articular step-off for individual and joint fragments

P Value	mAs	Images	Distance(m)	# of Motion	OSATS	Both	Frag2	Frag3
Pre-Post	0.680	0.965	0.878	0.743	0.025	0.411	0.764	0.609
Pre-Ret	0.026	0.036	0.433	0.183	0.007	0.645	0.595	0.861
Post-Ret	0.006	0.009	0.641	0.858	0.141	0.418	0.479	0.794

Note: Yellow highlight denotes significant difference for $p<0.05$, and green highlight denotes significant similarity for $p>0.95$

Note: mAs is fluoroscopic exposure as a rate of electric current–time for fluoroscope usage, Images is the comparison of the number of fluoroscopic images

Note: Both is the combination of articular step-off for the two floating fragments, and Frag 2 and Frag 3 are the individual fragments

Note: All T tests performed were two-tailed and paired

The OSATS scores for the residents improved between pretest and posttest simulation, and were similar for the posttest and retention test (table 6). The mean score went from 18.67 during the pretest to 25.43 during the posttest to 27.14 during the retention test. As seen in figure 15, variability among the scores increased as the resident became more experienced, with the maximum score increasing more than the minimum. However the minimum score during the retention test was higher than the maximum score during the pretest, indicating that improvement occurred during the simulations. As seen in previous generations of the simulator, the results show that with practice on the simulator, the skills that expert traumatologists deem important improve.

The articular step-off measurement for each of the simulation trials was performed using the MATLAB method to obtain a true point-to-point surface comparison. The articular deviation was calculated for each of the two floating fragments individually and for both together, in order to determine if the residents were better able to reduce one fragment compared to the other. For the most part, the posterior fragment, labeled fragment 2 (frag 2) from the fracture pattern A (figure 2, purple fragment), proved more difficult than fragment 3 (frag 3), the light blue fragment. Due to the small sample size, the mean can be easily influenced by an outlying step-off from a single case. For example, the large deviation from subject purple in the retention test for fragment 2 skewed the data, making the median a better representation of the step-off as a group (table 7 & figure 16). The standard deviation (SD) of the step-off gives an indication of how rotated a fragment is from the natural position. If the fragment is greatly rotated, the SD is higher than a fragment that is not rotated out of position. The translation of a fragment appears to affect the mean but not to have an effect upon the SD. This is because the point-to-point measurement is the Euclidean distance, and for a purely translated fragment, the SD would be zero; whereas rotation would increase the SD depending on the center of rotation. However, a misrotated fragment will still have an

increased mean step-off compared to a correctly rotated fragment, as the misrotation will still increase the point to point distance between ideal and reconstructed point pair.

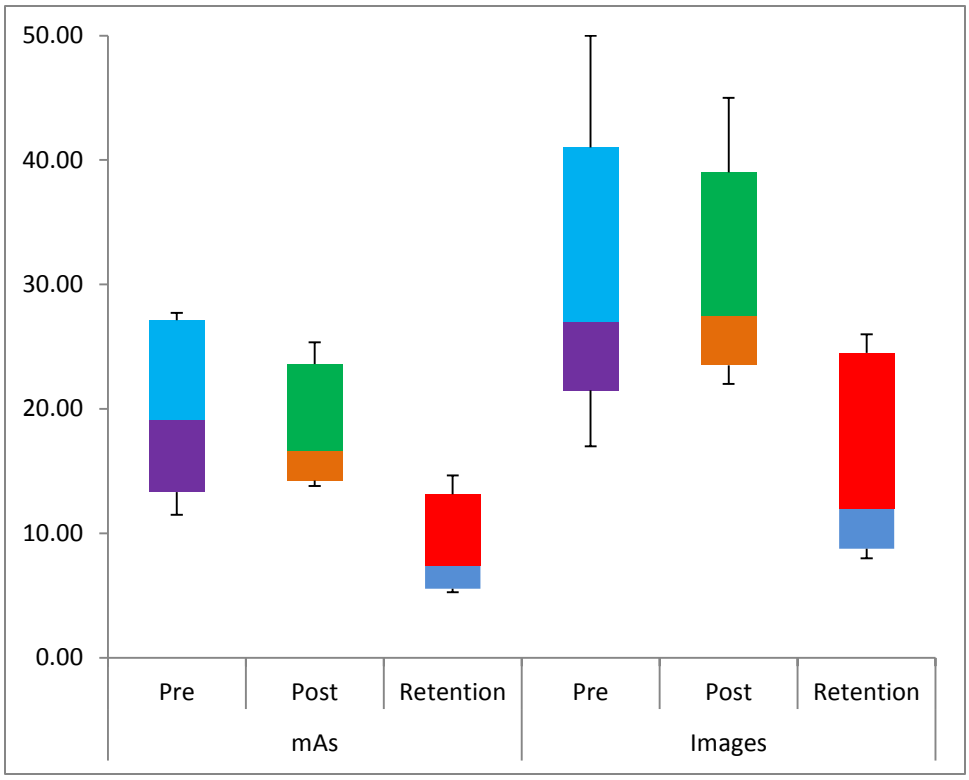

Figure 13: Box and whisker plot for 2013 University of Iowa simulation-fluoroscopic data, with exposure in milliamperes-seconds and number of images.

Note: The decrease in fluoroscopic use is obvious between the posttest and retention test when the coaching of fluoroscopic use occurred for the residents.

Note: The improvement in consistency of the whole group between the post and retention is also evident.

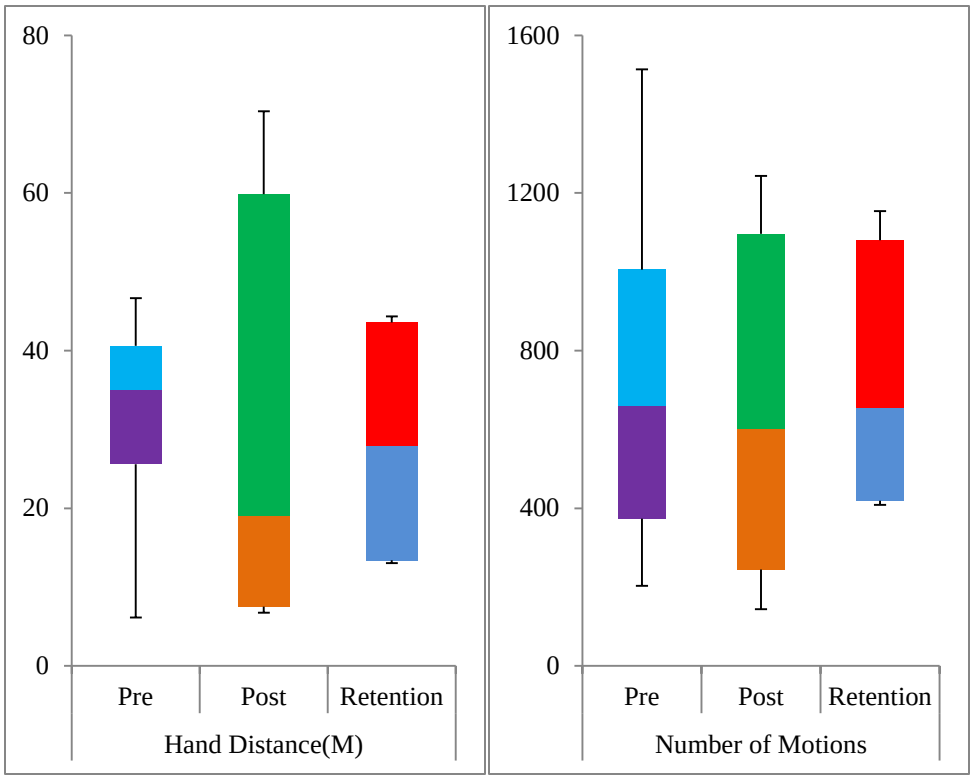

Figure 14: Box and whisker plot of 2013 University of Iowa simulation for hand motion capture data. Left: cumulative hand distance traveled during the trial for both hands. Right: Number of discrete motions performed by each hand.

Note: The difference in distance between the pretest and posttest is evident as the median decreased greatly, but the variability increase, while the change for the post - retention interval saw an increase in median distance but more consistency.

Note: The median number of motions stayed consistent between the trials, but variability decreased.

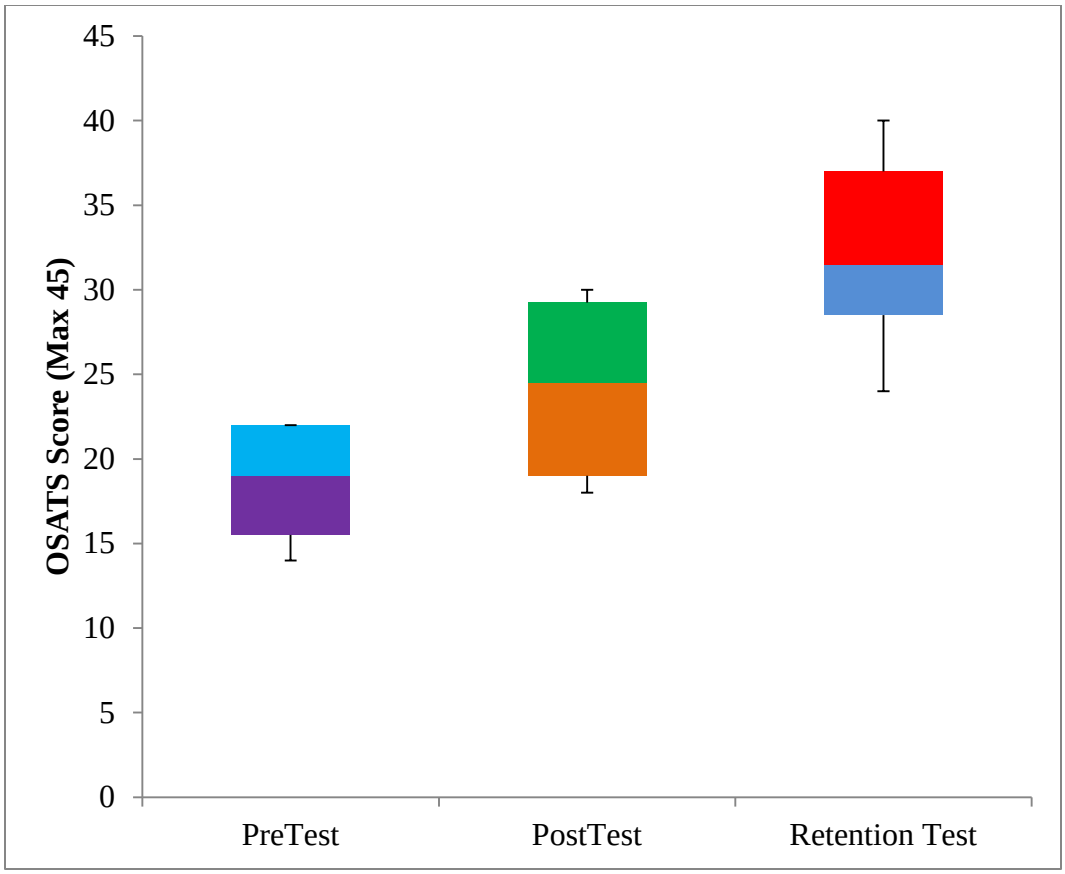

Figure 15: 2013 University of Iowa simulation OSATS score box and whisker plot.

Note: Reported Score is the sum of all nine categories scored on a 1-5 point scale given a maximum possible point score of 45

Note: The upward trend between trials indicates improvement which was significant between the pretest and posttest, but not between the posttest and retention test.

Note: The max score for each trial increased at a greater rate than the minimum score indicating that while the residents improved as a group, some improved at a greater rate than others.

Table 7: 2013 University of Iowa articular step-off results calculated using MATLAB point to point analysis method, to find the mean step-off and standard deviation of the step-off

All Units (mm)	Pretest						Posttest						Retention test					
	Frag 2		Frag 3		Both		Frag 2		Frag 3		Both		Frag 2		Frag 3		Both	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Blue	5.650	0.70	1.793	0.30	4.028	0.53	3.395	0.81	5.142	0.39	4.130	0.63	2.350	0.25	3.619	0.31	2.883	0.27
Brown	3.378	0.57	3.619	0.31	3.479	0.46	2.955	0.48	2.677	0.50	2.838	0.49	4.376	0.12	1.179	0.28	3.031	0.18
Green	3.643	0.47	3.661	0.33	3.650	0.41	7.539	0.38	0.757	0.34	4.686	0.36	3.266	0.66	7.385	1.16	4.998	0.87
Orange	3.576	0.58	7.518	1.16	5.234	0.83	3.766	0.35	3.303	0.72	3.571	0.51	4.190	1.74	1.823	0.72	3.194	1.31
Purple	1.652	0.24	2.454	0.41	1.989	0.31	2.740	0.51	2.369	0.37	2.584	0.45	17.901	2.87	2.891	0.63	11.588	1.93
Red	6.288	0.50	1.583	0.14	4.309	0.35	1.499	0.31	2.725	0.33	2.014	0.31	2.404	0.08	2.217	0.49	2.325	0.25
Mean	4.031		3.438		3.782		3.649		2.829		3.304		5.748		3.186		4.670	
Max	6.288		7.518		5.234		7.539		5.142		4.686		17.901		7.385		11.588	
Min	1.652		1.583		1.989		1.499		0.757		2.014		2.350		1.179		2.325	
Median	3.609		3.036		3.839		3.175		2.701		3.205		3.728		2.554		3.113	

Note: Frag 2 refers to the posterior fragment in purple on Pattern A (figure 2), and frag 3 refers to the anterior-lateral fragment in light blue. Both is the combined mean articular step off value

Note: Green highlight is the minimum step-off for that column, and red highlight is the maximum step-off for the column.

Note: The bottom four rows highlight the variability within each group, as the difference between the minimum step-off and maximum step-off can be quite large and due to a small sample size of 6, how a single malreduced fragment can significantly shift the group mean step-off.

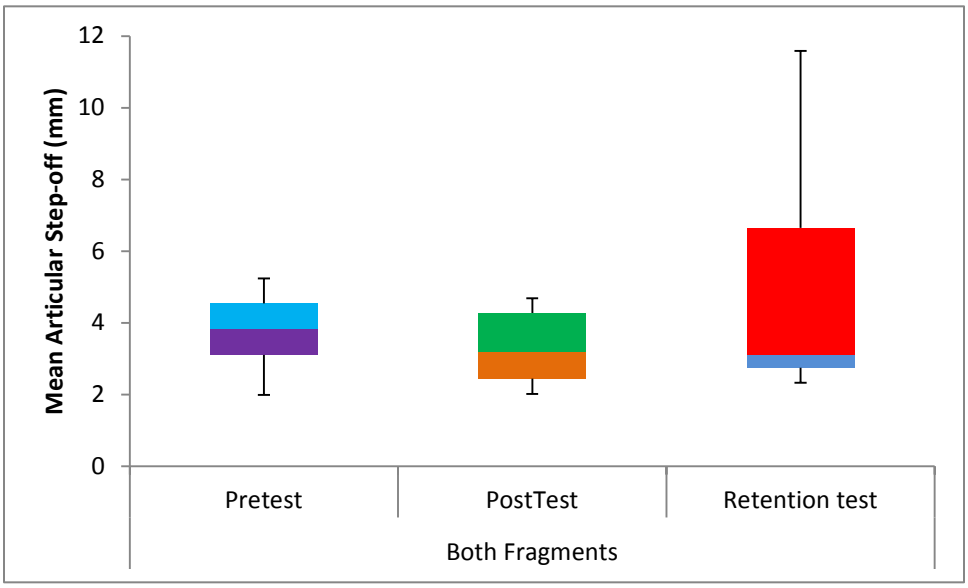

Figure 16: 2013 University of Iowa articular step-off box and whisker plot.

With use of the DEA algorithm, contact stress was computed for all eighteen trials of the simulation. Raw contact stress was considered for each of the thirteen steps of the stance phase of the gait cycle. Using subject Blue, we compared the pretest, posttest, and retention test contact stresses generated during step 10 of the stance phase; this is the step with the greatest loading, at 3.18 times body weight (600N for this contact analysis). Figure 17 shows the contact area for each of the three trials performed by subject Blue and what percentage of the contact area experienced greater than a contact stress level threshold. In figure 17, the posttest has the smallest percentage of contact area engaged at the higher contact stress thresholds. The retention had lower than the pretest engagement areas at high stress (>8 MPa), but the levels were greater than the post test at very high levels (>12 MPa) of stress. Building upon the work by Anderson et al[2], Brown et al[9], Hadley et al[98], Maxian et al[99], and Kern[94], the results would indicate that the lower levels of contact stress will lead to better clinical outcomes and that the results

from the posttest and retention would indicate that subject Blue improved their surgical skills, and will provide better outcomes than a surgeon who has no articular fracture reduction experience.

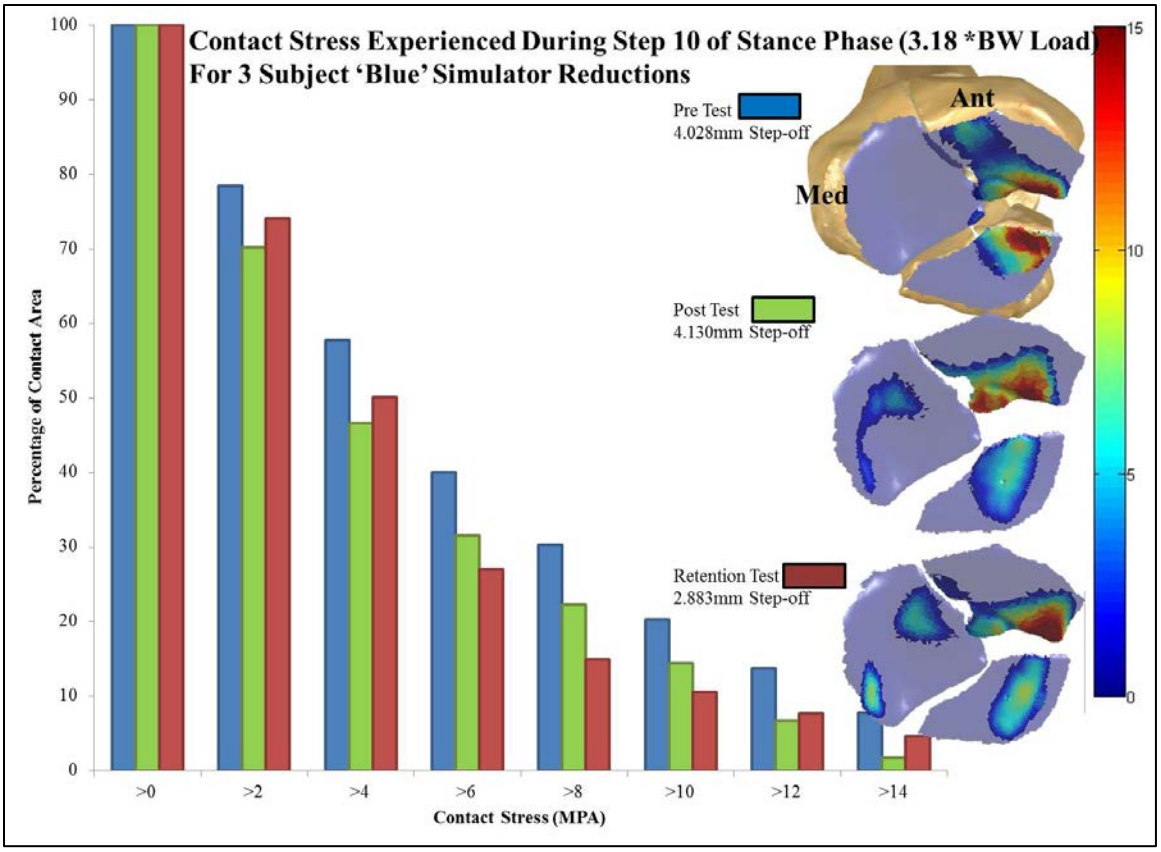

Figure 17: Contact stress experienced during step 10 of stance phase for subject Blue simulator trials. Contact Area illustrated on right.

Note: Step-off reported in figure is from MATLAB step-off analysis for comparison purposes

Note: The pretest contact area is overlaid on the digital fracture constellation reconstruction model from that simulation in order to orientate the contact surface

Note: The histogram contains the percent of contact area normalized to each individual trial's contact surface

In Figure 17**,** the MATLAB step-off value for the posttest is the highest of the 3 values, illustrating that although convenient, the practice of reporting a single value for step-off is not indicative of clinical outcome. In Figure 18, the step 10 loading percentage of contact area exposure was plotted for seventeen University of Iowa simulations and the ideal reconstruction. The ideal reconstruction has a greater decrease in percent of contact area engaged as contact stress increasing than the surgically performed reconstructions, and there is almost no area engaged that experiences greater than 8 MPa.

For the majority of posttest and retention tests there was a reduction of the contact area that experienced high contact stress compared to the pretest. It also appears that for most of the posttest trials a smaller contact area experienced the very high contact stress compared to the pretest trials. In terms of clinical outcome, the cases in which the contact area which experience high contact stress is small would have had better outcomes, based on the findings of Brown, Hadley, Maxian, Anderson, and Kern[2, 9, 94, 98, 99].

The contact stress time-exposure thresholds for the development of PTOA determined by Anderson et al and Kern [2, 94], were used to examine the DEA results as well. The contact area for the two thresholds of 4.1 MPa exposed for 5.1 MPa-s and 11MPa exposed for 1MPa-s were computed. Figure 19 shows the percentage of contact area that exceed the two thresholds, for each of the 17 cases from the University of Iowa 2013 simulation trials that were able to be processed by the DEA algorithm and the ideal reconstruction. There is a large variability among the results, and no clear overall improvement between the pretest and either the posttest or retention test. The ideal reconstruction has the smallest percentage of contact area that exceeds either threshold. There is also weak negative correlations between threshold contact areas and the subjects' OSATS scores, with the OSATS vs. 4.1MPA/5.1MPa-s Pearson's R = -0.233 and the OSATS vs. 11MPA/1MPA-s Pearson's R = -0.218, for the 17 cases able to be processed by the DEA algorithm.

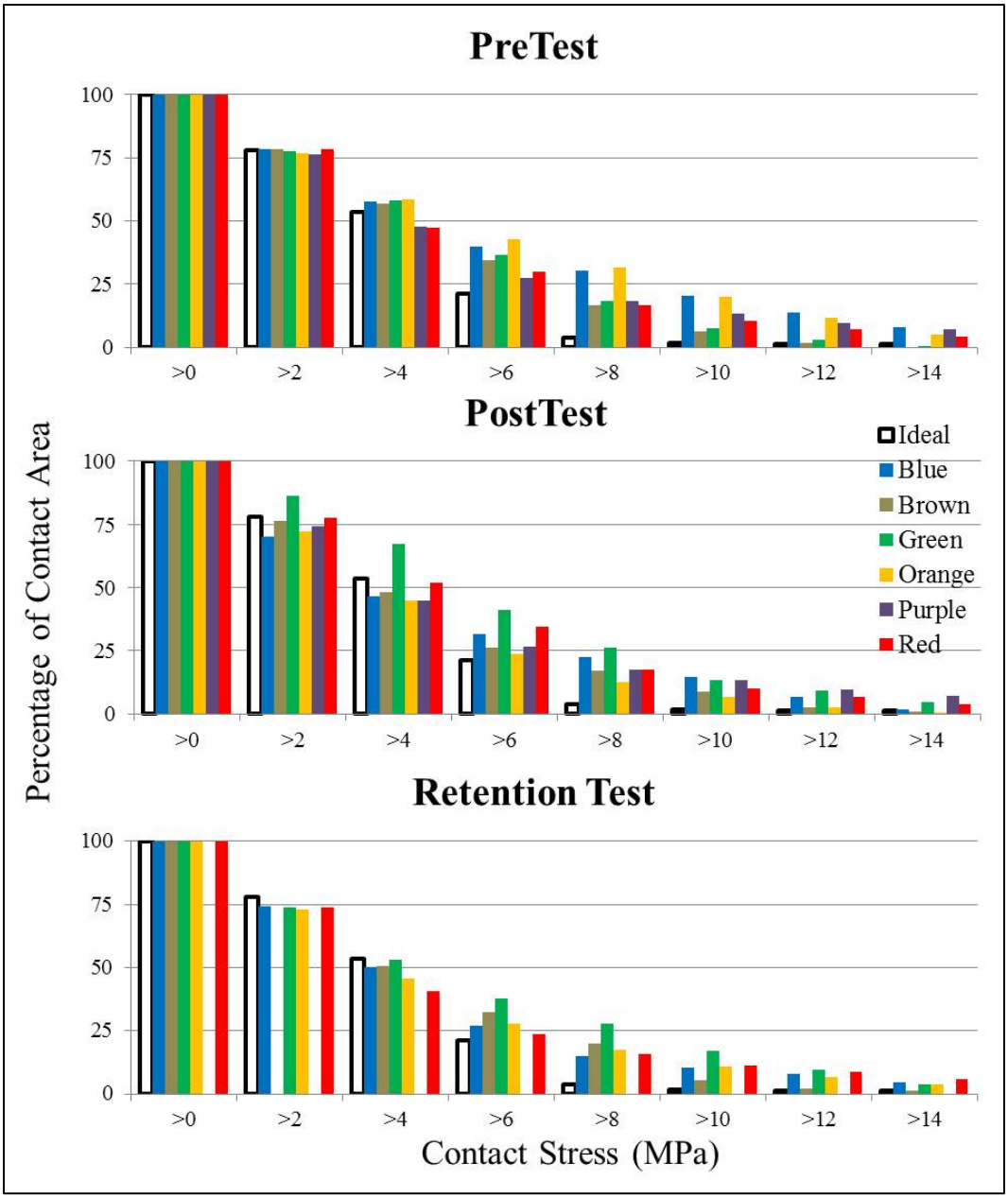

Figure 18: 2013 University of Iowa DEA step 10 of stance phase contact area engagement for a given contact stress threshold.

Note: Ideal reconstruction is included for comparison.

Note: Contact Stress area normalized for each trial individually

Note: Purple retention values not plotted as DEA could not be performed, due to gross malreduction

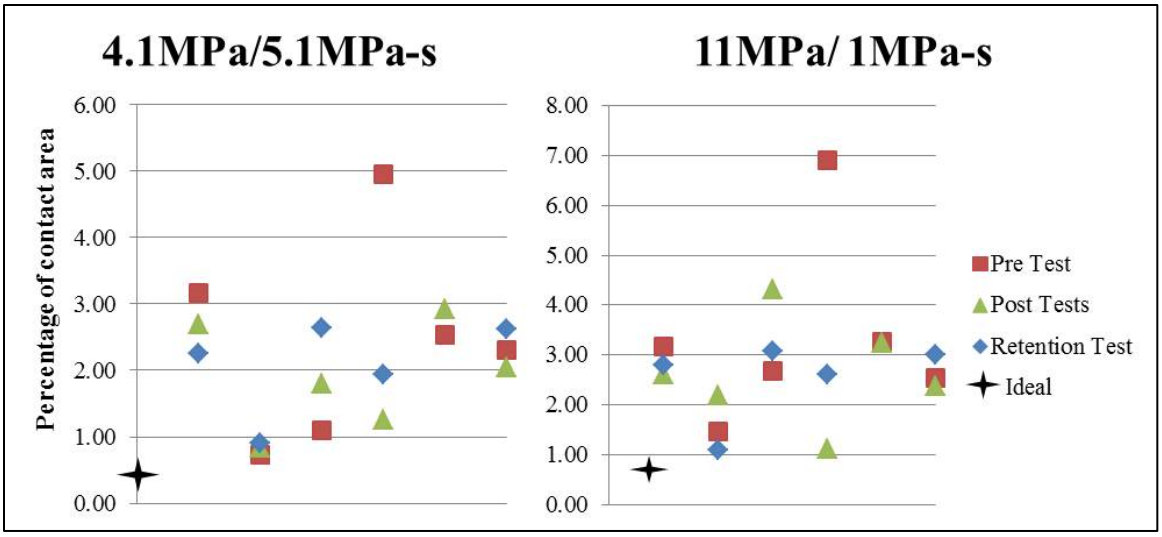

Figure 19: 2013 University of Iowa, percent of contact area exceeding contact stress-time exposure thresholds
Left: threshold value of 4.1MPa exposed for 5.1 MPa-s
Right: threshold value of 11 MPa exposed for 1 MPa-s

Note: Each vertical grouping is a single subject's 3 simulation results

Note: Pearson's R between the two thresholds is R=0.854

Note: Percentage of contact area scales are different between the two thresholds

The improvements in OSATS scoring, fluoroscopic usage, and contact stress exposure during the 2013 University of Iowa simulation trials indicate that the simulator does have a positive effect on the residents' surgical skills, supporting goals 1, 3, and 4. With a larger sample size, an improvement may be shown in the hand motion distance, number of discrete actions and the MATLAB articular step-off values. Larger sample size would also allow a clearer picture on whether the contact stress-time exposure thresholds can be used as a metric to differentiate between novice and expert surgeons.

4.3 2013 University of Minnesota Simulation Results

The 2013 University of Minnesota simulation trials took place in February and April 2013. The test included seven first year residents (R1) and eight second year residents (R2). In February, all 15 residents underwent the first attempt at the simulator (pretest). During the month of March, a senior orthopaedic traumatologist performed a one-on-one review of eight of the residents' first simulation performance during which coaching on surgical technique, fluoroscopic use, and time management was provided. All fifteen residents then performed a second attempt on the simulator in April (posttest). During the month following the posttest, the residents who were not provided coaching between trials were coached as to not hinder their development compared to their peers who previously received the coaching. Video was captured using the GoPro head camera and wide angle cameras located around the simulation area. Fluoroscopic data were recorded and counted images taken and radiation exposure. OSATS scoring was performed live by a senior traumatologist, blinded to resident year and which residents received coaching between simulator attempts. Motion capture of the hands was performed using the G4 system, for which the magnetic source was placed at a location, offset along the surgical table's short axis to the center of the table to ensure that the entire table was within the recommend range of the G4 system.

The University of Minnesota simulation trial fluoroscopic results unfortunately do not show any significant results. This however may be a result of previous knowledge provided during the Minnesota residents' standard training or natural skill, as the mean number of fluoroscopic images taken during the pretest was lower than the 2013 University of Iowa results pretest and posttest images taken. The Minnesota posttest mean number of images was 19.53 vs.15.73 for the Iowa retention test, which is not a significant difference by the student t-test for two-tailed homoscedastic groups. The only groups that are close to having a significant difference are the coached group which reduced the images taken from 25.25 to 18.75 and the uncoached group which increased

images from 15.14 to 20.43 (table 8). The Minnesota resident fluoroscopy usage during their pretest and posttest is similar to the Iowa residents after training. The Minnesota results show no significant improvement between trials. This lack of significance is expected and acceptable as a traumatologist attending from Minnesota completed an attempt on the simulator and only used 17 fluoroscopic images, within the range of images taken by the Minnesota residents and Iowa residents after training.

Table 8: University of Minnesota Simulation Fluoroscopic Results

Number of Images										
	Overall		R1		R2		Coached		Uncoached	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
Mean	20.53	19.53	21.43	21.00	19.75	18.25	25.25	18.75	15.14	20.43
SD	11.62	8.37	14.60	10.00	8.09	6.36	13.02	4.82	6.42	11.06
P-value		0.75		0.94		0.61		0.19		0.09
mAs										
	Overall		R1		R2		Coached		Uncoached	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
Mean	18.18	15.76	18.52	17.36	17.88	14.35	22.20	15.33	13.58	16.25
SD	9.59	5.88	11.92	6.85	6.93	4.43	10.66	3.81	5.25	7.56
P-value		0.31		0.80		0.19		0.09		0.17

Number of Images				mAs			
R1 vs. R2		Coached vs. UC		R1 vs. R2		Coached vs. UC	
Pre	Post	Pre	Post	Pre	Post	Pre	Post
0.80	0.56	0.11	0.72	0.91	0.36	0.094	0.781

Note: No significant results were observed for the fluoroscopic data

Note: For comparison between the pretest and posttest, a two-tailed paired t-test was performed

Note: For intergroup, same trial comparisons, a two-tail homoscedastic t-test was performed

Note: mAs is the measure of radiation exposure used

The Minnesota resident OSATS scoring showed significant improvement for all fifteen residents, with a p-value of < 0.001, for a mean improvement from 26.93 to 35.53

between the pretest and posttest. Figure 20 Illustrates the differences between the sub groups of residents, the R1 and R2 groups both improved, as did the coached and uncoached groups. The R1 group goes from a level that is significantly worse than the R2 group to having no significant difference (figure 20). The coached and uncoached group performed similarly in the pretest but the coached group performed significantly better than the uncoached group during the posttest (figure 20). These findings demonstrate the benefit that comes from practice on the simulator and from having feedback on the simulator from an expert, partially fulfilling goals 1, 3, and 4. The attending traumatologist did not wish to have OSATS scoring performed on their attempt, so no score was collected to compare to the residents.

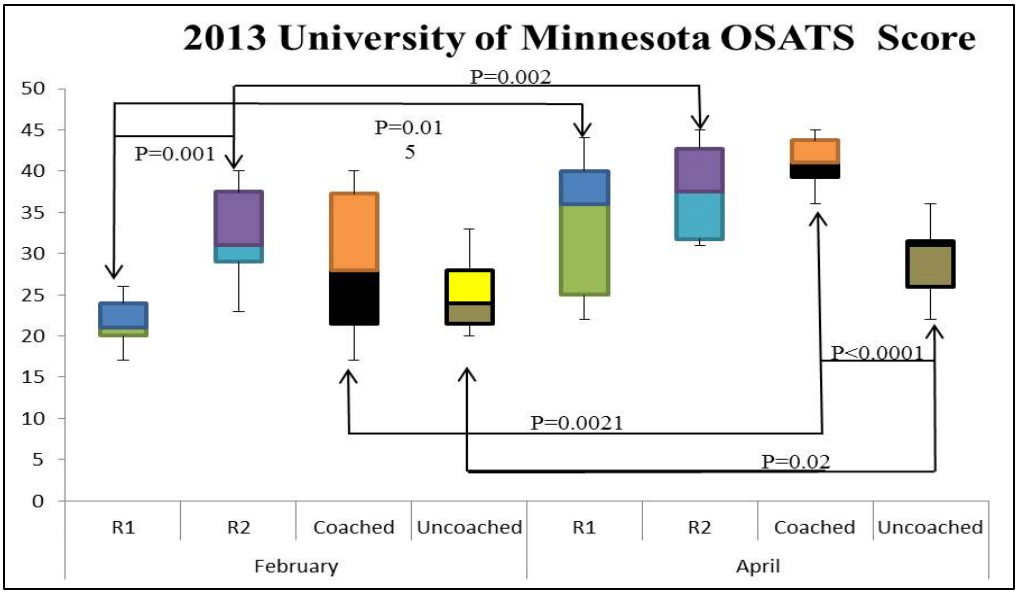

Figure 20: Box and Whisker Plot of 2013 University of Minnesota OSATS Global rating score. Significant results are indicated for a $p < 0.05$.

Table 9: 2013 University of Minnesota Hand Motion Results.

	Number of Motions									
	Overall		R1		R2		Coached		Uncoached	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Mean	920.4	586.27	810.43	603.57	1016.63	571.13	986.88	588.38	844.43	583.86
SD	316.82	168.81	360.28	199.99	234.07	133.95	304.09	145.83	313.99	191.70
P-Value		0.002		0.176		0.004		0.027		0.035
	Motion Distance (m)									
	Overall		R1		R2		Coached		Uncoached	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Mean	527.02	313.62	49973	311.31	550.90	315.63	584.81	317.58	460.98	309.09
SD	200.87	87.62	245.66	70.80	147.07	100.00	240.96	85.57	109.23	89.69
P-Value		0.003		0.088		0.021		0.032		0.023

	Number of Motions				Motion Distance			
	R1 vs. R2		Coached vs. UC		R1 vs. R2		Coached vs. UC	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
	0.24	0.77	0.422	0.617	0.16	0.93	0.734	0.864

Note: Significant results highlighted in yellow

Note: Comparison for a group between pretest and posttest performed using a two-tailed paired t-test

Note: Intergroup comparison performed using a two-tailed homoscedastic t-test

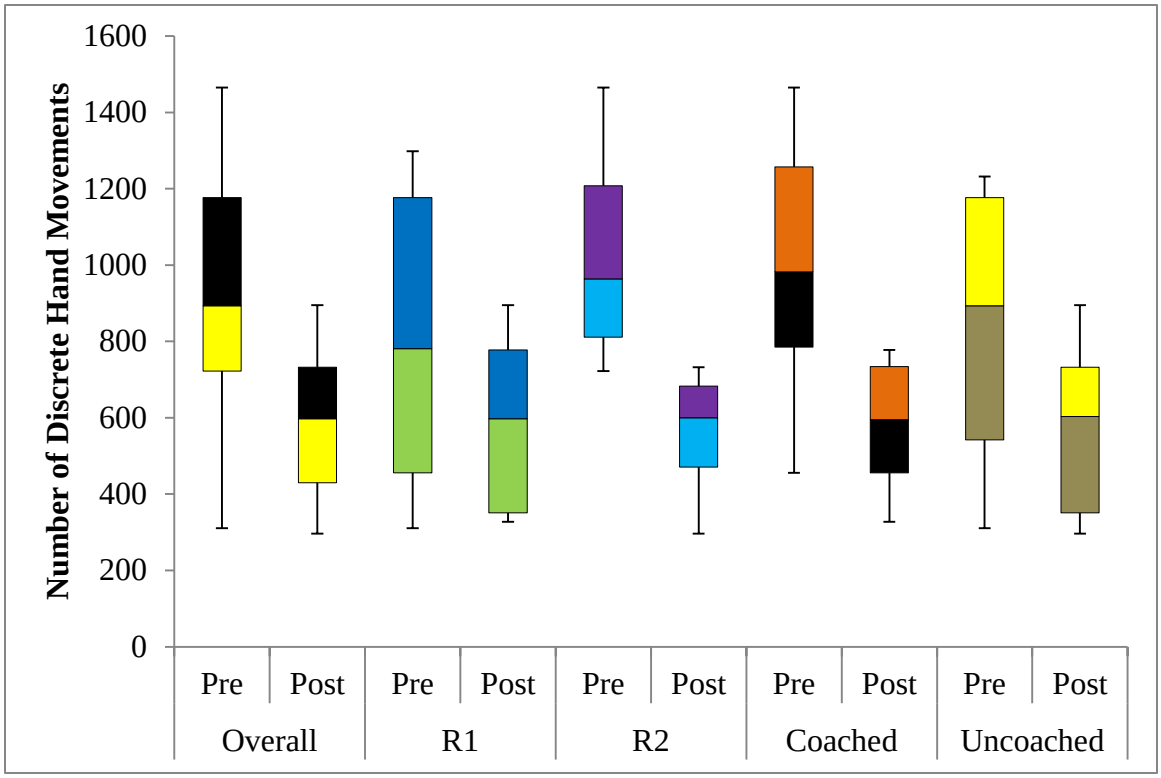

Figure 21: Box and whisker plots for 2013 University of Minnesota discrete hand movements from the simulator trials.

Note: Overall group consists of all fifteen Minnesota residents that participated in the simulation trials

Note: Groups with significant reductions in the number of movements include the overall group, the R2 group, the coached group, and the uncoached group.

Note: The R1 group has a visible reduction in movement that is not significant, but does trend down as four of the seven first year residents decrease movements, and two of the seven increased by <60 movements. The other resident increased by 287 movements.

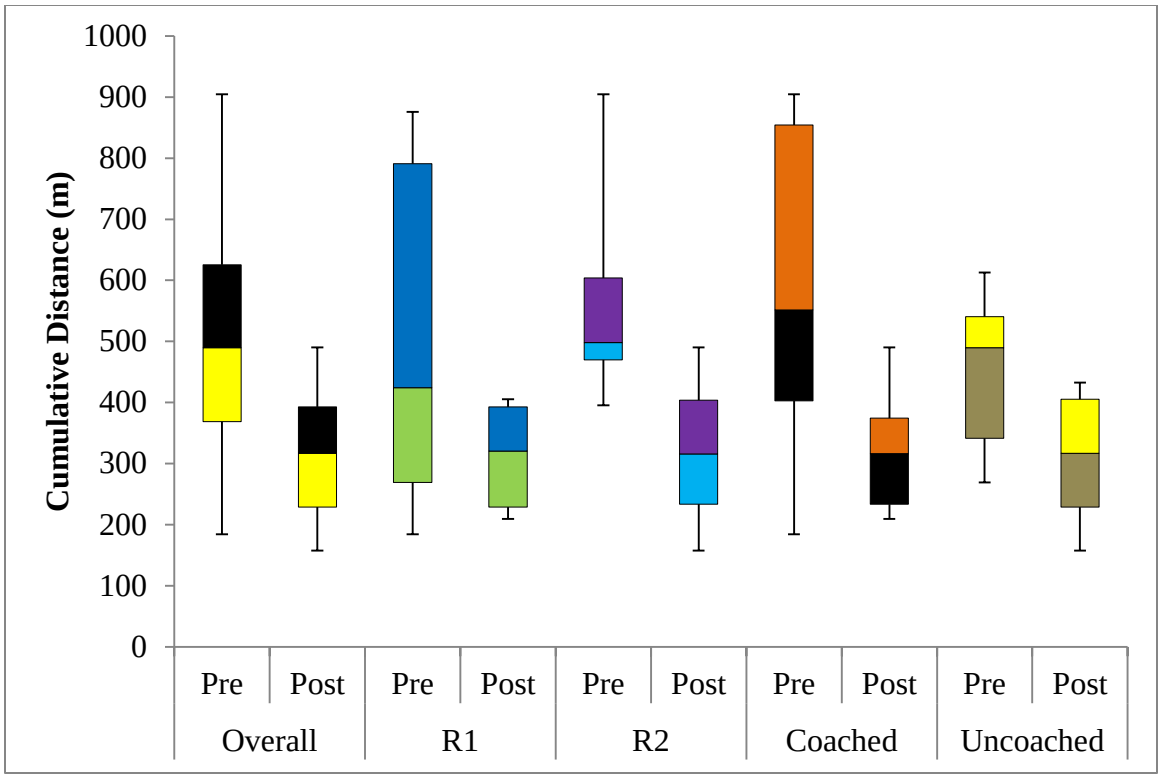

Figure 22: Box and whisker plot of 2013 University of Minnesota cumulative hand motion distance.

Note: Overall group consists of all fifteen residents who participated in the simulator

Note: Groups with significant reduction in distance include the overall group, the R2 group, and the uncoached group.

Note: The R1 group had a trend of decreasing the distance, and was close to significance. Five of the seven residents decreased the distance, one increased by < 5 m, and the last increased by approximately 140 meters.

The hand motion results also illustrate that the simulator provides skill benefits to the residents. Table 9 shows there is a significant reduction in the number of motions for all fifteen residents as a group, the R2 group, the coached group, and the uncoached group. There is also a significant reduction in the distance for all residents as a group, the R2 group, the coached group, and the uncoached group. The R1 group decreased both the

number of motions and the distance, just not significantly. As Figures 21 & 22 show, the variability for the residents as a whole and each subgroup became more consistent between performers. Table 9 also confirms the inter-resident increase in consistency as the standard deviation (SD) decreased across all groups as well. The attending performed 533 discrete motions and traveled a distance of 129.61 meters. The attending's numbers are consistent with the senior residents' testing in 2010. The Minnesota residents' discrete actions, during the posttest, are similar to the senior residents and attending, but distance is still much greater than the more experienced surgeons. This is similar to results for all three simulation trials with the University of Iowa residents.

Due to time constraints the articular step-off and contact stress results remain to be calculated for the University of Minnesota residents. The reduced fractures have all been CT scanned, and processed through the automated segmentation algorithm. The human post processing of the segmentations has not been completed.

4.4 G4 and Qualisys Motion Capture Comparison Results

Following completion of the 2011 simulator trials, a decision was made to change motion capture systems. The Qualisys optical motion capture system that had been used up to that point in time had several issues that led to this decision. First, the burden of transporting the entire four-camera system to each new location made the expansion to additional test sites more difficult. Second, the requirement that line-of-sight be maintained between the cameras and the hands being tracked was problematic, with substantial loss of data during simulations. Third, the time investment necessary for the post-processing of collected motion data (in part because of marker dropout due to blocked line of sight during the simulations) was restricting the frequency with which simulations could be done. Fourth, limitations in the Qualisys system related to reliance upon the on-board memory occasionally led to a loss in motion data.

The new motion capture system chosen was a Polhemus G4 electromagnetic motion tracking system. The G4 system was more compact, did not require maintenance of line-of-sight during data collection, required less time-intensive post processing of data, and did not rely upon on-board memory for data collection. However, the G4 system is susceptible to data corruption when ferrous metallic objects distort the generated electromagnetic field, and it has a fairly limited spatial volume for data collection. In order to ensure that the data collected by the two different systems was equivalent, a set of comparison tests was undertaken. The goals in these tests were (1) to ensure that the number of discrete actions counted in the data collected by both the Qualisys and G4 system were similar; (2) to ensure that the measured distance traveled were similar; and (3) to confirm that previous analysis data from the Qualisys system could be compared against future data from the G4 system. The tests performed here are part of the goal of assessing of new metrics for surgical skills.

During the comparison trials for the G4 and Qualisys motion capture system, the motions were counted live (Live). The day after the trial, another count was performed by watching the motion capture trace of the Qualisys data within the native Qualisys software at 50% speed (3D). The counts were compared against the MATLAB code results for each of the two systems as shown in table 10. Figure 23 also provides a comparison of the amount of hand motion as a percent of the number counted live. Except for the repeated circle motion test, the counts are quite close. The difference is attributable to jerky motion that is not apparent live, but when viewing the 3D motion trace becomes readily apparent. The discrimination between the 3D count and the two motion capture systems is due to the observer being forced to make arbitrary judgments on whether the motion is jerky enough to be split into two separate motions whereas with the motion capture data processing there is a clear and distinct threshold that is constant. The error from the circle motion is also easily attributed to a different issue, where during the live count each circle was counted as two motions, an up and a down. During the 3D

trace recount, jerky motion was not readily apparent during the circular motion so it too was counted as two motions per circle.

Table 10: The Number of Hand Motions During the Motion Capture Tests for Different Counting Sources

Test	Hand Motions (#)			
	Live	3D	G4	Qualisys
1-Static	2	2	2	2
2-UP & Down	26	29	36	31
3-Circles	26	26	77	55
4-Back&Forth	54	54	52	54
5-Random	190	206	183	194
6-Random	194	216	191	202

Note: Live refers to count performed while the motions were occurring.

Note: 3D refers to the count performed by watching the motion trace in the Qualisys viewer at 50% speed.

Note: G4 refers to the algorithm's calculation of motions using G4 Data.

Note: Qualisys refers to the calculation of motions using Qualisys data.

Note: The large difference between the circle motion counts is due to the way counting of motions was performed. For live and 3D each circle was divided into two motions (up arc and down arc) whereas the algorithm counted motion by local velocity minimums

Table 11: The Measured Distance Obtained by Each of the Motion Capture Systems for Each of the 6 Comparison Motion Tests

Test	Distance(m)	
	G4	Qualisys
1-Static	0.08	0.06
2-UP & Down	5.50	4.81
3-Circles	13.08	11.64
4-Back&Forth	11.12	9.86
5-Random	37.16	35.28
6-Random	37.80	36.09

Note: T-Test P value: 0.9011

Note: Mann-Whitney P Value: 0.5887

Note: Pearson's Correlation: 0.998

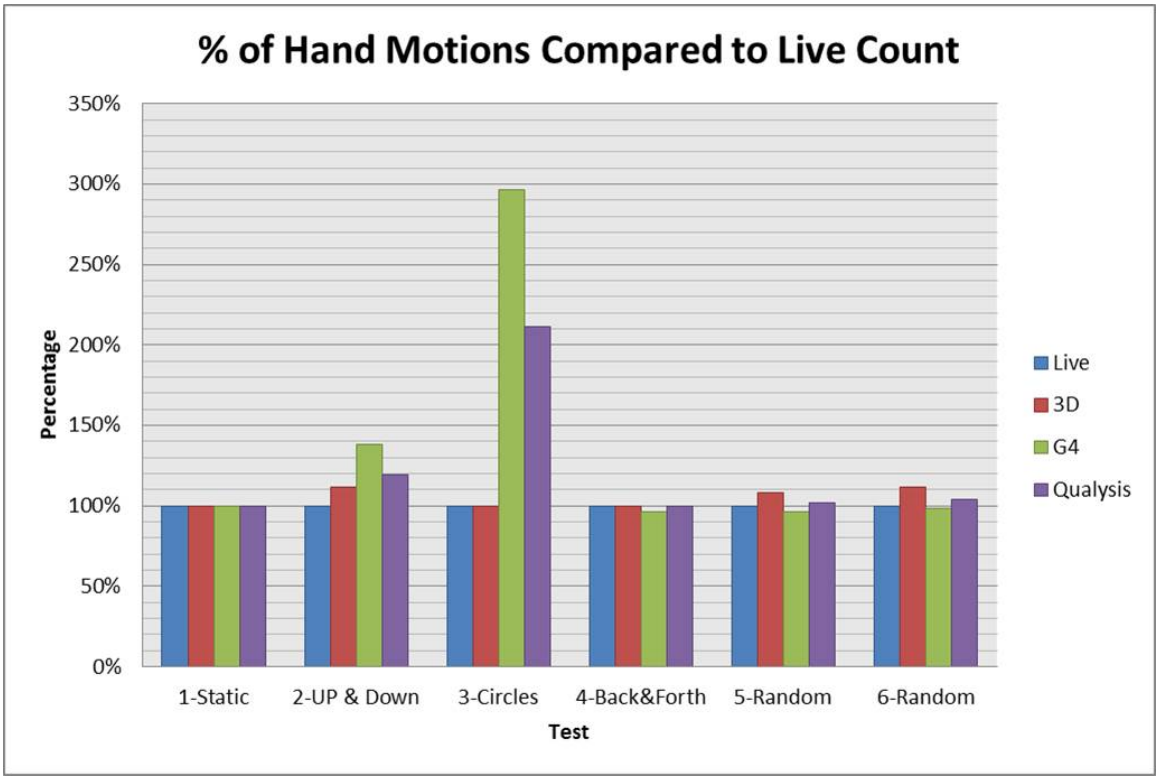

Figure 23: Percent of hand motions counted compared to live count for the two motion capture systems and a recounting done by watching the motion traces in a 3D digital viewer

Note: The large difference in the circle motion count is due to the human counting dividing a circle in to two hemi circle arcs, whereas the algorithm checked for velocity minimums to count motions.

The circle motion calculated counts are different from the human counts due to the algorithm counting discrete motions as local minima in velocity which can be caused the circle motion speed or radius changing while not being noticeable to the human eye, yet being apparent to the algorithm. This finding was not totally unexpected, as the circular hand motion would be the most problematic to analyze in terms of counting discontinuities in motion.

The distance both hands traveled was also calculated and is reported in table 11 The distance traveled during the static test is due to noise from both systems that is not filtered out, and is not significant compared to the distances covered by the other tests and during the simulation. Both the Student T-Test and Mann Whitney Rank Test indicate that the difference in motion distance between the G4 and Qualisys system is not significant. The Pearson's correlation coefficient r is also almost 1 indicating that near perfect correlation in that as distance traveled scales, both systems report consistently increasing values. For distance measurements, the Qualisys system was considered to be the gold standard, and with no significant differences between the Qualisys system and the G4 system, we can consider the G4 motion data equivalent to Qualisys results assuming that the G4 trackers stayed within an acceptable range of the G4 source.

4.5 Comparison of Articular 'Step-off' Results

During the iterations of the simulator, multiple methods of calculating articular step-off were used: First using the 3D comparison tool in Geomagic Qualify, second exporting the raw surface deviation data from Geomagic and importing the data into Microsoft excel and calculating the step-off, and third exporting the surgical simulation reconstructed articular surface and the ideal reconstruction, obtained through Dr. Thomas's puzzle solving algorithm, and importing it into MATLAB to compute the surface comparison.

The reason for originally using the Geomagic 3D comparison to compute the step-off was that it involved minimum user interaction and could be performed quickly once the surgical reconstruction was laser scanned, only taking 5-10 minutes of user time and 1 hour of compute time. The results gave a single number that due to the GD&T method of computing the mean deviation (step-off) gave values that were distorted from a normally computed mean, suffered from the fact the points on the surgical reconstruction were mapped to the closest point on the surface of the idea reconstruction regardless to

whether that ideal point was corresponding to a similar surface location or not, and could not be directly correlated to clinical outcome. This method also conveyed no information of the relative rotation of the articular surfaces and provided only a Euclidean distance. The excel method was a stop-gap that was only applied to the 2011 simulation results to see how much the Geomagic method of GD&T affected the final step-off value. The results in Table 12, show that except for three cases, the effects of using the GD&T mean was greater than 1 mm of step-off and up to 3.5 mm of difference, illustrating the artificial nature of the values. However, the excel method still suffered from the lack of true point to point comparison, lack of rotational information, and still has no direct clinical relevance.

The step-off calculation method developed in MATLAB eliminates the problem of not comparing the same point on the ideal versus reconstructed surface. By using a transformation of the same surface to position the articular surface of the reconstructed fragments into their ideal location and then comparing the Euclidean distance between each matching set, a true point to point step off measure is created. This also solves for the lack of accounting of rotational position since the distance calculation is independent of whether rotation of the reconstructed surface places the test point closer to another part of the ideal location. The MATLAB step off method does not, however, solve for how the step-off is directly clinically relevant. All that can be noted is that a lower number should be better. Histographic data in figure 24 illustrates the MATLAB method's ability to account for rotation as the histograms would be in just two bins (one for each fragment) if the reduction error was pure translation. A comparison was done between the Geomagic method and the MATLAB method for six cases from the 2011 simulation trial in which 3 sets of reconstructions with similar Geomagic step-off values were run through the MATLAB algorithm and the mean and standard deviation for each of the 6 cases was found and the deviation spectrums were plotted as shown in figure 25 and figure 26.

Table 12: The articular step-off values from the 2011 simulation trials for the 12 residents for both the months of November and December, using both the Geomagic and Excel methods and calculating the difference in mean value and standard deviation for each method for each trial.

Articular Step-off values for 2011 Simulation in Geomagic and Excel						
	Geomagic		Excel		Difference	
	Mean	SD	Mean	SD	Mean	SD
November						
Alpha	3.25	2.54	1.77	2.10	1.48	0.44
Echo	6.70	4.52	3.35	3.21	3.35	1.31
Golf	3.96	4.97	2.80	2.31	1.16	2.66
India	6.24	1.49	6.24	1.49	0.00	0.00
Kilo	6.08	4.31	3.30	3.33	2.78	0.98
Lima	6.02	3.75	3.09	3.00	2.93	0.75
November	5.30	3.06	3.48	2.77	1.82	0.29
Papa	3.86	3.06	1.88	1.38	1.98	1.68
Sierra	3.69	2.01	1.82	0.95	1.87	1.06
Uniform	3.71	2.37	1.93	1.65	1.78	0.72
Whiskey	5.92	3.77	3.05	2.49	2.87	1.28
Zulu	3.24	2.23	1.85	1.75	1.39	0.48
December						
Alpha	3.39	2.20	1.70	1.66	1.69	0.54
Echo	3.66	2.00	1.99	1.24	1.67	0.76
Golf	3.02	2.03	1.51	1.37	1.51	0.66
India	3.93	2.00	1.80	0.90	2.13	1.10
Kilo	3.54	1.68	1.70	1.14	1.84	0.54
Lima	2.89	2.01	1.50	1.46	1.39	0.55
November	6.63	3.07	3.10	1.01	3.53	2.06
Papa	4.52	2.38	1.86	1.50	2.66	0.88
Sierra	3.48	1.67	2.15	1.01	1.33	0.66
Uniform	3.28	1.35	2.70	1.26	0.58	0.09
Whiskey	1.81	1.28	0.90	0.92	0.91	0.36
Zulu	5.64	2.79	2.56	1.38	3.08	1.41

Note: The three smallest differences are highlighted in green and the three largest differences are highlighted in red.

Note: The six cases used for comparison of the MATLAB and Geomagic Methods are highlighted in yellow.

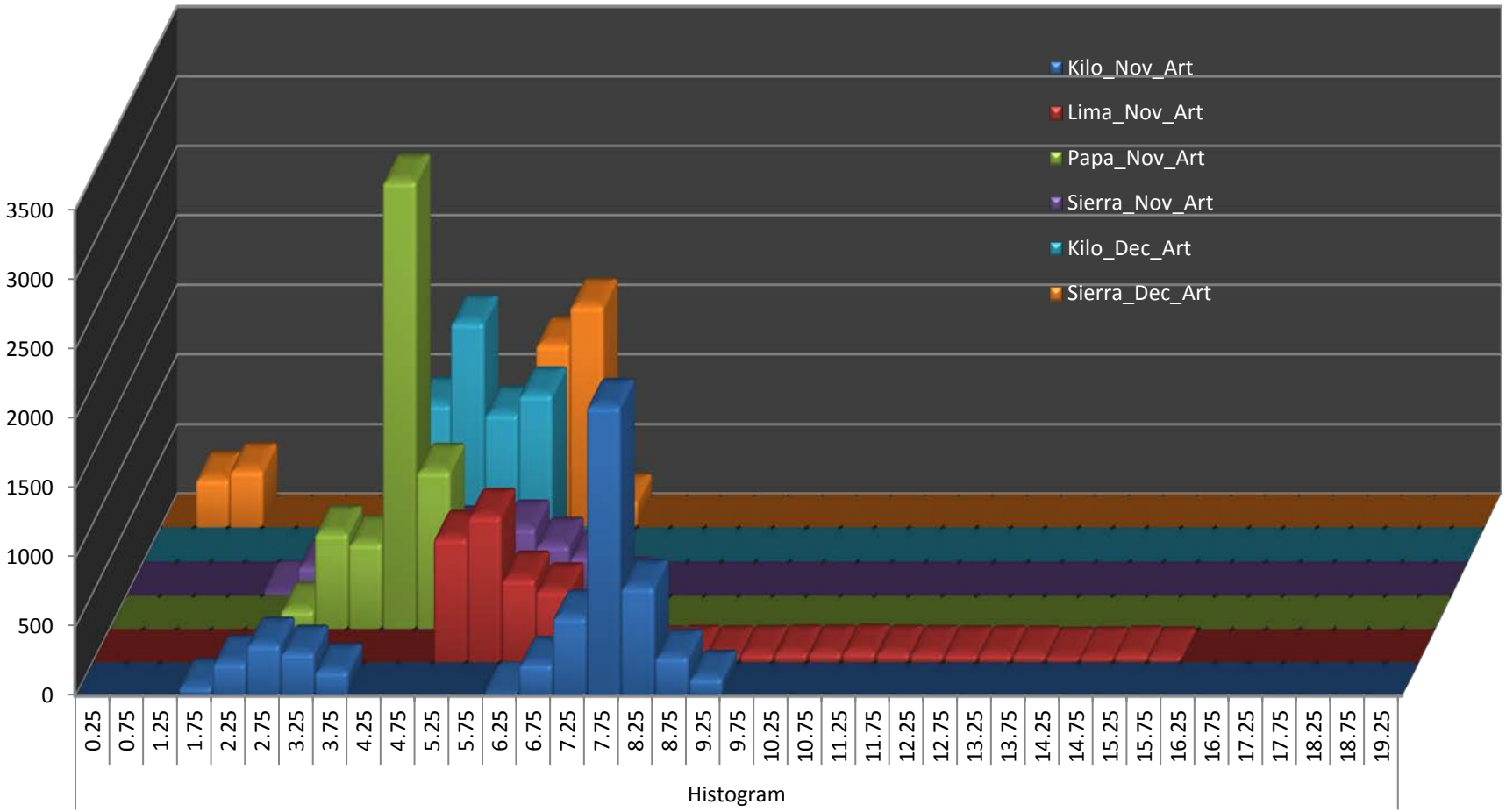

Figure 24: Histogram data from MATLAB Point to Point comparison of Step-off for the articular surface of six 2011 simulation trials

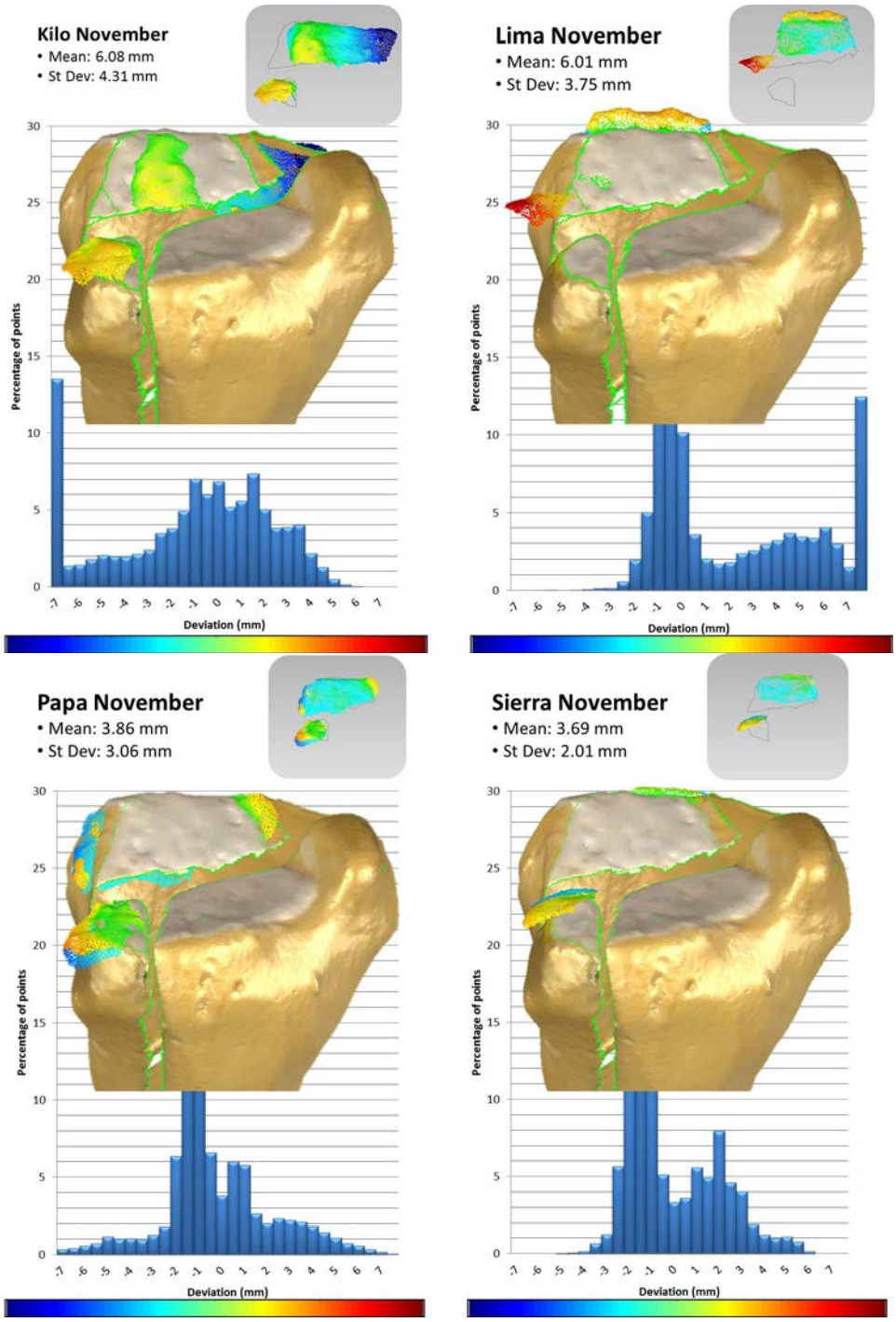

Figure 25: Six 2011 simulation cases from Geomagic that show articular surface mapping of deviation on ideally reconstructed distal tibia in the center of each sub image.

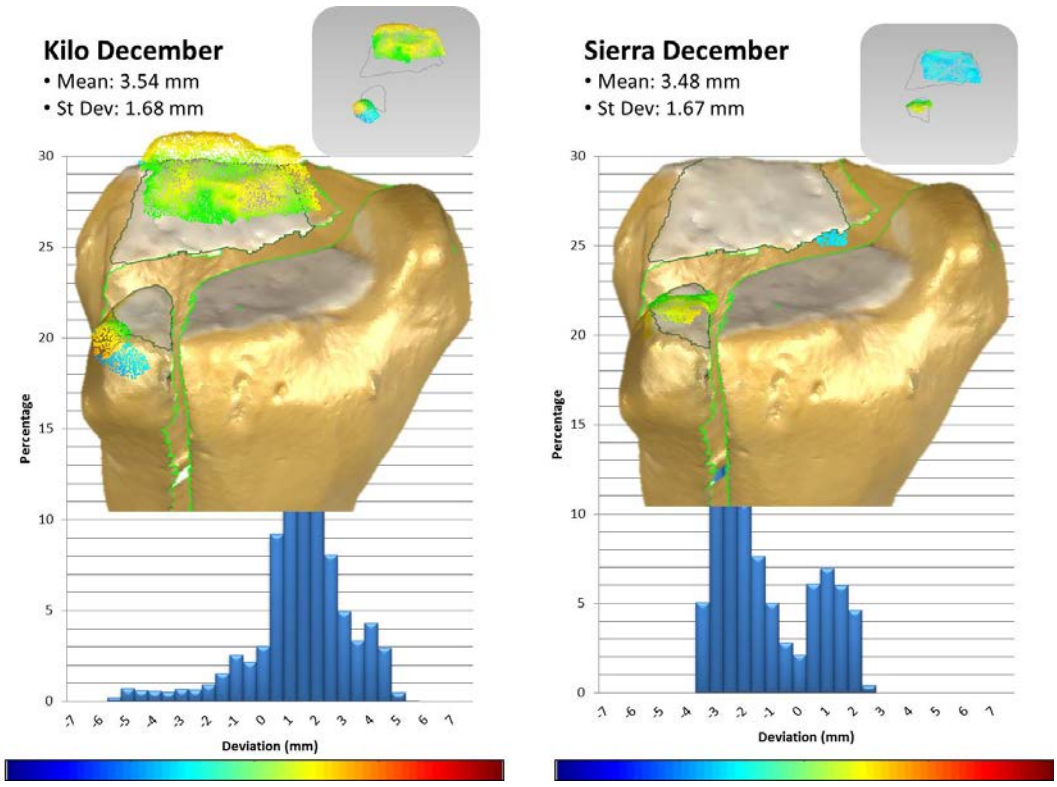

Figure 25: Continued

Note: Geomagic method uses test point to nearest surface point to compute deviations

Note: In the upper right, the same deviation is mapped with only an outline of the ideal articular surface for each of the two fragments.

Note: Behind each image is the histographic data for that particular reconstruction.

Note: The subject name and month performed, mean and standard deviation obtained from Geomagic are reported in upper left of each sub image.

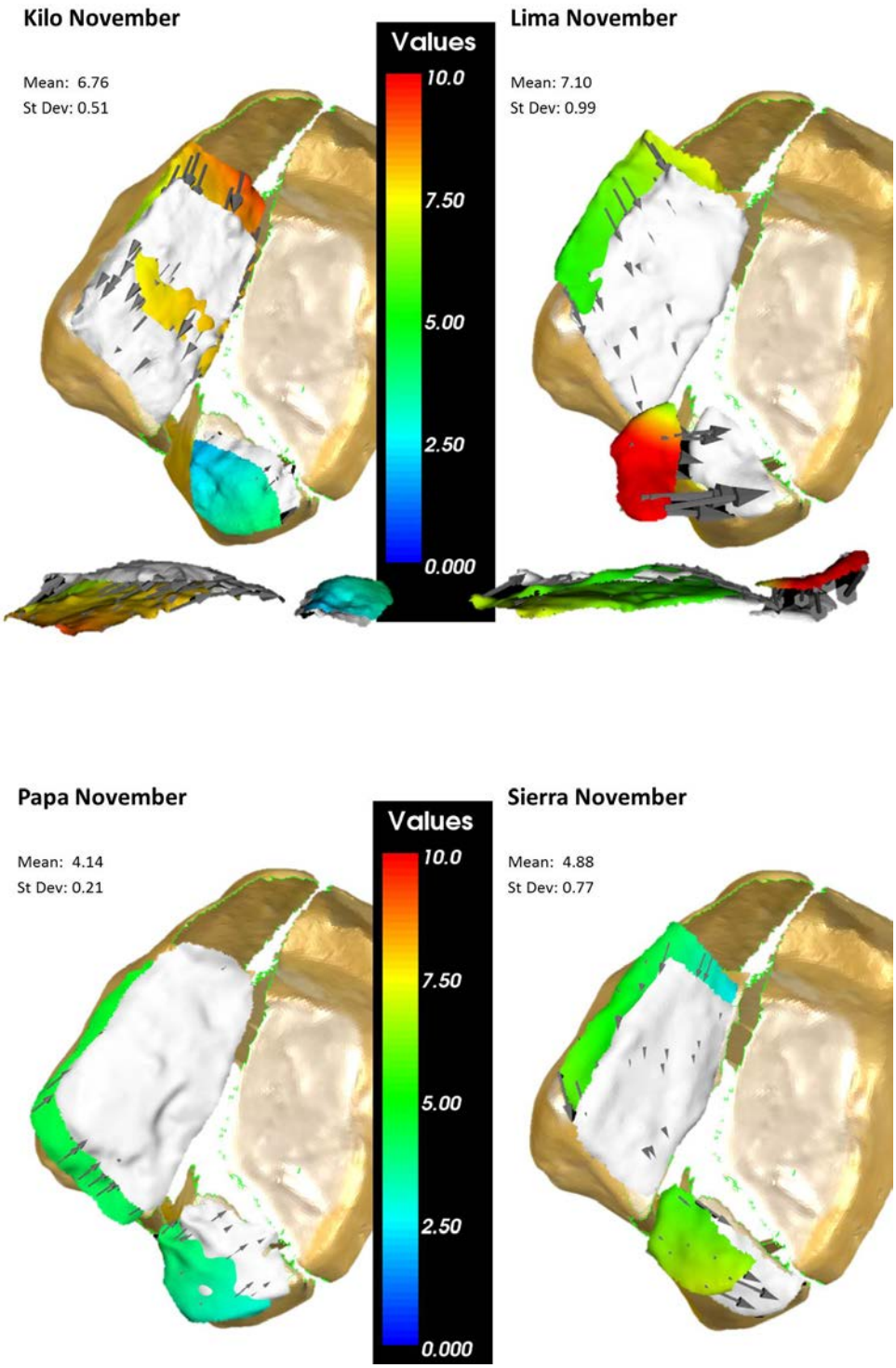

Figure 26: MATLAB point-to-point step off mapping showing the deviation spectrum on the surgical reconstructed surface and the white being the ideal reconstruction location mapped onto a rendering of the ideal tibia articular surface and in an axial plane view (Medial to Lateral).

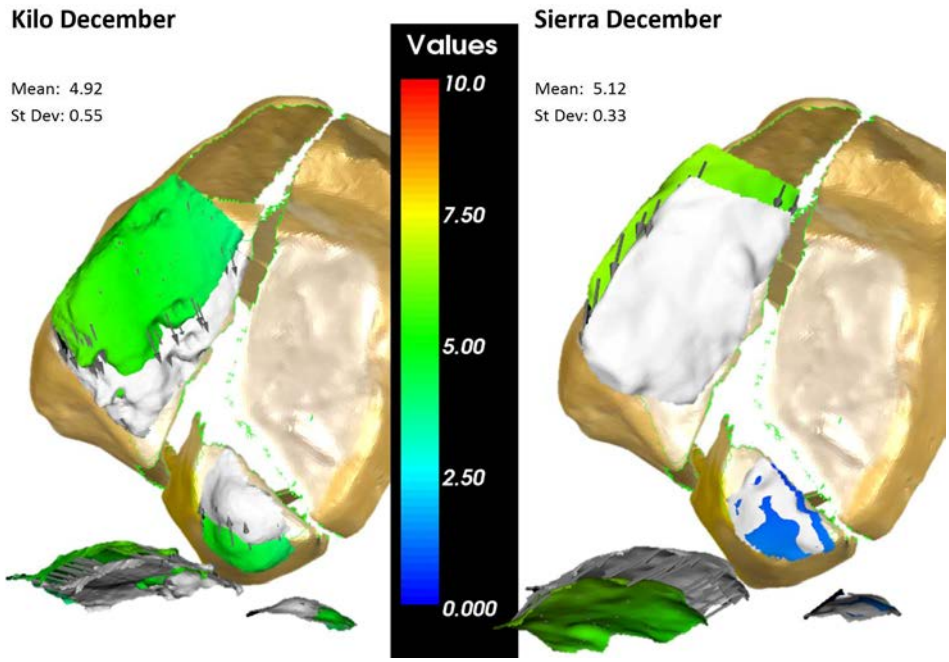

Figure 26 Continued.

Note: The ideal surface shown is the laser scanned surgical reconstruction surface mapped to the ideal location of the fragments articular surface

Note: The arrows shown highlight the direction and distance that the reconstructed surfaces needed to be moved in order to obtain ideal placement.

Note: Case Identification and Step-off mean value and standard deviation are also reported.

In figures 24 and 25**,** we can see that the some of the cases have a bimodal distribution of deviation, this is most likely due to the difference in how well each of the two floating fragments were reduced, as one could be well reduced without ensuring that the other has been reduced satisfactorily. With figures 25 and 26**,** the difference between a true point-to-point step-off measurement and a point-to-surface is readily apparent as the point-to-surface can significantly alter the measured deviation. With all six cases we see that the mean step-off measure in the MATLAB point to point differs at least 0.5mm from the reported value with the Geomagic point to surface. In the Sierra December case

we see a mean value of 5.12mm of step-off for MATLAB, 3.48mm for Geomagic, and 2.15mm for Excel. This illustrates the artificial lowering of the step-off measurement, using the point to surface comparison especially when not using the GD&T mean calculation found in Geomagic, since the algorithm just searches for the nearest point on the reference surface and not the point that corresponds to the test point. In the Lima November case, the smaller fragment clearly illustrates the problem this causes as the surgical reduction placed the small fragment nearer to the ideal location for the larger fragment so the algorithm calculated the distance to that surface rather than the correct fragment articular surface, shown in figure 25.

Of the three articular step-off deviation analysis methods performed, the point to point algorithm used in MATLAB provides the most relevant step-off value as it is not sensitive to gross malreduction and provides the most accurate mean step-off value. This development of an articular deviation calculation method was driven by the goal of developing and assessing new metrics to measure surgical performance. Although all three methods are more precise than the current clinical practice of using a ruler to measure the distance between fragments' articular surfaces on a radiograph or a single distance measurement from a CT scan, none provide a true indicator of long term patient success that is possible with the contact stress analysis developed by Anderson et al and implemented using DEA by Kern [2, 94].

CHAPTER 5 –IMPACT AND LIMITATIONS

5.1 Tibial Plafond Fracture Reduction Simulation's effect on Orthopaedic Surgery

Over the last five years, there has been a renewed interest in simulation in the orthopaedic community. The new guidelines mandated ensure that simulation will play a larger role in R1 education. When Dr. Marsh presented the early work with the simulator to the AOA council of residency directors[133] and to the AAOS[130], the community seemed very receptive to a simulator that could teach articular fracture reduction skills. A national survey, conducted by Dr. Karam et al [134], indicated that a majority of residency program directors and a larger majority of residents believe that surgical skill is not currently measured objectively with orthopaedic surgery. Over 80% of both residency directors and residents believe that surgical skills simulations should be a required part of training, and over three quarters believe it should be standardized[134]. This means that a simulation that has been validated to show improvement on objective metrics would be accepted by the orthopaedic community and readily integrated into the national curriculum requirements. Many companies are developing commercial simulators that have not been validated, are not scored using well defined metrics (most use a custom developed, untested scoring systems), and are expensive.

The fracture reduction simulator is poised to become a part of standard resident education due to several factors. The first factor is the cost; compared to currently available simulators, the fracture reduction simulator is inexpensive. For approximately an initial cost of six thousand dollars used to purchase a PC and laser scanner and an ongoing cost of approximately 500 dollars per resident (two sawbones models and fluoroscopy time), an institution can have the simulator up and running. Motion capture equipment is not needed in the clinical setting as other metrics have shown to demonstrate skill improvement more clearly in resident performance. Also the motion

capture systems represent a significant capital investment (>$10,000 for the G4 system and >$50,000 for the Qualisys) that most orthopaedic programs could not or would not pay. The residents would still be judged on fluoroscopic use, OSATS scoring, and articular step-off / contact stress.

The second factor is the skills taught: fluoroscopic usage, surgical tool handling, closed reduction fragment manipulation, fracture reduction techniques, care of soft tissue, k-wire placement, and surgical time management. In terms of fluoroscopic usage, residents learn to read the images in order to form the 3D model of the fracture site; this involves using subtle shadows present on the image to determine position of the fragments and being able to stitch together a lateral image and an anterior-posterior image into the 3D mental model. The tool handling, fragment manipulation, reduction techniques and care of soft tissue are intertwined skills, which facilitate a surgeon's ability to cause a positive outcome for patients. K-wire placement is an important skill as a few well-placed k-wires will better hold a fracture, during plating or insertion of surgical screws, than many poorly placed k-wires. Surgical time management is important because the longer a patient is under anesthesia the more risk there is, also the longer the surgical incision is open there is a greater chance of infection occurring. Having a time limit teaches the surgeon to prioritize their approach and have a pre-operative plan. The time limit also forces the young surgeons to realize that 'perfection is the enemy of good enough,' in that although the reduction they achieve could be better, by trying to better reduce the fracture, the risk of making it worse is high. This is demonstrated during the retention test for subject Purple during the 2013 Iowa trials.

The third factor is objective scoring. The fluoroscopic usage is straightforward; did the resident take more or less shots? Did the resident cause a higher exposure? OSATS scoring has been validated as semi-objective across multiple surgical procedures and simulations outside of orthopaedics, and a higher score correlates with better patient outcomes, and other objective and semi-objective metrics such as hand motion in the

ICSAD simulator and the OSCE clinical metric. OSATS is a way to measure the knowledge gained through experience that traumalogists value, and believe to be necessary to perform a successful reduction. Step-off is a clinically used metric to gauge surgical success, and it is measured on a single comparison done by radiograph or CT scan. The MATLAB point-to-point algorithm provides a more robust metric that is not subject to location bias as is the clinical method. The algorithm provides a mean and standard deviation that account for both translation and rotation of the fragments. It provides an objective measure, as a higher mean and/or standard deviation is worse than lower values. The contact stress metric is also objective, as a high percentage of surface area experiencing a high contact stress is a worse than a lower percentage of contact area. Previous work showing that high contact stress and high exposure lead to worse patient outcomes leads to the ability to say that a reduction with high stress or high exposure is not as well reduced as one with low stress or low exposure, without bias. Also, the exposure thresholds found by Anderson et al[2] and Kern[94] indicate that it may be possible to predict the PTOA outcome of the simulator reductions.

Since there are multiple objective performance metrics available with this simulator, once a standard competence threshold is established, it would be possible to use this simulator as a part of board certification or recertification. Using objective metrics that directly measure surgical skill would give the public reassurance that a board certified orthopaedic surgeon has a minimal level of competence to successfully reduce the patient's fracture. This would alleviate the concern present when not being able to pick which surgeon performs the reduction as is common within other surgical fields. Although it would not be needed, hand motion data could also be captured. As the certification board would only need a few systems, with the cost spread over all orthopaedic surgeons, it would not be too burdensome to purchase and use during certification, providing one more objective metric to judge competence. This would also allow for post-stroke or other disease stricken surgeons to be tested to make sure the

surgeon's still have the physical abilities necessary to perform surgery. Tests such as the ones used to compare the two motion capture systems could be used to determine how fluid/jerky the surgeons' motions are, since differences are present between stroke and normal motions[85].

5.2 Future and Concurrent Work

The articular fracture reduction simulator step-off and contact stress analysis for the 2013 University of Minnesota residents will be completed in the fall of 2013. The new group of orthopaedic residents at the University of Iowa will also participate in the simulator during the year one surgical skills month as done in January 2013 for the 2012 residents [130]. Plans are being made to take the simulator to other teaching hospitals around the Midwest, including returning to the University of Minnesota. This will allow for a larger population in order to test for improvement. Plans are also to have multiple traumatologists perform OSATS scoring on the same simulations and find inter-observer correlation, and to have the same grader rerate a group to determine intra-observer reliability. Plans to have a group of five to ten traumatologists who are considered experts by their peers complete the simulation in order to develop standards for the metric are also being formed. In future simulations, the plan is to move the location of the magnetic source to the underside of the center of the surgical table, in order to ensure that the hand motions occur within the magnetic field without interfering with the resident's ability to perform the simulation. Ongoing work also includes breaking the fracture reduction procedure into specific subtasks and developing trainers for those skills.

The counting of k-wires placed during simulation as a metric was considered. However, since k-wire usage is accounted for in the OSATS scoring and the fact that different number of k-wires may be placed and still create a stabilized joint and successful reduction, the direct use of the k-wire count was not used. Another considered metric was time to completion. However, during the 2010 experiment with a ½ hour time

limit, no resident exceeded twenty minutes, and most finished in between 13-15 minutes. This 13-15 minute time was noted by multiple trauma faculty members at the University of Iowa to be typical of the time it takes an attending to perform the procedure. A 15 minute time limit was implemented for the rest of the studies to ensure that the residents would perform the procedure in a similar time to the senior surgeons.

Due to the time involved in performing a laser scan or CT on the reconstructed tibia surrogate, a method to create a better step-off measurement than that currently obtained clinically is needed. Since during the clinical procedure, two final fluoroscopic images are taken from the AP and Lateral views, taking a mean of 10 step-off measurements from each view could provide a more realistic value for surface deviation that is still quick and cost-effective to perform clinically.

5.3 Limitations of Tibial Plafond Fracture Reduction Simulation

The single greatest limitation in the evaluation of the simulator is the limited number of orthopaedic residents that have used it to date. The University of Iowa accepts six new residents each year, and the University of Minnesota accepts eight. Once the residents have participated in a trial of the simulator, the residents cannot be considered true novices anymore, as they have experience with the simulator and articular fracture reduction. Also once the residents who have participated in the simulator as junior residents have reached the last year of their residency, and are senior residents, they cannot be used since they will have an experience bias which could taint the results. The small sample sizes limit the power of the studies, and make it hard to find significance in the small changes that occur during improvement from a novice surgeon to expert.

Another limitation is cost: financially, time-wise, and with respect to radiation exposure. The lower leg and foot with embedded fracture costs around $200 from Sawbones. There is also the monetary cost of booking use of a fluoroscopy machine and

technician. The motion capture system also takes time to set-up and calibrate along with instrumenting the residents, as well as the time to process the results from the simulation. A senior traumatologist must either take time to score the simulation live or by watching the video, as well as coaching the residents. In the current form, the simulator relies on the use of a fluoroscopy machine which emits radiation to create the images. Current OSHA standards and University of Iowa policy limit the amount of radiation that a person may be exposed to, limiting the number of times that a resident could perform the simulation. Along with the radiation, the use of a technician to run the machine means that a resident could not perform the simulation in his or her spare time.

In terms of data collection, limitations include the issues with the motion capture system, segmentation accuracy, small sample sizes, OSATS scoring, and differences in natural skill or previous training.

Each of the two motion capture systems has limitations. The Qualisys system requires line of sight for data collection, which means that anytime the fluoroscopy occluded view of the markers, no data were collected. The line of sight requirement also means that the markers could be obscured by a tool, the surgeon's body or other hand, or the ankle surrogate. The Qualisys also had persistent issues with the onboard memory not transferring the motion capture data to the laptop for analysis. The G4 system has the issue of being sensitive to ferro-magnetic material within the magnetic field. Another problem with the G4 is that the system has accuracy issues when the sensor is near the edge of the magnetic field. A third issue with the G4 system is a variable frame rate that must be accounted for when calculating distance and discrete motion. The accuracy of data from the 2013 Iowa simulator trials obtained using the G4 system was compromised by errors in measurements associated with the measurements being made at the edge of the recommended spatial volume (a function of distance from the electromagnetic source Although trends in the hand motion data are here noted, any conclusions drawn must be interpreted with caution. Although the recommended maximum distance for accurate

measurements was 6 feet, during data analysis, an observed maximum of 4 feet was shown to maintain accurate measurements. This limitation was corrected for in the Minnesota simulations as the source was place directly adjacent to the surgical area, ensuring that the motion envelope would be well within 4 feet.

In terms of creating the digital surface model for calculating articular step-off or contact stress, the laser scanner has an accuracy of 0.005 inch or 0.127 mm. The CT scanner has a voxel size of 0.37 x 0.37 x 0.37 mm, meaning that the best surface representation has an error of approximately 0.37 mm. The segmentation of the CT scan is also subject to the parameters set up in the automatic segmentation and the human accuracy during refinement. This error will affect the step-off measure, but since the range of step-off is 2 mm to 10 mm, 0.37 mm is not a large enough error to be concerned about, as the error will be minimized by using the mean step-off since the error in different directions will cancel out some of the effect.

The small sample sizes for the trials means that a single outlier value can substantially distort the mean and standard deviation. A single bad performance by one of the residents could negate significance for the test metric. The small sample size also means that to find significance, a bigger change in values must happen; this is an issue when some of the differences that separate experienced surgeons from novices are quite small. For example, in terms of step-off, ½ mm of change could have a profound effect on contact stress, but with less than 20 subjects would most likely not be significant.

The OSATS score is a validated metric for measuring surgical performance, yet it is not without its faults. The scoring is performed by an individual, who can make mistakes, such as letting inadvertent bias influence the scoring, becoming distracted during the scoring and missing something that is important to the scoring, have personal ideas of what is important that may cause a different score compared to other raters, being more generous or withholding of positive scores. During the 2011 and 2013 University of Iowa trials, the OSATS rater was also the expert surgeon who performed

the coaching of the residents. The rater had also previously had professional exposure to the residents and may have had preconceived perceptions of which residents were more skilled.

Even though the residents were all located at either the University of Iowa or the University of Minnesota, they are a diverse group. Coming from different backgrounds and medical schools, which may have influenced the skills presented during the simulation. Non-professional experience, such as a carpentry hobby, may have led to an increase in skill in certain residents. Baseline abilities that differed between residents may have also affected the results. This is would be corrected for with a large population as the skill level difference would be averaged out.

CHAPTER 6 – CONCLUSION

The articular fracture reduction simulator, presented within this thesis, represents the first step towards creating a comprehensive surgical skills curriculum that can improve resident abilities in orthopaedic trauma while objectively measuring their improvement. The simulator builds upon the ideas of previous surgical simulators that use an objective rating system that has been validated for differentiating between skills levels regardless of rater, such as the MISTELS and ICSAD laparoscopic simulators. The analysis of the fracture reduction within the simulator has led to the development of a more robust measurement for articular step-off that accounts for both fragment translation and rotation. Building upon previous work within the University of Iowa Orthopaedic Biomechanics Laboratory, the new functional analysis of the articular reduction computes contact stress using DEA for an objective metric of the quality of reduction and the risk for development of PTOA.

The simulator and the approach to assessing its value in facilitating skills acquisition have some clear limitations. The inclusion of larger numbers of residents in future studies would allow for a more definitive analysis of which metrics best reflect competence. The need for fluoroscopic images will also be a limiting factor in how many times a resident could use the simulator to improve their surgical skills, due to the radiation emitted during image generation. Even though the use of fluoroscopy during a typical procedure by an experienced surgeon only causes approximately an extra day's worth of background radiation, the standard of care use of CT and radiographs to detail this injury necessitate any possible reduction in radiation exposure.

The articular fracture reduction simulator, in its current form, meets the objectives that were originally laid out. The ability to differentiate between junior and senior residents seen in the 2010 experiment indicates construct validity. The fact that the OSATS scoring improves after practice and coaching in the 2011 experiment and both

trials in 2013, suggests that the simulator has content validity. The validation of the OSATS scoring metric against currently implemented simulator that are used for certification in general surgery means that that metric has criterion validity. It improves residents' fracture reduction skills (the skills may be transferable to other surgical procedures), the cost of training has decreased compared to cadaveric (a single cadaveric ankle would cost more than two Sawbones ankles) or operating room (OR) training (during which a mistake could cost thousands of dollars), objective metrics have been shown to correlate with resident improvement (OSATS scores), and because the residents no longer have their first fracture reduction attempt or learn fluoroscopy techniques in the OR, patient safety has been improved.

The goal of developing a simulator that improves residents' surgical performance has been accomplished as shown with the improvement with OSATS scoring in 2011 and both 2013 groups, the fluoroscopic improvement since in the Iowa trials, and the motion improvement seen in the Minnesota residents. The OSATS, articular deviation and contact stress exposure, and fluoroscopic usage metrics have been implemented and shown to demonstrate improvement, fulfilling the goal of developing new performance metrics. The residents' improvement in the metrics leads to increased patient safety. Since this is all accomplished, for a lower cost than a mistake in the OR or the cost of cadaveric limbs, the goal of creating a more cost effective training is accomplished.

Although the simulator has met the desired goals, it is a complex approach to surgical skills acquisition. The complexity has made it difficult to prove that skills are learned and improved. Currently, there are ongoing efforts to deconstruct the articular fracture reduction simulator into the fundamental skills that are needed to successfully perform the procedure. Once the subtasks have been identified, trainers for the individual skills can be created and validated. After trainers for the specific tasks, the fracture reduction simulator could be used to stitch the tasks together and teach the residents how to complete the procedure, rather than trying to do all of that and teach the basic skills as

well. This would allow for a better education, and should lower the need to perform the simulator several times, decrease cost and radiation exposure for resident education.

REFERENCES

1. Brown, T.D., et al., *Posttraumatic osteoarthritis: a first estimate of incidence, prevalence, and burden of disease.* Journal of orthopaedic trauma, 2006. **20**(10): p. 739-44.

2. Anderson, D.D., et al., *Is elevated contact stress predictive of post-traumatic osteoarthritis for imprecisely reduced tibial plafond fractures?* Journal of orthopaedic research : official publication of the Orthopaedic Research Society, 2011. **29**(1): p. 33-9.

3. Heim, U., *The Pilon Tibial Fracture.* Berlin:Springer-Verlag, 1995.
4. Ruedi, T.P. and M. Allgower, *The operative treatment of intra-articular fractures of the lower end of the tibia.* Clinical orthopaedics and related research, 1979(138): p. 105-10.

5. Marsh, J.L., et al., *Articular fractures: does an anatomic reduction really change the result?* The Journal of bone and joint surgery. American volume, 2002. **84-A**(7): p. 1259-71.

6. Watson, J.T., et al., *Pilon fractures. Treatment protocol based on severity of soft tissue injury.* Clinical orthopaedics and related research, 2000(375): p. 78-90.
7. Dirschl, D.R., et al., *Articular fractures.* The Journal of the American Academy of Orthopaedic Surgeons, 2004. **12**(6): p. 416-23.

8. Matta, J.M., *Fractures of the acetabulum: accuracy of reduction and clinical results in patients managed operatively within three weeks after the injury.* The Journal of bone and joint surgery. American volume, 1996. **78**(11): p. 1632-45.
9. Brown, T.D., et al., *Contact stress aberrations following imprecise reduction of simple tibial plateau fractures.* Journal of orthopaedic research : official publication of the Orthopaedic Research Society, 1988. **6**(6): p. 851-62.

10. Li, W., et al., *Patient-specific finite element analysis of chronic contact stress exposure after intraarticular fracture of the tibial plafond.* Journal of orthopaedic research : official publication of the Orthopaedic Research Society, 2008. **26**(8): p. 1039-45.

11. ACGME, *ACGME Progrma Requirements for Graduate Medical Education in Orthopaedic Surgery*, 2012.

12. Surgery, A.o.R.C.i.O. *PGY1 Year Changes - A Message from Dr. Marsh of the RRC*. 2013 2013 [cited 2013 June 3]; Available from: http://www.arcosonline.org/index.php?option=com_content&view=article&id=139:pgy1-year-changes-a-message-from-dr-marsh-of-the-rrc&catid=1:latest-news&Itemid=93.

13. Rentschler, E.H., *Porphyria and its simulation of surgical conditions.* The American surgeon, 1952. **18**(8): p. 758-69.

14. Zilm, F. and R.B. Hollis, *An application of simulation modeling to surgical intensive care bed need analysis in a university hospital.* Hospital & health

services administration, 1983. **28**(5): p. 82-101.

15. Shalimov, A.A., et al., *[Pathogenesis and major principles of simulation and surgical treatment of pancreatitis. I. Simulation of acute pancreatitis]*. Klinicheskaia khirurgiia, 1983(11): p. 30-4.

16. Capperauld, I. and J. Hargraves, *Surgical simulation for general practitioners.* Annals of the Royal College of Surgeons of England, 1991. **73**(5): p. 273-5.

17. Munro, A., et al., *Skin simulation for minor surgical procedures.* Journal of the Royal College of Surgeons of Edinburgh, 1994. **39**(3): p. 174-6.

18. Pier, A. and F. Gotz, *[Simulation trainer for endoscopic surgical techniques]*. Der Chirurg; Zeitschrift fur alle Gebiete der operativen Medizen, 1992. **63**(4): p. 387-92.

19. Jackson, A.P. and B.G. Tidmarsh, *Simulation models for teaching endodontic surgical procedures.* International endodontic journal, 1993. **26**(3): p. 198-200.

20. Noar, M.D., *The next generation of endoscopy simulation: minimally invasive surgical skills simulation.* Endoscopy, 1995. **27**(1): p. 81-5.

21. Edmond, C.V., Jr., G.J. Wiet, and B. Bolger, *Virtual environments. Surgical simulation in otolaryngology.* Otolaryngologic clinics of North America, 1998. **31**(2): p. 369-81.

22. Hill, J.W., et al., *Telepresence interface with applications to microsurgery and surgical simulation.* Studies in health technology and informatics, 1998. **50**: p. 96-102.

23. Zhang, X., X. Wang, and Z. Zhang, *[Accuracy of computerized aid diagnosis, surgical simulation and facial appearance prediction in orthognathic surgery]*. Zhonghua kou qiang yi xue za zhi = Zhonghua kouqiang yixue zazhi = Chinese journal of stomatology, 1998. **33**(1): p. 6-9.

24. Gorman, P.J., A.H. Meier, and T.M. Krummel, *Simulation and virtual reality in surgical education: real or unreal?* Archives of surgery, 1999. **134**(11): p. 1203-8.

25. Haluck, R.S. and T.M. Krummel, *Simulation and virtual reality for surgical education.* Surgical technology international, 1999. **8**: p. 59-63.

26. Kneebone, R.L., *Twelve tips on teaching basic surgical skills using simulation and multimedia.* Medical teacher, 1999. **21**(6): p. 571-5.

27. Gallagher, A.G. and C.U. Cates, *Approval of virtual reality training for carotid stenting: what this means for procedural-based medicine.* JAMA : the journal of the American Medical Association, 2004. **292**(24): p. 3024-6.

28. Ziegler, R., et al., *Virtual reality arthroscopy training simulator.* Computers in biology and medicine, 1995. **25**(2): p. 193-203.

29. Poss, R., et al., *Development of a virtual reality arthroscopic knee simulator.* The Journal of bone and joint surgery. American volume, 2000. **82-A**(10): p. 1495-9.

30. Fujioka, M., et al., *Computer-aided interactive surgical simulation for craniofacial anomalies based on 3-D surface reconstruction CT images.*

Radiation medicine, 1988. **6**(5): p. 204-12.

31. Lo, L.J., et al., *Craniofacial computer-assisted surgical planning and simulation.* Clinics in plastic surgery, 1994. **21**(4): p. 501-16.

32. Abe, M., et al., *Model-based surgical planning and simulation of cranial base surgery.* Neurologia medico-chirurgica, 1998. **38**(11): p. 746-50; discussion 750-1.

33. Ozawa, Y., et al., *Surgical simulation of Class III edentulous patient using a 3D craniofacial model: report of a case.* The Bulletin of Tokyo Dental College, 2000. **41**(2): p. 73-7.

34. Nakajima, H., et al., *Craniofacial surgical simulation system in the 3 dimensional CT SurgiPlan system.* The Keio journal of medicine, 2001. **50 Suppl 2**: p. 95-102.

35. Kurihara, T., *The fourth dimension in simulation surgery for craniofacial surgical procedures.* The Keio journal of medicine, 2001. **50 Suppl 2**: p. 155-65.

36. Lin, C.C., et al., *Craniofacial surgical simulation: application of three-dimensional medical imaging and rapid prototyping models.* Chang Gung medical journal, 2001. **24**(4): p. 229-38.

37. Teschner, M., S. Girod, and B. Girod, *3-D simulation of craniofacial surgical procedures.* Studies in health technology and informatics, 2001. **81**: p. 502-8.

38. Morris, D., et al., *An interactive simulation environment for craniofacial surgical procedures.* Studies in health technology and informatics, 2005. **111**: p. 334-41.

39. Xia, J.J., et al., *Cost-effectiveness analysis for computer-aided surgical simulation in complex cranio-maxillofacial surgery.* Journal of oral and maxillofacial surgery : official journal of the American Association of Oral and Maxillofacial Surgeons, 2006. **64**(12): p. 1780-4.

40. Fried, M.P., et al., *Identifying and reducing errors with surgical simulation.* Quality & safety in health care, 2004. **13 Suppl 1**: p. i19-26.

41. Fried, G.M., et al., *Proving the value of simulation in laparoscopic surgery.* Annals of surgery, 2004. **240**(3): p. 518-25; discussion 525-8.

42. Bann, S., et al., *The reliability of multiple objective measures of surgery and the role of human performance.* American journal of surgery, 2005. **189**(6): p. 747-52.

43. Anastakis, D.J., et al., *Evaluating the effectiveness of a 2-year curriculum in a surgical skills center.* American journal of surgery, 2003. **185**(4): p. 378-85.

44. MacRae, H., et al., *A comprehensive examination for senior surgical residents.* American journal of surgery, 2000. **179**(3): p. 190-3.

45. Flexner, A., *Medical education in the United States and Canada. From the Carnegie Foundation for the Advancement of Teaching, Bulletin Number Four, 1910.* Bulletin of the World Health Organization, 2002. **80**(7): p. 594-602.

46. Michelson, J.D., *Simulation in orthopaedic education: an overview of theory and practice.* The Journal of bone and joint surgery. American volume, 2006. **88**(6): p. 1405-11.

47. Examiners, N.B.o.M. *USMLE Bulletin*. 2006 [cited 2005 Dec 21]; Available from: www.usmle.org/bulletin/2006/TOC.htm.

48. Yehyawi TM, T.T., Ohrt GT, Marsh JL, Karam MD, Brown TD, Anderson DD, *A Simulation Trainer For Complex Articular Fracture Surgery.* J Bone Joint Surg AM, 2012. **In Press**.

49. Pedowitz, R.A., J. Esch, and S. Snyder, *Evaluation of a virtual reality simulator for arthroscopy skills development.* Arthroscopy : the journal of arthroscopic & related surgery : official publication of the Arthroscopy Association of North America and the International Arthroscopy Association, 2002. **18**(6): p. E29.

50. Strom, P., et al., *Validation and learning in the Procedicus KSA virtual reality surgical simulator.* Surgical endoscopy, 2003. **17**(2): p. 227-31.

51. Strom, P., et al., *Training in tasks with different visual-spatial components does not improve virtual arthroscopy performance.* Surgical endoscopy, 2004. **18**(1): p. 115-20.

52. Hayter, M.A., et al., *Validation of the Imperial College Surgical Assessment Device (ICSAD) for labour epidural placement.* Canadian journal of anaesthesia = Journal canadien d'anesthesie, 2009. **56**(6): p. 419-26.

53. Datta, V., et al., *The relationship between motion analysis and surgical technical assessments.* American journal of surgery, 2002. **184**(1): p. 70-3.

54. Sansregret, A., et al., *Choosing the right physical laparoscopic simulator? Comparison of LTS2000-ISM60 with MISTELS: validation, correlation, and user satisfaction.* American journal of surgery, 2009. **197**(2): p. 258-65.

55. Vassiliou, M.C., et al., *The MISTELS program to measure technical skill in laparoscopic surgery : evidence for reliability.* Surgical endoscopy, 2006. **20**(5): p. 744-7.

56. Dauster, B., et al., *Validity of the MISTELS simulator for laparoscopy training in urology.* Journal of endourology / Endourological Society, 2005. **19**(5): p. 541-5.

57. Fraser, S.A., et al., *Evaluating laparoscopic skills: setting the pass/fail score for the MISTELS system.* Surgical endoscopy, 2003. **17**(6): p. 964-7.

58. Moorthy, K., et al., *Bimodal assessment of laparoscopic suturing skills: construct and concurrent validity.* Surgical endoscopy, 2004. **18**(11): p. 1608-12.

59. Moorthy, K., et al., *An innovative method for the assessment of skills in lower gastrointestinal endoscopy.* Surgical endoscopy, 2004. **18**(11): p. 1613-9.

60. Seymour, N.E., et al., *Virtual reality training improves operating room performance: results of a randomized, double-blinded study.* Annals of surgery, 2002. **236**(4): p. 458-63; discussion 463-4.

61. Datta, V., et al., *The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model.* Journal of

the American College of Surgeons, 2001. **193**(5): p. 479-85.

62. Datta, V., et al., *Relationship between skill and outcome in the laboratory-based model.* Surgery, 2002. **131**(3): p. 318-23.

63. Mackay, S., et al., *Practice distribution in procedural skills training: a randomized controlled trial.* Surgical endoscopy, 2002. **16**(6): p. 957-61.

64. Mackay, S., et al., *Electromagnetic motion analysis in the assessment of surgical skill: Relationship between time and movement.* ANZ journal of surgery, 2002. **72**(9): p. 632-4.

65. Martin, J.A., et al., *Objective structured assessment of technical skill (OSATS) for surgical residents.* The British journal of surgery, 1997. **84**(2): p. 273-8.

66. Reznick, R., et al., *Testing technical skill via an innovative "bench station" examination.* American journal of surgery, 1997. **173**(3): p. 226-30.

67. Reznick, R.K. and H. MacRae, *Teaching surgical skills--changes in the wind.* The New England journal of medicine, 2006. **355**(25): p. 2664-9.

68. Szalay, D., et al., *Using operative outcome to assess technical skill.* American journal of surgery, 2000. **180**(3): p. 234-7.

69. Reznick, R.K., et al., *Process-rating forms versus task-specific checklists in an OSCE for medical licensure. Medical Council of Canada.* Academic medicine : journal of the Association of American Medical Colleges, 1998. **73**(10 Suppl): p. S97-9.

70. Regehr, G., et al., *Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination.* Academic medicine : journal of the Association of American Medical Colleges, 1998. **73**(9): p. 993-7.

71. Harden, R.M. and F.A. Gleeson, *Assessment of clinical competence using an objective structured clinical examination (OSCE).* Medical education, 1979. **13**(1): p. 39-54.

72. Taffinder, N., et al., *Validation of virtual reality to teach and assess psychomotor skills in laparoscopic surgery: results from randomised controlled studies using the MIST VR laparoscopic simulator.* Studies in health technology and informatics, 1998. **50**: p. 124-30.

73. Darzi, A. and S. Mackay, *Assessment of surgical competence.* Quality in health care : QHC, 2001. **10 Suppl 2**: p. ii64-9.

74. Bouldin, J., *The body, animation and the real: Race, reality and the rotoscope in Betty Boop.* Proceedings of affective encounters: rethinking embodiment in feminist media studies, 2001: p. 48-54.

75. Menache, A., *Understanding motion capture for computer animation and video games*2000: Morgan Kaufmann Pub.

76. Moeslund, T.B. and E. Granum, *A Survey of Computer Vision-Based Human Motion Capture.* Computer Vision and Image Understanding, 2001. **81**(3): p. 231-

268.

77. Moeslund, T.B., A. Hilton, and V. Krüger, *A survey of advances in vision-based human motion capture and analysis.* Computer Vision and Image Understanding, 2006. **104**(2–3): p. 90-126.

78. Auer, T., S. Brantner, and A. Pinz, *The integration of optical and magnetic tracking for multi-user augmented reality*, in *Virtual Environments' 99*1999, Springer. p. 43-52.

79. Polhemus. http://www.polhemus.com/. 2013 6/1/2013 [cited 2013 June 10]; Available from: http://www.polhemus.com/.

80. Qualisys, A., *Qualisys track manager user manual.* Gothenburg: Qualisys AB, 2006.

81. Barnett, T.G., J.G. Logan, and J.M. Paterson, *Estimation of 3-Dimensional Joint Angles Using Qualisys Position Sensors.* Journal of Physiology-London, 1993. **467**: p. P7-P7.

82. Ren, L., R.K. Jones, and D. Howard, *Predictive modelling of human walking over a complete gait cycle.* Journal of biomechanics, 2007. **40**(7): p. 1567-1574.

83. Eriksson Crommert, M., M. Ekblom, and A. Thorstensson, *Motor control of the trunk during a modified clean and jerk lift.* Scandinavian Journal of Medicine & Science in Sports, 2013.

84. Ping, U.E. and S. Parasuraman, *Modeling of Wrist and Hand Motion while Performing Functional Task.* Advanced Materials Research, 2012. **433**: p. 2316-2320.

85. Liebermann, D.G., et al. *Arm path fragmentation and spatiotemporal features of hand reaching in healthy subjects and stroke patients*. in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*. 2010. IEEE.

86. Jiang, R. and Z. Wang. *Analysis on Smashing Motion in Badminton*. in *Proceedings of the International Conference on Information Engineering and Applications (IEA) 2012*. 2013. Springer.

87. Aziz, A., et al., *Development of gesture database for an adaptive gesture recognition system.* International Journal, 2012.

88. Brigstocke, G., et al., *In-vivo confirmation of the use of the dart thrower's motion during activities of daily living.* Journal of Hand Surgery (European Volume), 2012.

89. Polhemus, J.T., J.E. Morgan, and A. Mandell, *Condition sensor system and method*, 1978, Google Patents.

90. Mandell, A., J. Morgan, and J. Polhemus, *Rocking-motion sensor for the blind.* 1977.

91. Polhemus, J., J. Morgan, and A. Mandell, *A rocking motion sensor for the blind.* ISA transactions, 1976. **15**(2): p. 192.

92. Kuipers, J., *Tracking and determining orientation of object using coordinate transformation means, system and process*, 1976, Google Patents.

93. Kuipers, J., *Apparatus for generating a nutating electromagnetic field*, 1977, Google Patents.

94. Kern, A.M., *Large Population Evaluation of Contact Stress Exposure in Articular Joints for Prediction of Osteoarthritis Onset and Progression*, in *Biomedical Engineering*2011, University of Iowa: Iowa City. p. 110.

95. Thomas, T.P., et al., *Objective CT-based metrics of articular fracture severity to assess risk for posttraumatic osteoarthritis.* Journal of orthopaedic trauma, 2010. **24**(12): p. 764-9.

96. Buckwalter, J.A. and T.D. Brown, *Joint injury, repair, and remodeling: roles in post-traumatic osteoarthritis.* Clinical orthopaedics and related research, 2004(423): p. 7-16.

97. DeCoster, T.A., et al., *Rank order analysis of tibial plafond fractures: does injury or reduction predict outcome?* Foot & ankle international. / American Orthopaedic Foot and Ankle Society [and] Swiss Foot and Ankle Society, 1999. **20**(1): p. 44-9.

98. Hadley, N.A., T.D. Brown, and S.L. Weinstein, *The effects of contact pressure elevations and aseptic necrosis on the long-term outcome of congenital hip dislocation.* Journal of Orthopaedic Research, 1990. **8**(4): p. 504-513.

99. Maxian, T.A., T.D. Brown, and S.L. Weinstein, *Chronic stress tolerance levels for human articular cartilage: Two nonuniform contact models applied to long-term follow-up of CDH.* Journal of biomechanics, 1995. **28**(2): p. 159-166.

100. Brekelmans, W.A., H.W. Poort, and T.J. Slooff, *A new method to analyse the mechanical behaviour of skeletal parts.* Acta orthopaedica Scandinavica, 1972. **43**(5): p. 301-17.

101. Huiskes, R. and E.Y. Chao, *A survey of finite element analysis in orthopedic biomechanics: the first decade.* Journal of biomechanics, 1983. **16**(6): p. 385-409.

102. Kloosterman, G., *Contact methods in finite element solutions*, in *Reseach Scool of integrated Manufacturing*2002, University of Twente: Twente. p. 134.

103. Li, W., et al., *An evolutionary approach to elastic contact optimization of frame structures.* Finite Elements in Analysis and Design, 2003. **40**(1): p. 61-81.

104. Grosland, N.M., et al., *IA-FEMesh: an open-source, interactive, multiblock approach to anatomic finite element model development.* Computer methods and programs in biomedicine, 2009. **94**(1): p. 96-107.

105. Anderson, D.D., et al., *Intra-articular contact stress distributions at the ankle throughout stance phase-patient-specific finite element analysis as a metric of degeneration propensity.* Biomechanics and modeling in mechanobiology, 2006.

5(2-3): p. 82-9.

106. Fitzpatrick, C.K., M.A. Baldwin, and P.J. Rullkoetter, *Computationally efficient finite element evaluation of natural patellofemoral mechanics.* Journal of biomechanical engineering, 2010. **132**(12): p. 121013.

107. Anderson, D.D., et al., *Quantifying tibial plafond fracture severity: absorbed energy and fragment displacement agree with clinical rank ordering.* Journal of orthopaedic research : official publication of the Orthopaedic Research Society, 2008. **26**(8): p. 1046-52.

108. Anderson, D.D., et al., *Physical validation of a patient-specific contact finite element model of the ankle.* Journal of biomechanics, 2007. **40**(8): p. 1662-9.

109. Kellgren, J.H. and J.S. Lawrence, *Radiological assessment of osteo-arthrosis.* Annals of the rheumatic diseases, 1957. **16**(4): p. 494-502.

110. Li, G., M. Sakamoto, and E.Y. Chao, *A comparison of different methods in predicting static pressure distribution in articulating joints.* Journal of biomechanics, 1997. **30**(6): p. 635-8.

111. Chao, E.Y., et al., *Discrete element analysis in musculoskeletal biomechanics.* Molecular & cellular biomechanics : MCB, 2010. **7**(3): p. 175-92.

112. Elias, J.J., et al., *Evaluation of a computational model used to predict the patellofemoral contact pressure distribution.* Journal of biomechanics, 2004. **37**(3): p. 295-302.

113. Blankevoort, L., et al., *Articular contact in a three-dimensional model of the knee.* Journal of biomechanics, 1991. **24**(11): p. 1019-31.

114. Bei, Y. and B.J. Fregly, *Multibody dynamic simulation of knee contact mechanics.* Medical engineering & physics, 2004. **26**(9): p. 777-89.

115. Schuind, F.A., et al., *A biomechanical study of the ulnar nerve at the elbow.* Journal of hand surgery, 1995. **20**(5): p. 623-7.

116. An, K.N., et al., *Pressure distribution on articular surfaces: application to joint stability evaluation.* Journal of biomechanics, 1990. **23**(10): p. 1013-20.

117. Anderson, D.D., et al., *Implementation of discrete element analysis for subject-specific, population-wide investigations of habitual contact stress exposure.* Journal of applied biomechanics, 2010. **26**(2): p. 215-23.

118. Fregly, B.J., et al., *Computational wear prediction of a total knee replacement from in vivo kinematics.* Journal of biomechanics, 2005. **38**(2): p. 305-14.

119. Schuind, F., et al., *Force and pressure transmission through the normal wrist. A theoretical two-dimensional study in the posteroanterior plane.* Journal of biomechanics, 1995. **28**(5): p. 587-601.

120. Halloran, J.P., et al., *Comparison of deformable and elastic foundation finite element simulations for predicting knee replacement mechanics.* Journal of biomechanical engineering, 2005. **127**(5): p. 813-8.

121. Caruntu, D.I. and M.S. Hefzy, *3-D anatomically based dynamic modeling of the human knee to include tibio-femoral and patello-femoral joints.* Journal of biomechanical engineering, 2004. **126**(1): p. 44-53.

122. Van Heest, A., et al., *Assessment of technical skills of orthopaedic surgery residents performing open carpal tunnel release surgery.* The Journal of bone and joint surgery. American volume, 2009. **91**(12): p. 2811-7.

123. Insel, A., et al., *The development of an objective model to assess arthroscopic performance.* The Journal of bone and joint surgery. American volume, 2009. **91**(9): p. 2287-95.

124. Thomas, T.P., *Virtual Pre-Operative Reconstruction Planning For Comminuate Articular Fractures*, in *Biomedical Engineering*2010, University of Iowa: Iowa City. p. 130.

125. Beardsley, C.L., et al., *High density polyetherurethane foam as a fragmentation and radiographic surrogate for cortical bone.* The Iowa orthopaedic journal, 2000. **20**: p. 24-30.

126. Frank MC, H.C., Anderson DD, McKinley TO, Brown TD. *Rapid manufacturing in biomedical materials: Using subtractive rapid prototyping for bone replacement*. in *19th Annual International Solid Freeform Fabrication Symposium*. 2008. Austin, Texas.

127. Frank, M.C., R.A. Wysk, and S.B. Joshi, *Rapid planning for CNC milling - A new approach for rapid prototyping.* Journal of Manufacturing Systems, 2004. **23**(3): p. 242-255.

128. Faulkner, H., et al., *Validation of an objective structured assessment of technical skill for surgical residents.* Academic medicine : journal of the Association of American Medical Colleges, 1996. **71**(12): p. 1363-5.

129. Karam, M.D., et al., *Application of surgical skill simulation training and assessment in orthopaedic trauma.* The Iowa orthopaedic journal, 2012. **32**: p. 76-82.

130. Marsh JL, K.M., Yehywai TM, Westerlind BO, Kho J, Anderson DD. *A Surgical Skills Training Curriculum for PGY-1 Orthopaedic Residents*. in *American Academy of Orthopaedic Surgeons*. 2012.

131. El-Khoury, G.Y., et al., *Cartilage thickness in cadaveric ankles: measurement with double-contrast multi-detector row CT arthrography versus MR imaging.* Radiology, 2004. **233**(3): p. 768-73.

132. Stauffer, R.N., E.Y. Chao, and R.C. Brewster, *Force and motion analysis of the normal, diseased, and prosthetic ankle joint.* Clinical orthopaedics and related research, 1977(127): p. 189-96.

133. Marsh JL, Y.T., Ohrt GT, Thomas TP, Anderson DD,!Brown TD, Karam MD, *In Articular Fracture Reduction Training for Orthopaedic Residents: Surgical Simulation and Performance Assessment*, in *AOA Council of Residency Directors Conference*2011: Boston, MA.

134. Karam, M.D., et al., *Current and future use of surgical skills training laboratories in orthopaedic resident education: a national survey.* The Journal of bone and joint surgery. American volume, 2013. **95**(1): p. e4.