Proficiency-Based Progression Training: Quality Assured Preparation for the Practice of Surgery

Richard L. Angelo, M.D.

PhD Thesis

National University of Ireland, Cork

Department: Surgery

June, 2017

Department Head: Professor Henry Paul Redmond

Supervisors:
Anthony G. Gallagher, PhD, DSc
Eugene Dempsey, M.D.

Table of Contents

Thesis Abstract

Purpose: To investigate the effectiveness of proficiency-based progression (PBP) training coupled with simulation for the acquisition of surgical skills and to define the essentials for the development of a comprehensive, validated, PBP training curriculum.

Methods: The nature and rationale for the paradigm shift in surgical skills training from the apprenticeship model to one of proficiency-based progression training is reviewed along with the intent to move from process-based to outcomes-based medical education and training. A review of the diagnostic assessment of shoulder instability and the evolution of shoulder arthroscopy in the management of unidirectional anterior glenohumeral instability is presented along with effective, current techniques essential to obtaining a successful repair. The proficiency-based progression curriculum design was initiated with a task deconstruction of an arthroscopic Bankart repair (ABR) in which step, error, and sentinel (more serious) error metrics along with phases of the procedure were defined. A modified Delphi panel of senior experienced shoulder surgeon faculty members (N =27) was convened to determine face and content validity of the metrics as an accurate and reliable evaluation of operative performance for an ABR. To determine the construct validity of the ABR metrics coupled with a medium fidelity shoulder model simulator, the performance of novice (N = 7) and experienced (N = 12) surgeons was compared using full-length videos of the subjects performing a 3 anchor ABR on a shoulder simulator. Trained reviewers scored the videos in blinded fashion. To determine construct validity of the ABR metrics coupled with a cadaver shoulder, the performance of novice (N = 12) and experienced (N = 10) surgeons was again compared using full-length ABR videos scored in blinded fashion. An investigation to objectively evaluate knot-tying performance was conducted using the 'Fundamentals of Arthroscopic Surgery Training' (FAST) workstation and knot tester. Knot loop

constructs were stressed to 15# for 15 seconds. The performance of faculty (N = 20), experienced (N = 30), and novice (N = 44) surgeons were compared. The novice surgeons were randomized into 3 groups to compared the effectiveness of a various curricula for knot tying: A) standard training representing the apprenticeship model (N = 14), B) FAST workstation enhanced (N = 14), and C) PBP curriculum employing the knot tester (N = 16). Finally, a randomized, controlled, blinded trial was conducted comparing the performance of 4th and 5th year orthopedic residents exposed to one of 3 different training curriculums for an ABR on a shoulder cadaver: 1) a traditional AANA residency training program representing the apprenticeship model (N = 14), 2) a simulator enhanced curriculum (N = 14), and 3) a proficiency-based progression curriculum coupled with a model simulator (N = 16). In the latter, the instructors 'taught to the metrics' and provided proximate feedback enabling the trainee to engage in deliberate practice.

Results: Face and content validity were confirmed and consensus achieved for the ABR metrics through the modified Delphi panel deliberations. Construct validity of the metrics coupled with the model simulator was verified as a training tool. The experienced group made 63% fewer errors, committed 79% fewer sentinel errors, and performed the procedure in 42% less time than the novice group (all significant differences). A proficiency benchmark for the shoulder model simulator was specified as completing a 3 anchor arthroscopic Bankart repair with no more than 4 errors in total, and no more than 1 sentinel error. Construct validity of the metrics coupled with a cadaver shoulder was verified as an accurate assessment tool. Novice surgeons made 54% more errors, showed significantly more performance variability (SD, 3.5 v 1.6), and took significantly longer to perform the procedure (45.5 minutes v 25.9 minutes). A proficiency benchmark for the metrics coupled with a cadaver shoulder consisted of completing a 3 anchor arthroscopic Bankart repair with no more than 3 errors in total, and no more than 1 sentinel error. The FAST workstation and tester proved to be accurate and reliable. In the faculty group, 24% of knots "failed" under load and performance was inconsistent. In the experienced group, 22% of knots failed and for the novice group, 26% of knots failed. The novice subgroup of PBP trained residents demonstrated an 11% knot failure rate (half the faculty rate). The randomized trial comparing three different training curricula for performance of an ABR on a cadaver demonstrated unambiguous superiority of the PBP protocol coupled with a medium fidelity simulator. The PBP-trained group (Group C) made 56% fewer objectively assessed errors than the traditionally trained group (Group A) and 41% fewer than Group B (both comparisons were statistically significant). The proficiency benchmark was achieved on the final repair by 75% of the Group C^{PBP} residents (who met all of the intermediate proficiency benchmarks) and 68.7% of the participants in the entire C Group, compared with 36.7% in Group B and 28.6% in Group A. When compared with Group A, Group B participants were 1.4 times, Group C participants were 5.5 times, and Group C^{PBP} participants 7.5 times as likely to achieve the final proficiency benchmark.

Conclusions: Task deconstruction of an arthroscopic Bankart procedure facilitated the creation of a validated proficiency based progression training program that was metric based. Construct validity was demonstrated for the metrics with the model simulator as a training tool and the cadaver shoulder as a performance assessment tool. In a randomized trial, a PBP curriculum coupled with simulation training was dramatically more effective in training the skills necessary to reach the proficiency benchmark for an ABR than both simulation-based training and current AANA training methods. Initial meetings for the metric developers should be conducted in person, but cost-effective internet-based methods for subsequent communication substantially reduces cost. The index procedure selected for task deconstruction and metric development should be uncomplicated and employ commonly used techniques. Performance metrics must be unambiguously defined and able to be reliably scored. Error metrics are the most valuable in discriminating between levels of operative performance. Thorough orientation and training for video reviewers is essential to ensure acceptable inter-rater reliability among scoring pairs. The establishment of a fair, clear, and objective proficiency benchmark serves as a reference standard and provides an intermediate assessment for the trainee to specify deficiencies requiring correction. Simulators are most useful when they serve as a vehicle to deliver a strong, metric based curriculum, which must be developed prior to the selection of specific simulations. The fidelity of a particular simulator should be matched to the specific skill or task to be trained. Task deconstruction along with metric development and construct validation are time-consuming endeavors that will involve substantial cost, but offer the potential for superior, objective based surgical skills training.

1. Introduction

1.1 Paradigm Shift

The most important mission of the Arthroscopy Association of North America (AANA) is the education of its membership. In July of 2009 the President's Council of AANA met, in part to discuss potential methods and strategies to upgrade the society's educational programming. In the past, this had been accomplished primarily using a "bottom up" approach, i.e. "what are our most effective and highest rated programs", and "how can we then improve upon our current offerings". Tasked by the President's Council, a group of thought leaders from AANA elected to employ a distinctly unique methodology using a different approach or "top down" strategy. They proposed to go out into the world and examine other businesses, industries, and professions that trained highly skilled individuals, and examine the 'best practice' strategies that those other entities employ. In May of 2010, a task force was appointed with First Vice President Richard Angelo, M.D. as its Chair. The mandate for that task force was to "sail around world", across varied disciplines and professions seeking answers to the question, "what are the most effective methods being used to educate and train individuals to work in highly-skilled technical professions." This effort to "sail the world" in search of educational pearls became known as the AANA Magellan Project (although Magellan didn't complete the journey, his expedition was credited with being the first to circumnavigate the globe, or 'sail around the world'). As promising educational strategies were discovered, the intent was for AANA to adapt and apply those methods to training surgeons in the principles and best practice of arthroscopic surgery.

The Magellan Project included six subcommittees: Didactic, Surgical Skills, Electronic Media, Simulation, Outcomes / Metrics, and Health Policy / Advocacy. Within the focus of each of the subcommittees, the members conducted extensive research into potential concepts and ideas that AANA might employ to enhance surgeon education. For example, the Outcomes / Metrics subcommittee was provided with several contacts at the National Aeronautics and Space Administration (NASA). The committee posed the questions to those responsible for astronaut education, "With the significant costs related to preparation and evaluation, how do you determine which astronauts to train to perform highly skilled maneuvers such as docking the lunar landing module", and "How do you assess whether the necessary skills were mastered or not?" An exhaustive document was returned entitled, "Development and Implementation of an Extravehicular Activity Skills Program for Astronauts".[1] This document was used as a template to develop an Arthroscopic Bankart Skills Assessment Tool. A Bankart repair addresses the most common pathology encountered in unidirectional anterior shoulder instability – capsulolabral tearing and detachment from the anterior and inferior glenoid rim. The pilot program sought to evaluate the learner's skill development using an 'Alex Shoulder Model Professor' (Sawbones, Inc., Vashon Island, Wash.) as a medium fidelity dry model "simulator".

From a different Magellan subcommittee, Surgical Skills, the question was posed, "Is there a better way to train surgical skills than our current methods?" AANA has conducted over 300 arthroscopic lab skills courses using both models and cadaver tissue, but has been unable to make any reliable determination as to the effectiveness of those programs with respect to the registrants completing the courses actually acquiring improved arthroscopic skills. The curriculum has often varied from course to course with the content of handouts and outlines largely dependent on which Master and Associate Master Faculty were teaching the course. Despite listening to lectures, viewing videos demonstrating the various procedures to be learned, and practicing endoscopic knot tying, registrants were often unprepared to work on cadaver shoulders. For those trainees who dismissed faculty guidance, portals were often improperly placed, which limited their utility and rendered suboptimal views of the joint being studied. Inefficiencies due to lack of familiarity with the sequence of steps for preparation, insufficient understanding of the techniques necessary to execute the procedure, and the inability to properly handle various arthroscopic instruments often resulted in marginal progress and significant fluid extravasation into the cadaveric soft tissues. With distorted anatomy, the practice of key procedural steps proved to be difficult if not impossible. Rather than gaining confidence in their improving abilities, some registrants were left discouraged and less confident. At the completion of many courses, substantial variance existed for the attendees with respect to the value of the instruction and the profitability realized in improving their surgical skill. No objective assessment of the registrant's skill was made at the completion of the various programs. Concern existed that the AANA lab courses were simply providing an educational 'experience' that failed to lead to substantive improvements in surgical skill for many of the registrants.

Predominantly through "sailing" the Internet, the Surgical Skills subcommittee became aware of 'proficiency based progression' (PBP) training for surgical skills as an alternative to the 'apprenticeship' model. The PBP training protocol dictates that the trainee master and be able to demonstrate increasingly more complex skillsets before being able to progress in training. The principles and validation of the PBP concepts for surgery have been evolving over the past 25 years, predominantly in the laparoscopic and general surgery realms.

At approximately the same time as the subcommittee Magellan voyages were taking place, I was serving on an advisory board for the first "World Congress on Surgical Skills Training" held in Goteborg, Sweden. We submitted an abstract for the meeting, detailing the Magellan Project efforts. Dr. Anthony Gallagher (whom I did not know at the time) was also serving on the advisory board and read the abstract of the Magellan Project. He emailed 3 related articles that "I might find of interest". Unbeknownst to me, he is likely the world's authority on PBP training for procedural skills. After several months of communication, I indicated to him that AANA was interested in studying PBP training to determine if it was a methodology that should be considered by AANA in optimizing their educational programs. We believed that prior to proposing to the BOD that a substantially different curriculum for surgical skills education be adopted, evidence of its superiority needed to be proven and presented. Dr. Gallagher accepted our invitation to serve as a consultant. We embarked on an investigation to study the merits of the 'paradigm shift' from the apprenticeship model to proficiency based progression training in order to determine the latter's effectiveness. Thus, the effort became known as the AANA 'Copernicus Initiative' (Nicolai Copernicus is credited with being the major influence in the paradigm shift from the earth to the sun being the center of the universe). Dr. Gallagher guided the primary investigators, Richard Ryu, M.D., Robert Pedowitz, M.D., PhD, and myself, along the path of study design and implementation.

1.2 Surgical Skills Training

"See one, do one, teach one" is reflective of the conventional or apprenticeship training model used to prepare physicians for surgical practice.[2] In the past, the sheer volume of exposure to surgical procedures and progressive involvement in patient cases during training led to reasonable preparation for most resident surgeons. Current safety concerns related to trainees operating on patients [3-6], the costs associated with prolonged operative times in training facilities[7], and training inefficiencies[8] have created pressures on training programs and made sufficient resident preparation for surgical practice less certain today. Two primary issues have contributed to this dilemma. The introduction of significantly more technically demanding minimally-invasive surgical procedures (MIS) require a new and more sophisticated skill set, and the reduction in trainee hours and opportunities to develop essential surgical skills have led to potential deficiencies in trainee preparation.[9-12] With the advent of endoscopic techniques, the challenge of preparing trainees has escalated significantly. The demand for a thorough knowledge of gross anatomy has always existed, but an appreciation for microanatomy has become essential. Critical to the practice of any surgical discipline is a clear and accurate understanding of anatomical relationships, which permits the safe dissection and manipulation of tissues. The limited field of view (FOV) afforded by an endoscope, however, requires that those anatomical relationships be appreciated at greater magnification for the surgeon to maintain orientation. In addition, important reference landmarks, previously relied on during open surgery, are often outside the endocopist's FOV and therefore, less available as key reference structures.

1.3 Learning Curve

The execution of any highly technical skill requires both a conceptual understanding of the objective and how to accomplish it, as well as the physical ability to perform the necessary techniques well on a consistent basis. The more demanding the skills required, the greater the need for practice, rehearsal, and repeated training in an effort to obtain and maintain specific skills. A learning curve exists for all endeavors requiring the execution of a technical skill, endoscopic surgery likely even more so.

Progress along the learning curve is dependent on many factors including the difficulty of the skill or technique being acquired, the training tools and educational methods employed, and the innate visuospatial, perceptual, and psychomotor talents possessed by the trainee. The more challenging the skill, the higher the learning curve and more difficult it is to achieve proficiency. Training exercises and operative experience are the primary means of moving up the learning curve. The more accurate the training tools are at emulating a specific task or simulating real context, the more effective they will be at assisting the student in acquiring the necessary skills. Some trainees will progress up the learning curve more efficiently with a higher slope (Figure 1, trainee 'A'), while others will take longer and have a lower slope to their learning curve (trainee 'B'). The slower the acquisition of the component abilities, the longer the learning curve extends over time. Simple practice and repetition, however, will not ensure success if the training exercises fail to provide an accurate synthetic experience. Repetitive practice alone may result in the student moving more horizontally along, rather than vertically up the curve. As a result, the slope of progress decreases and the curve elongates (trainee 'C') with the trainee ultimately failing to achieve proficiency. Further, skills acquired have a tendency to degrade over time if not sufficiently practiced and utilized.[13-15] A student may periodically make good progress, but experience degradation of skill over time, which can also prevent the attainment of proficiency (trainee 'D').

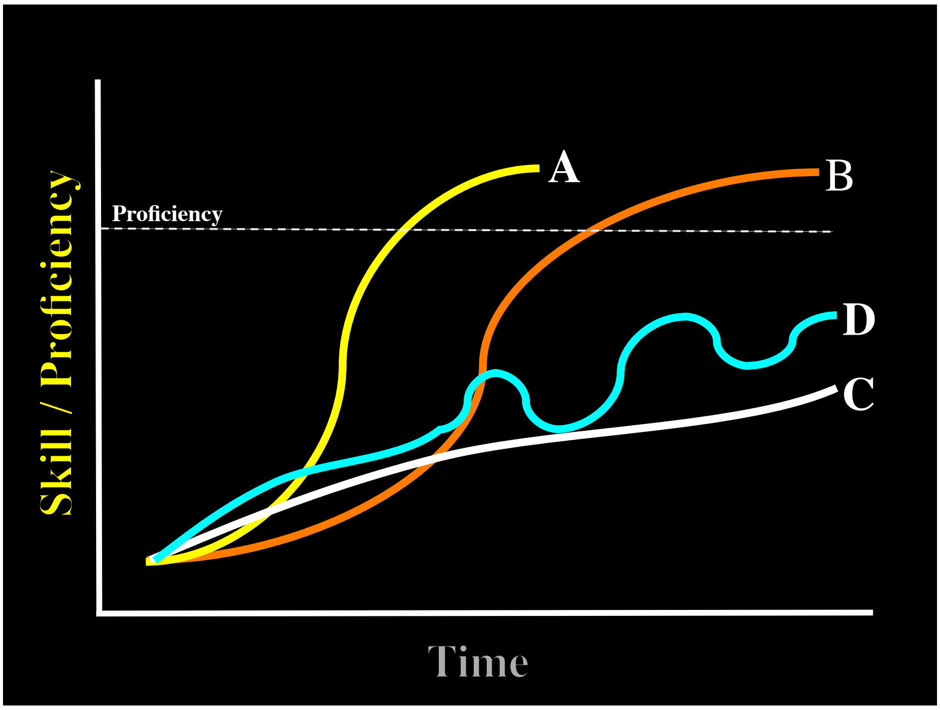


Figure 1: A = efficient progress up the learning curve; B = slower progress but trainee eventually attains proficiency; C = progress too

slow to attain proficiency; D = progress in spurts, but skill degradation prevents attainment of proficiency.

Morbidity (suboptimal outcomes, complications and the associated costs to manage them) is inversely related to the acquisition of skill and progress up the learning curve. While there may be a financial cost associated with various aspects of surgical training, morbidity only begins when the surgeon commences operating on patients. The morbidity curve M is inversely related to the learning curve (L) (Figure 2). The area (MA) under the morbidity curve (M) increases significantly as training becomes more inefficient and a longer period of time is required to attain proficiency.

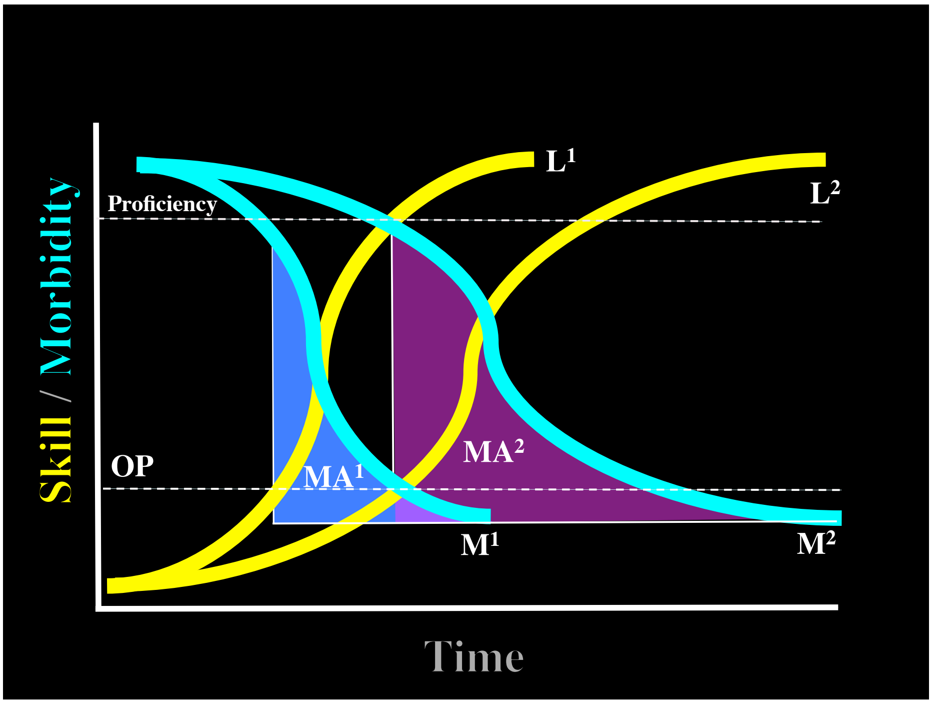


Figure 2: L^1 = learning curve for trainee 1; M^1 = morbidity curve for trainee 1 inversely related to the corresponding learning curve L^1; L^2 = learning curve for trainee 2; M^2 = morbidity curve corresponding to learning curve L^2; MA^1 = morbidity area for trainee 1 representing complications, suboptimal outcomes, and associated costs; MA^2 =

morbidity area for trainee 2 (a much greater morbidity area for trainee 2 compared to trainee 1 due to elongated learning curve).

Typically, early progress up the learning curve for surgical skills begins in the laboratory setting. As long as the training occurs ex vivo, there would be an associated financial cost, but no patient morbidity. Historically, however, much of the new surgeon's experience and a significant portion of the progress up the learning curve toward proficiency have taken place after entering practice and operating on live patients. Despite completing their surgical training, residents may have been relatively low on the learning curve for specific techniques and procedures as they initiated their clinical practice. Refinements in their surgical skills and the knowledge of how to avoid complications gradually lead to improved patient outcomes. The earlier the point on the learning curve that the surgeon begins to operate on patients, the larger the area beneath the morbidity curve (Figure 3 – MA^1). The same pattern often exists for experienced surgeons studying to acquire a new skill or technique. Figure 3 depicts 2 trainees with simular learning curves, one who begins to operate on patients while relatively low on the learning curve (OP^1), and a second trainee who acquired a large component of their surgical skill in the laboratory with repetative practice and rehearsal (OP^2). The trainee who begins to operate on patients from a position significantly higher on the learning curve has a much smaller area beneath the morbidity curve (MA^2 vs MA^1).

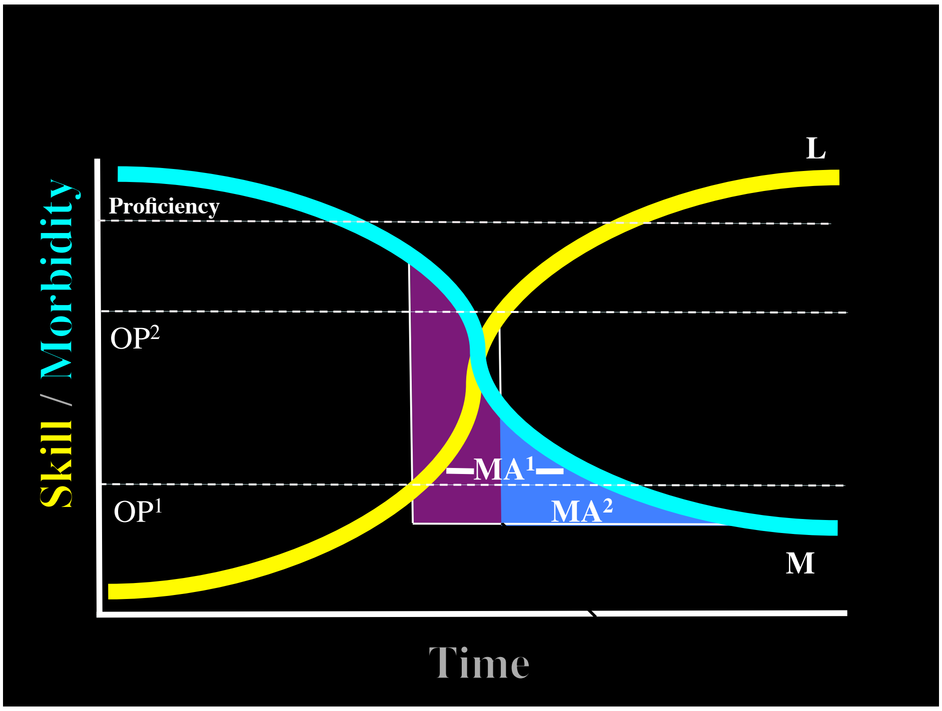


Figure 3. L = learning curve; M = morbidity curve; OP1 = trainee 1 beginning to perform surgery on patients while relatively low on the learning curve; MA1 = morbidity area for patients of trainee 1 (entire area beneath morbidity curve – purple and blue); OP2 = trainee 2 who begins to perform surgery on patients when relatively higher on the learning curve; MA2 = smaller associated morbidity area (blue) for trainee 2).Research into laparoscopic surgery complications in general supports those principles and has revealed that the majority of a surgeon's complications occur in the early segment of their laparoscopic experience.[16] Complications for a laparoscopic cholecystectomy are most likely to occur during the first 50 MIS cases with the greatest risk related to the first 10 cases. Acceptable skill for basic laparoscopic surgery may require between 10 and 50 procedures[16], cystoscopy 25 - 100[17], and gastrointestinal endoscopy as many as 300 procedures[18].

While it has been shown that practitioners with larger surgical volumes tend to have fewer complications,[19-21] the extent of morbidity and cost associated with acquiring that experience is unknown. Differences in innate abilities and individual performance variability[22-24] are at least partly responsible for each individual having a unique learning curve for any particular skill or procedure. Although surgeons in practice will always continue to improve and refine their skills, the ultimate goal would be for the trainee to have reached an advanced level of skill proficiency prior to operating on live patients. The investigation of a structured training and assessment (STAC) curriculum for the performance of a laparoscopic cholecystectomy has shown that a significant portion of the learning curve can, indeed, be transferred from the operating room to the simulation laboratory.[25] Follow-up practice and training are also necessary to maintain proficiency and prevent skill degradation. In a review of laparoscopic complications 3 months following initial training, surgeons who performed procedures without additional training were more than 3 times as likely to have at least one complication compared with surgeons who sought additional training.[4] Similarly, at 12 months, surgeons without further training were over 7 times more likely to have a complication than those who engaged in additional instruction. Further, at both 3 and 12 months, laparoscopic complication rates of individual surgeons demonstrated a significant inverse correlation with the number of laparoscopic procedures performed during that period.

1.4 Surgical Simulation

Over 2 decades ago, Satava[26] appreciated the potential for virtual reality (VR) simulation to enhance surgical skills training in an effective and cost efficient manner. Expense, insufficient computing power, limited sensory feedback, and lack of an understanding of how to harness and utilize simulation initially limited its impact and implementation into surgical training. At that time, general surgery and in particular the laparoscopic surgery discipline demonstrated the effectiveness of simulation for training and skills transfer to the operating room.[27-30] Orthopedic surgery and the subspecialty of arthroscopy are just beginning to study the role of simulation for surgical education. Previously, bench top (BT) models have been evaluated[31-33] and, while they lack a degree of face validity[34], they can be very effective and economical depending on the specific task being trained, i.e. endoscopic knot tying.[35] One of the negative aspects related to the use of physical models is the requirement for a significant amount of faculty time to provide instruction and particularly, feedback. Limited use is another drawback. Cadaver training offers the opportunity to practice using 'real' anatomy.[36] Associated disadvantages include cost, procurement, single use, disease transmission, disposal, and comorbid pathology compromising the region of anatomical study. In an investigation using the Arthroscopy Surgical Skill Evaluation Tool (ASSET) and time to completion, an assessment of training value was conducted comparing cadaver versus VR training for knee arthroscopy at a single institution.[37] While the cadaveric based training was twice as efficient, the VR simulator was more cost-effective if employed at least 300 hours per year.

VR simulators have been studied[33, 38-48], although the definition of VR is somewhat loosely defined and may or may not include haptics and the simulated feeling of touch. For active haptic devices, a computer generates and controls artificially created mechanical resistance caused by the trainee's actions. Passive haptic feedback occurs with the use of the instrument itself, i.e. passing and manipulating a hook probe through simulated skin and soft tissues.[49] The incorporation of active haptics for VR simulators is both challenging and costly and is just beginning to be implemented. The lack of haptic feedback may introduce a level of risk for the trainee who, lacking tactile feedback, may be unaware that they are potentially contacting anatomic structures and creating tissue damage (i.e. by aggressive handling of delicate tissues, over-penetration of an instrument, or inappropriate tool trajectory).

Arthroscopic procedures are particularly suited to the use of VR simulation, as a 2D image that is projected on a monitor constitutes representation of a 3D anatomical region. As the complexity of arthroscopic procedures has evolved, mastery of the required skills has become a greater challenge and requires more extensive training. It has been difficult for orthopedic residency programs to provide sufficient arthroscopic training in the curriculum due to the associated costs, time constraints, and potential patient safety issues. As a consequence, concerns exist as to whether residents are adequately prepared to perform arthroscopic surgery on entering practice.[50] To be valuable and effective in helping to train individuals to perform arthroscopic procedures safely and effectively, it must be shown that simulators are able to train essential skills that can be transferred to the OR. In addition to training for fundamental arthroscopic skills[51-53], simulation as a training tool for arthroscopy has predominantly been studied in the knee[31, 37, 40, 41, 43, 45, 54, 55], shoulder[42, 52, 54, 56, 57], and hip[58]. Although simulator based training for arthroscopic surgery in particular holds great promise and has ardent supporters, the effectiveness of surgical simulation for arthroscopic surgery remains largely unproven. To date, research metrics are almost exclusively surrogates (time to completion, instrument motion analysis, frequency of collision with anatomic structures, etc.) for actual proficiency in accomplishing a specific arthroscopic task or surgical intervention. In all probability, as various simulations and simulators become better at emulating the necessary skills, they will become more effective in preparing the surgeon to optimize treatment for their patients. In addition, it is expected that the efficiency of the training process using simulators will be enhanced and prepare trainees well in a shorter period of time. Greater surgical proficiency and the reduction of operative errors are certain to reduce patient morbidity and the costs associated with managing related complications.

1.5 Validating Surgical Skills

We must search for and identify how best to employ simulation to enable the trainee to master the requisite skills. It is imperative, however, that in our investigations, validation methodologies, rating scales, and statistical tools be rigorously and accurately employed lest we be led astray by acceptance and assumption too freely given. The concept of 'validity' (i.e. face, content, construct, and concurrent validity) can be applied with a brush too broad and should not be accepted as though it were a clearly and rigidly defined standard. Construct validity refers to the ability of a simulator to discriminate between different levels of technical expertise. For most simulator construct validity studies to date, however, global rating scales or checklists along with time to completion, motion analysis, and avoidance of collisions have been used as proficiency metrics. [38, 59-61] These surrogate metrics of performance, however, may only loosely correlate with the ability to perform a specific surgical task or procedural component well. Too great of a reliance has been placed on global rating scales (GRS) as they relate to the establishment of construct validity. In 1932, Likert described a global rating scale in an effort to assess a 'range of attitudes'.[62] The numbers on the scale were used to clarify the breadth of responses with every other digit on the spectrum having no description at all. These scales were not designed to be, nor are they objective assessments. Their use involves a significant component of subjective interpretation by the scorer. The fact that a numeric scale has been assigned does not, in and of itself, 'objectify' the scoring. Although the digits identified lend themselves to statistical analysis, the basis for their assignment in this instance remains subjective. One of the challenges imposed on subjectively based assessments such as those used with GRSs is that it becomes very difficult to gain acceptable inter-rater reliability (IRR) among blinded assessors.[63, 64] What performance would justify a score of 2 or 4? Even with 1, 3, and 5, the identifiers are subjective descriptions rather than objective definitions. What "graceful and dexterous", or "confident clear economy" means to one reviewer might mean something different to another observer because subjective interpretation is required. In contrast, objective assessments which use clearly defined (not described) metrics that enable binary scoring, i.e. the event either 'was' or 'was not' observed to occur, provide a much more robust evaluation.[65] Binary scoring makes the calculation of a very precise (and high) IRR achievable, i.e. there was either agreement or disagreement between 2 blinded raters on each item of an evaluation (IRR = # agreements / # agreements + # disagreements [agreements + disagreements = the total number of items being scored]). Since each specific metric either was or was not agreed upon, the differences in scoring are clear and precise – there is no averaging or pooling of responses.

As noted above, performance metrics often include an analysis of motion, which assesses economy and efficiency more than the acquisition of technical skill. While motion analysis adds a degree of objectivity to the evaluation of training tools, it is not synonymous with evaluating the skills necessary to perform a procedural technically well. Documenting the avoidance of collisions, however, is a more useful metric as the creation of errors is often the best discriminator between novice and experienced operators.[66]

Predictive validity is determined by the extent to which the performance exhibited on a simulator correlates with surgical performance in the OR. Evidence for transfer of training pertaining to arthroscopic techniques is just beginning to be evaluated. Data exists, however, to support that performance on a VR simulator does correlate with the extent of surgical experience and operative skill.[67-69] Further, skills exhibited on a VR simulator have a strong correlation with demonstration of the same skills in a cadaver model.[52] Correlations, however, don't necessarily prove that skills acquired during training directly transfer to the OR. Although transfer of training has been demonstrated for laparoscopic surgery[22, 70-73], only one study has evaluated the effectiveness of transfer of simulator trained skills for arthroscopy to the operating room.[74] The investigation revealed that when performing a diagnostic knee arthroscopy on live patients, trainees who were involved in a simulator training program outperformed those who underwent traditional training without simulation. We must be careful, however, not to extrapolate to freely from diagnostic efficiency to arthroscopic surgical skill in general. Surgical skill as it relates to obtaining a clear surgical field of view, preparing a repair site, or passing and tying sutures arthroscopically requires a much more sophisticated set of talents than a straightforward diagnostic procedure

1.6 Human Factors

The challenges of performing endoscopic surgery are related in large measure to the loss of 3D imaging that creates perceptual, spatial, and psychomotor difficulties. 3D information must be interpreted from a 2D monitor. In the absence of binocular input, limited information is provided as to the depth of images displayed on the screen. In addition, compared to the natural view afforded by open surgery, scaling difficulties due to magnification and image degradation contribute to perceptual impairments. Pistoning or moving the arthroscope closer to or further away from structures also alters the perspective. As the lens moves closer, the FOV becomes restricted and limits the number of reference structures available for orientation. A number of additional challenges are unique to endoscopic surgery. The paradoxical movement of instruments is due to the 'fulcrum effect'[75] caused by the passage of tools through a tissue plane such as the body wall or more superficial tissues surrounding a joint. Thus, movement of an instrument handle to the right results in the instrument tip moving to the left in the body cavity and on the video monitor. Triangulation refers to the ability to place the working component of separate instruments in the same operative space and field of view (FOV). To accomplish that task, the operator must conceptualize the surgical field in 3D, accurately appreciate the location of instruments in space, and manipulate them into the desired location even when the tools are not initially in view. Finally, tactile feedback degrades when tissues are manipulated with an instrument rather than the gloved hand.

The most significant contribution, however, to the difficulties in attaining proficiency for minimally invasive endoscopic procedures is insufficient training.[76] Regardless of how sophisticated the instrumentation or how autonomous it is designed to work, operators must still be trained to use it. Although the brain will eventually tend to automate to the performance of many of the skills needed to perform endoscopic surgery, time, meaningful practice and repetition are required to master the necessary skills. Furthermore, the more novice and inexperienced the operator, the greater the portion of finite attentive resources that are consumed to focus on the recognition of anatomy, maintain an acceptable FOV, and assess pathology as well as accurately deliver, manipulate, and control instruments.[77-80] For the more skilled surgeon, these tasks are automated for the most part. Thus, the experienced surgeon retains sufficient attentive resources to recognize potential difficulties, avoid pitfalls, and anticipate subsequent steps. For the novice, those perceptions are sacrificed at the expense of performing routine arthroscopy tasks. It is highly probably that as the technical complexity of surgical procedures intensifies, particularly with endoscopic and MIS, not all residents and or surgeons in practice will have the prerequisite physical, conceptual, or 3 dimensional abilities to become accomplished surgeons for technically demanding procedures.

1.7 Arthroscopic Bankart Repair

For unidirectional anterior shoulder instability, capsulolabral / ligamentous disruption from the anterior and inferior glenoid is a significant contributor to recurrent anterior subluxation or dislocation of the shoulder. This constellation of pathology has been termed the Bankart lesion.[81-83] Traditionally, the repair of the capsulolabral tissue to the glenoid was performed in an open manner with suture tunnels through the rim of the glenoid[83, 84], or suture anchors placed along the glenoid margin.[85, 86] Methods were developed to perform a suture anchor Bankart repair arthroscopically.[87, 88] As the arthroscopic technique has been refined, it has become a commonly accepted method for effectively stabilizing the shoulder in the absence of bone deficiency.[89, 90] Variables do exist according to surgeon preference with respect to whether the patient is placed in the supine or beach chair orientations, the specific placement of the portals, and the choice of viewing perspective.[91] The treatment of contact athletes and those with glenoid bone deficiency or multidirectional laxity is more controversial.[92] Depending on the extent of the pathology, 3 suture anchors are routinely used to accomplish a Bankart repair.[89] For the AANA Copernicus Initiative designed to investigate the merits of proficiency based progression training, we elected to use an arthroscopic suture anchor Bankart repair. It is a procedure, which is commonly performed, employs a predominant technique, is arthroscopic and able to be captured on video, and employs skills common to many arthroscopic procedures including tissue manipulation, debridement with a shaver, suture passage, and endoscopic knot tying.

The first 4 publications presented in Chapter 2 of this thesis present papers reviewing the evaluation and diagnosis of shoulder instability, recommendations for shoulder arthroscopy set-up / approaches, considerations for arthroscopic vs. open Bankart repair, and a presentation of the principles and technique for an arthroscopic Bankart repair.

1.8 AANA Copernicus Investigation

In order to accurately study the merits of PBP training, 3 tools needed to be developed and rigorously validated; a metrics tool[65], a training tool[93], and an assessment tool[66]. The 'metrics tool' was needed to accurately and objectively evaluate the performance of an arthroscopic Bankart repair. The 'training tool' or simulator was needed to enable trainees to practice the necessary steps / skills for the Bankart procedure, to identify specific skill deficiencies, and to serve as an intermediate evaluation tool to verify that the necessary skills had been mastered before advancing to the use of a cadaver for training. The shoulder model simulator afforded the trainee the opportunity to commit errors in a consequence-free environment, to learn from them, and make corrections. Finally, the 'assessment' tool was needed to be able to accurately and objectively evaluate the performance of an arthroscopic Bankart repair in a cadaver shoulder. This final tool provided the ability to compare operative performance for 4th and 5th year orthopedic residents who participated in one of 3 different surgical training protocols, and to determine the relative effectiveness of the different curricula. The development and validation of these 3 essential tools and a knot testing protocol are reported in the 5th – 8th publications contained in this thesis, detailing the AANA Copernicus Initiative: 1) "Metric Development for an Arthroscopic Bankart Procedure - Assessment of Face and Content Validity" (Chapter 3); 2) "The Bankart Performance Metrics Combined with a Shoulder Model Simulator Create a Precise and Accurate Tool for Measuring Surgeon Skill" (Chapter 4); 3) "The Bankart Performance Metrics Combined with a Cadaver Shoulder Create a Precise and Accurate Assessment Tool for Measuring Surgeon Skill" (Chapter 5); and 4) "Objective Assessment of Knot-Tying Proficiency With the Fundamentals of Arthroscopic Surgery Training Program Workstation and Knot Tester" (Chapter 6).

The 9th and final publication, "Results from the Arthroscopic Association of North America (AANA) Copernicus Initiative; A multicenter, prospective, randomized, blinded trial of proficiency-based progression training employing simulation for an arthroscopic Bankart procedure"[94] reports the findings of a study comparing; 1) AANA's traditional method of training residents to acquire arthroscopic skills representing the apprenticeship model, 2) a simulator-enhanced curriculum, and 3) a proficiency based progression curriculum coupled with the use of a shoulder model simulator (Chapter 7).

1.9 Training to Proficiency

Many MIS and endoscopic surgeons acquired their skills in the operating room. Currently, a typical preparation for a trainee learning an advanced MIS procedure would involve attending a course with lectures on indications and technique along with a video review of a master surgeon performing the techniques on a live patient in the OR. A practical skills laboratory would follow. The trainee would often then return home to his or her own practice and begin to perform the procedure. In some instances, a mentor might be available to proctor and assist the surgeon on the first several cases. This abbreviated approach to preparation fails to ensure that the requisite skills were mastered and is no longer acceptable to patients or the surgical profession.[6] Ideally, learning to perform a new skill would involve obtaining a cognitive understanding of the specific techniques involved as well as the appropriate clinical application for the procedure. Further, knowledge of not only what to do, but perhaps even more importantly, what to avoid in executing the technique would be essential. The component skills would then be learned, practiced, and mastered in the laboratory setting. Only when the skills could be demonstrated to a previously established objective performance standard or benchmark would the surgeon proceed to perform the techniques in the OR on live patients.

A framework for a surgical skills training curriculum has been developed and emphasizes 5 critical elements[95]: task deconstruction, creation of an evaluation tool, formulation of a comprehensive curriculum, an assessment as to whether the learning achieved transfers to a real environment, and the establishment of tools for credentialing. In one related investigation, residents participating in a structured training and assessment curriculum (STAC) demonstrated both superior technical proficiency in the OR as well as nontechnical skills compared with conventionally trained first and second year surgical residents.[25] The curriculum consisted of case-based learning, proficiency-based progression virtual reality training, laparoscopic box training, and OR participation. In a separate report, a proficiency based progression curriculum for obtaining the skills necessary to perform a laparoscopic cholecystectomy was shown to produce superior OR performance.[22] In that randomized, blinded trial, residents trained on a MIST-VR simulator to an established performance level on two consecutive trials performed a laparoscopic cholecystectomy 29% faster with 6 times fewer errors when compared to residents who did not undergo VR training. Few PBP curricula have been developed for complete surgical procedures.[95]

1.10 Outcomes Based Medical Education

A belief that medical education and training must move from process based (number of rotations, case exposure volume, etc.) to outcomes based assessment (demonstration of skill proficiency) is evolving.[96] It is neither possible to obtain a uniform level of performance nor quality-assure proficiency without establishing clearly defined goals and procedural skill benchmarks. The focus for education and surgical training research should shift from an emphasis on validating simulators to developing and validating comprehensive evidence-based curriculums capable of training the needed skills for surgical procedures in their entirety.[25] Despite the tremendous potential for simulation to contribute to surgical skill development, we must keep in mind the father of surgical simulation, Dr. Satava's admonition – "It is not the simulator, it is the curriculum".[97] The AANA Copernicus

Initiative rigorously investigates the merits of the evolving paradigm shift from apprenticeship-like training to one of proficiency-based progression training for endoscopic surgical skills. Based on the results of this investigation, AANA is proceeding with the task deconstruction and metric development for three additional arthroscopic procedures: 1) anterior cruciate ligament reconstruction for the knee, 2) rotator cuff repair for the shoulder, and 3) acetabular labral repair for the hip.

2. Shoulder Instability Management

2.1 Shoulder Instability Evaluation “The Overhead Athlete: How to Examine, Test, and Treat Shoulder Injuries. Intra-articular Pathology” Angelo R, Arthroscopy: The Journal of Arthroscopic and Related Surgery, Vol. 19, No 10 (December, Suppl. 1),2003: pp. 47-50 [92] Appendix 1

(Candidate is the sole author of this publication)

2.11 Etiologies for Shoulder Instability in the Overhead Athlete

The overhead athlete typically places demands on his or her shoulder that far exceed activities for the normal population. Intra-articular pathology in the overhead athlete includes microinstability (often anteroinferior), SLAP tears, internal impingement, biceps tendinopathy, and partial articular surface rotator cuff tears. The fact that these pathologies are interrelated and often coexist creates a challenge in identifying specific etiologies. The pathomechanics involved in the overhand throwing motion and the internal impingement phenomenon in particular are complex. Treatment of these entities in the past has produced varied outcomes. As the results of valuable research accumulate, more unified models are evolving that begin to explain the breadth of clinical findings better. With a more complete understanding arrives the promise of more effective treatment strategies.

2.12 Microinstability

Progressive acquired capsular and ligamentous attenuation, unrelated to specific traumatic events, may create progressive dysfunction in the shoulder. Symptomatic microinstability can be anteroinferior, straight anterior, or anterosuperior. Although posterosuperior instability has been described, the pathology is not truly one of instability. Rather than being caused by capsular and ligamentous laxity, it is due to posterior capsular tightness and aberrant posterosuperior glenohumeral translation. This issue is addressed later with internal impingement.

Repetitive overhead sports activities including throwing, volleyball, tennis, and gymnastics may create anteroinferior glenohumeral instability, which is the most common acquired symptomatic laxity. On examination, the “load and shift” test identifies excessive anteroinferior laxity but is difficult to quantitate. The anterior Jobe relocation test is a helpful sign. Patient apprehension is created when the examiner places the involved shoulder in abduction and external rotation. The test is considered positive if a posteriorly directed force applied to the proximal humerus by the examiner eliminates the patient’s apprehension. Arthroscopic findings may include a positive “drive-through” sign in which the arthroscope can be passed from posterior to anterior across the shoulder joint through a generous glenohumeral interval. Abrasion or wear of the labrum and glenoid chondromalacia are footprints of excessive translation of the humeral head on the glenoid.

Treatment options include an arthroscopic capsular plication that allows the surgeon, with some “guestimation,” to roughly quantitate the magnitude of capsuloligamentous shortening and volume reduction that is produced. A rasp or whisker shaver blade is used to lightly excoriate the capsule along a 1.5-cm band adjacent to the labrum. A tuck or fold of capsule is then created by inserting a curved, cannulated suture hook into the capsule 1.0 to 1.2 cm lateral to the glenoid rim, passing it immediately deep to the capsule, and exiting approximately 5 mm lateral to the glenoid rim. The suture hook is then delivered beneath the intact labrum to create a tuck or fold of capsule. Long-term follow-up is not available for this technique, but early results are encouraging and the risk is relatively low because minimal tissue destruction is created.
A second option to stabilize the shoulder is to perform a thermal capsulorrhaphy. A great deal of debate has surrounded this technique in the past several years regarding its safety and efficacy. The limited clinical studies that are available show widely varying success rates.

Levitz et al.[98] reported an 85% success rate in 122 throwing athletes when thermal capsulorrhaphy was used as an adjunctive tool. Others have reported failure rates from 15% to 60% depending on the primary clinical pattern of instability.[99] The visible tissue response to the heat probe is quite variable and is an unreliable guide to the magnitude of the thermal effect on the tissues. Reports of permanent capsular damage have led to recommendations for "striping" or creating a grid pattern with the wand rather than simply painting the tissues. It is advisable to leave as much healthy untreated tissue as that which is thermally altered. Capsular necrosis, stiffness, and axillary nerve injury are concerns and this modality must be used with caution.[100]

The open capsulolabral reconstruction has been reported to permit return of up to 75% of professional baseball players for at least one full season subsequent to shoulder repair. A transverse incision rather than a vertical detachment of the subscapularis avoids much of the morbidity associated with an open procedure and permits earlier and more aggressive rehabilitation. A "pants over vest" imbrication of the capsule along the glenoid rim is created to reduce the capsular volume and restore stability to the shoulder.

2.13 Straight Anterior Instability

Straight anterior glenohumeral instability is relatively uncommon and may result from tearing of the mid-labrum and detachment of the middle glenohumeral ligament origin. In addition to repetitive overhead activities, glenohumeral hyperextension at neutral rotation and 45° abduction may also result in direct anterior instability. Associated partial articular surface rotator cuff tears are identified in approximately 2/3 of patients with this diagnosis. Examination findings include a positive load and shift test and a positive anterior Jobe relocation test. A positive Whipple test is often present if there is associated supraspinatus tearing. This test is positive if pain results from resisted elevation of the arm in the scapular plane. Treatment includes a suture anchor repair of the anterior labrum and associated middle glenohumeral ligament complex along with arthroscopic debridement of the articular surface rotator cuff tear.

2.14 Anterosuperior Instability

Anterosuperior instability is also relatively rare. The eponym, SLAC (superior labrum anterior cuff) has been used to describe this lesion.[101] The constellation of associated pathology includes an anterosuperior labral lesion, a superior glenohumeral ligament tear, and a partial articular surface supraspinatus tear. Occasionally, chondromalacia of the anterosuperior glenoid quadrant is also present. Approximately 50% of the patients who have been recognized with this entity have been overhead athletes, and half have sustained significant shoulder trauma. The superior labral and glenohumeral ligament damage is either repaired or debrided along with the rotator cuff.

2.15 Slap-Biceps Lesions

The superior labrum is typically more meniscoid in configuration than the inferior region. The biceps anchor has a variable attachment to the supraglenoid tubercle with approximately 25% to 50% attaching to the bony tubercle and 50% to 75% attaching predominantly to the posterosuperior labrum. Normal variants include an anterosuperior sublabral foramen and a cord-like middle glenohumeral ligament or "Buford complex". Snyder[102] was the first to classify superior labral tears. Type I consists of superior labral fraying (20%); type II, biceps-labral detachment (55%); type III, a superior bucket-handle tear (9%); and type IV, a bucket handle tear with extension into the biceps tendon (10%). Complex, uncategorized tears make up the remainder (5%). Microinstability, internal impingement (discussed later), forced external rotation of the abducted arm,[103] and traction on the long head tendon of the biceps during deceleration of the throwing arm are possible mechanisms of injury creating SLAP lesions in the overhead athlete.[104, 105]

Many tests have been described to diagnose superior labral tears but they often lack sensitivity and specificity. The following examination tests tend to be more reliable. The posterior Jobe relocation test (for posterior or superior SLAP lesions) is initiated by placing the patient's arm in 90° abduction and full external rotation. The test is considered positive if the posterosuperior pain is relieved when the examiner applies a posteriorly directed force to the upper arm. The O'Brien test is performed by placing the patient's arm in 90° flexion, 25° adduction, and full internal rotation. Downward pressure on the arm applied by the examiner may create anterosuperior pain. The test is considered positive for an anterosuperior labral tear if the pain on resisted flexion is eliminated when the arm is in a similar position, but the forearm is fully supinated. Kibler's anterior slide test is initiated by asking the patient to their place hands on their hips with the elbows directed posteriorly. With one hand, the examiner supports the scapula. The other hand creates an anterosuperiorly directed force on the patient's elbow. If anterosuperior pain is generated, an anterior or superior labral tear is suspected.

A challenge is often presented to the arthroscopist in deciding which superior labral tears are significantly pathologic and require treatment. A large recess between the superior labrum and glenoid may be a normal occurrence. The findings that suggest significant pathology include hemorrhage and irregularity at the biceps anchor, superior labral arching with biceps traction, biceps "peelback" with abduction and external rotation of the shoulder, and a positive "drive-through" sign seen arthroscopically. Treatment includes debridement for type I tears, suture anchor repair for type II tears, resection versus repair for type III lesions, and repair, debridement, or biceps tenodesis for type IV tears.[106]

2.16 Internal Impingement

The constellation of pathology found with internal impingement includes posterosuperior SLAP tears, a partial articular surface tear of the posterior supraspinatus, and posterosuperior glenoid chondromalacia. Walch et al.[107] were among the first to describe this entity. Contact between the greater tuberosity and posterosuperior glenoid may occur normally in full abduction and external rotation. It was believed, however, that the repetitive frequency and intensity, with which it occurred, especially during throwing, led to labral and rotator cuff pathology. It was also believed that decreased humeral retroversion could exacerbate the problem. Jobe[108] and Davidson et al[109] attributed the pathologic findings of internal impingement to acquired anteroinferior microinstability that can compromise the obligate posterior rollback of the humeral head during abduction and external rotation.[110] The resulting anterior translation and lack of rollback of the humeral head were believed to permit increased impact of the greater tuberosity on the posterosuperior glenoid. In addition, hyper-angulation of the glenohumeral joint in the transverse plane was thought to increase the frequency and magnitude of the greater tuberosity-rotator cuff contact on the posterosuperior glenoid. Kibler[111] suggested that a loss of scapular synchrony with inefficient scapular elevation and retraction also contributed to hyper-angulation of the glenohumeral articulation. Eventually, with repetition, the increased stress on the anterior capsuloligamentous structures was believed to create an acquired anteroinferior microinstability. Components of several of these models likely coexist in any one particular shoulder patient suffering from internal impingement.

More recently, Burkhart et al.[112] offered a model that unifies a number of these concepts used to explain internal glenohumeral impingement. A key finding thought to initiate the pathologic cascade is a glenohumeral internal rotation deficit (GIRD) due to a contracted posterior capsule. As the arm moves into abduction and external rotation during the throwing motion, the contracted posterior capsule “slings” beneath the humeral head. After the elongation in the posterior capsule reaches it limit, the humeral head then begins to “roll up” the capsule much like a tire on a rope that results in an aberrant posterosuperior shift of the humeral head. This shift creates shear forces that produce posterosuperior labral tearing and glenoid chondromalacia. Accompanying the posterosuperior shift in the axis of humeral head rotation is a pseudolaxity of the anterior capsule. As the humeral head translates posterosuperiorly, it no longer “fills” the anterior capsule and results in redundancy of the capsule. The anterior pseudolaxity was thought to permit hyper-external rotation of the shoulder. There are 2 consequences of the hyper-external rotation of the humeral head. First, along with excessive torsion of the biceps tendon, the vector of the tendon becomes more posteriorly directed than normal and leads to a “peelback” of the posterior and superior labrum.[113, 114] Secondly, excessive torsion of the rotator cuff may contribute to tearing of the articular surface fibers. Finally, a “break in the ring” of the posterosuperior labrum (circle concept) is thought to add to the anterior capsular pseudolaxity that may be manifested as a positive “drive-through” sign during arthroscopy. Examination findings of significance include a loss of internal rotation greater than 25° compared to the normal side when the arm is in 90^0 of abduction. In addition, a positive posterior Jobe relocation test is often present as previously noted. Excessive anteroinferior glenohumeral translation may be detected but is often difficult to quantitate.

The initial treatment is directed toward activity modification, nonsteroidal anti-inflammatory medication (NSAIDs), focused posterior capsule stretching (sleeper stretches), and rotator cuff and periscapular strengthening. If posterior capsular stretching is unsuccessful, a limited arthroscopic posterior capsular release may be indicated for a small number of patients.[115] If significant anteroinferior microinstability is present, consideration may need to be given for an arthroscopic capsular plication, thermal capsulorrhaphy, or open anterior capsulolabral reconstruction.[116] When a posterior or superior SLAP

II lesion is present, it should be repaired and will usually eliminate the drive-through sign seen at diagnostic arthroscopy.

2.17 Partial Articular Surface Rotator Cuff Tears

The etiology for partial articular surface rotator cuff tears is likely multifactorial but may include repetitive traction on articular surface fibers during deceleration of the throwing arm,[117] and internal impingement as described above. Biomechanically, the articular surface of the cuff may be more likely to fail under tensile rather than compressive forces.[118] A grade 1 tear describes a defect less than 3 mm (~ < 25% of the rotator cuff thickness); grade 2, 3 to 6 mm defect (< 50%); and grade 3, greater than 6 mm defect (> 50%). Treatment includes an arthroscopic debridement to stable, healthy rotator cuff tissue. For the few that are grade 3 tears, consideration may need to be given for an arthroscopic or mini-open rotator cuff repair. An arthroscopic subacromial decompression may be considered part of the management for grade 1 and 2 articular surface tears as suggested by Payne et al. [119] Great caution should be exercised if instability is a component of the pathology because a subacromial decompression may aggravate the patient's symptoms.[120]

2.18 Summary

The pathology in the overhead athlete's shoulder is often complex, with substantial overlap between microinstability, labral pathology, internal impingement, and partial articular surface rotator cuff tears. An accurate diagnosis demands careful integration of the history, physical examination findings, imaging studies, examination under anesthesia, and findings at diagnostic arthroscopy. The treatment options described have relatively little intermediate or long-term follow-up and remain controversial.

2.2 Shoulder Arthroscopy Set-up: "Arthroscopic Setup: Approaches and Tips for Success" Angelo R. In: Johnson D. ed. Operative Arthroscopy, 4th Edition. New York: Lippincott, 2013. Appendix 2

(Candidate is the sole author of this publication)

2.21 Patient Positioning

The safety and ease with which an arthroscopic shoulder procedure is accomplished frequently relates to how the patient is positioned and how accurate and utilitarian are the portals that have been established. Although minor variations exist, most surgeons employ either the lateral decubitus or the beach chair position for the patient, and each has its proponents. The choice is largely influenced by the familiarity gained while the surgeon was learning shoulder arthroscopy, the ease and anticipated likelihood of converting to a mini-open procedure, and the availability of surgical assistants and supportive devices for arm positioning. Equipment is readily available to facilitate the use of either position.

2.211 Lateral Decubitus Orientation

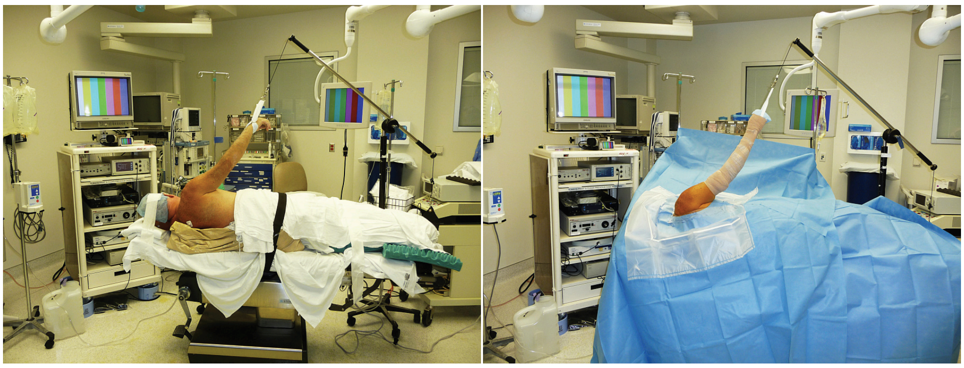
The supine position is used during the induction of general anesthesia. The patient is then repositioned in the lateral decubitus orientation on a vacuum bag (Figure 2.21, B). A gel pad can be layered on top of the beanbag, particularly if there is the anticipation that the procedure may be prolonged. A soft axillary roll is placed beneath the upper thorax to minimize direct pressure on the axilla, and the head is supported in a neutral orientation. The patient's thorax is allowed to roll back approximately 15 degrees orienting the glenoid roughly parallel with the floor. The vacuum bag is then evacuated to maintain support. All bony prominences must be appropriately padded, in particular the fibular head to protect the peroneal nerve. The operating table is then rotated to position the anesthesiologist and related equipment in an area at the middle of the operating table near the patient's abdomen (across the table from the surgeon). The surgeon is thus provided with unrestricted access to the involved shoulder. Monitors are located for easy viewing.

If the primary procedures are to be performed in the subacromial region, the primary monitor is positioned superior and anterior to the patient's head. A secondary monitor for use by the surgical assistant may be located in front of and above the patient's abdomen. When the work to be completed is primarily in the glenohumeral joint, i.e. a Bankart or SLAP repair, the monitor is set across from the surgeon near the patient's abdomen as the general viewing direction for glenohumeral procedures is anterior rather than superior. The arm is supported in 30 to 40 degrees of abduction and 15 degrees of forward flexion using 10 lb. (4.5 kg) to suspend rather than place significant traction on the arm. This shoulder position is varied during the case depending on the access necessary to specific locations. Numerous sterile sleeves and gauntlet devices are commercially available to support the arm.

Arthroscopic Bankart repairs may be facilitated by directing 10 lb. (4.5 kg) of accessory traction laterally (perpendicular to the proximal humerus) to distract the shoulder and improve access to the anterior aspect of the glenohumeral joint. Alternatively, an assistant can accomplish a similar manual maneuver. A routine sterile prep and draping are then performed. The lateral decubitus method eliminates the need for an assistant or mechanical device to support the arm. Internal and external rotation of the suspended arm affords acceptable access to the entire rotator cuff. If range of motion is to be assessed at the completion of surgery, i.e. following a Bankart repair, the arm is removed from suspension for the motion exam while maintaining sterility of the sleeve's suspension loop.

While working in the glenohumeral joint, the monitor view of the glenoid is typically oriented parallel with the floor. When working in the subacromial space, however, the surgeon may elect to either maintain this orientation (the acromion is vertical) or rotate the camera head to view the acromion in a position parallel with the floor (as it would appear with the patient standing).

If converting to an open procedure through a standard deltopectoral approach for the glenohumeral joint, subscapularis, or biceps tendon, the unsterile portion of the suspension apparatus is removed and the patient's arm is allowed to rest on the ipsilateral hip. The vacuum bag is at least partially inflated (softened) and the patient allowed to roll back into a more supine position. The draw sheet is used to center the patient on the operating table. The table is then configured to a gentle beach chair orientation with acceptable position and support for the head and neck verified. Although it is unnecessary to completely re-prep and re-drape, it is prudent to replace the clean, sterile barrier sheet anterior to the shoulder to shield the anesthesiologist and related equipment.



A B

Figure 2.21 A: Patient positioned in the lateral decubitus orientation: anesthesia setup is near the abdomen. Dual monitors are helpful, particularly to provide a comfortable view for the surgical assistant. **B**: Once draping is complete, easy access to the entire shoulder is afforded; the arm is "suspended with 10 lb. through a disposable arm sleeve.

If the surgeon elects to convert to a mini-open approach to the subacromial region, repositioning is unnecessary although some prefer to tilt the table posteriorly toward the surgeon to improve access to the anterior shoulder. An approach to the supraspinatus and infraspinatus is readily obtained by extending the lateral subacromial (LSA) portal superiorly. An absorbable suture is introduced transversely through the deltoid at the inferior extent of the portal defect to prevent inadvertent distal extension of the deltoid split and iatrogenic injury to the axillary nerve. The deltoid is then divided proximally along its fibers to the level of the acromion.

2.212 Beach Chair Orientation

Some surgeons prefer the beach chair position due to its more anatomic orientation, which conforms to the familiar open approach.[121] The patient's thorax is positioned to permit the involved shoulder to overhang the side of the table. Once the hips are flexed 70° to 80° and the legs 30°, the back is elevated approximately 70°. After padding bony prominences, a vacuum pack supports the hips and thorax, but is pushed medially and displaced from the ipsilateral periscapular region before evacuating air. Alternatively, a specially designed table with a removable wing for exposure of the operative shoulder may be employed (Figure 2.22A, B). A more vertical orientation for the back will minimize the dependent position of the camera when the scope is in the posterior portal minimizing the chance that the lens will fog. However, a more upright position for the thorax increases the hydrostatic pressure gradient between the head and the brachium. The anesthesiologist sets up near the patient's uninvolved shoulder, and the viewing monitor is placed opposite the surgeon near the foot of the table. A surgical assistant or a sterile, maneuverable mechanical arm holder adjusts the position of the shoulder during the procedure, depending on the access necessary. Somewhat greater mobility of the arm and shoulder exists when compared with the lateral decubitus position.

A. B.

Figure 2.22. A: Patient positioned in the beach chair orientation; anesthesia setup is near the contralateral shoulder; a table with a removable wing affords easy access to the entire shoulder. **B:** The anterior and posterior aspects of the shoulder are readily accessed; a sterile arm positioner can be employed if desired.

The upright (anatomic) orientation for the arthroscope and monitor view is maintained while working in both the glenohumeral and the subacromial regions. Conversion to an open procedure for all regions of the shoulder is relatively simple and only requires reducing the elevation of the thorax. The vacuum pack must be at least partially inflated in order to safely change the patient's position without creating pressure points. Alternatively, a relatively more supine position for the thorax can often be accomplished by tilting the entire table into a greater Trendelenberg position.

A recent case report series identified four patients who underwent shoulder surgery in the beach chair position, which resulted in one death and three patients with severe brain damage. [122] Cerebral hypo-perfusion, rather than cardiovascular risk factors, was believed to be the cause and may be attributable to differences in blood pressure reference points. A blood pressure difference as great as 90 mm Hg between the head and the calf may exist in the sitting position based on hydrostatic factors alone. Potentially catastrophic cerebral hypoperfusion can be avoided by taking precautions including placing the blood pressure cuff on the brachium rather than the calf, [123] maintaining perioperative blood pressure values at a minimum of 80% of preoperative resting values, and ensuring that the intraoperative blood pressure is at a minimum of 100 mm Hg at the level of the head. Losses of vision and ophthalmople-gia have also been reported following general anesthesia with the patient in the beach chair position, but the exact mechanisms for this pathology are unclear. [123] Lower extremity thromboembolic events are also possible with the patient in the beach chair position and make the use of cyclical pneumatic compression cuffs around the calves prudent.

2.22 Anesthesia Choices

2.221 General Anesthesia

Both endotracheal intubation and a laryngeal mask airway (LMA) provide safe, reliable options for maintaining the airway during general anesthesia. No durable analgesia is afforded once the patient awakens, and nausea/vomiting can sometimes be difficult to manage in the perioperative period.

2.222 Interscalene Regional Block

Interscalene blocks (ISBs) provide anesthesia, muscle relaxation, and postoperative analgesia although supportive parenteral pain medication may be necessary during the immediate postoperative period. [124] An ISB can be used as the primary means of anesthesia or as an adjunct to general anesthesia. As with any invasive procedure, the risk/ benefit ratio determines its use. Proponents note its effectiveness despite the frequent need for some additional narcotic support during the immediate postoperative period and its relatively low risk of serious complications. Dedicated anesthesia teams committed to regional anesthesia and who perform a large number of blocks will help to minimize untoward events. [125] Potential serious complications have been reported including cardiac arrest, grand mal seizures, hematoma, and pneumothorax. Possible neurologic injuries include damage to the recurrent laryngeal, vagal, and axillary nerves. Phrenic nerve dysfunction is common and can give rise to significant respiratory distress. Brachial plexus pathology may include transient paresthesias (which have been reported to be as high as 9% at 24 hours and 3% at 2 weeks post-op), [126] or a brachial plexus palsy, which may be transient, require prolonged recovery or be permanent in a very small number of cases.

It is essential that the block be performed with the patient awake so they are able to provide critical feedback during administration of the block. More recently, the use of ultrasound to guide placement of the needles has added a measure of safety. Even with a successful block, the duration of pain relief averages only 12 to 18 hours following surgery, which may make pain management challenging in an outpatient setting. [124] A thorough disclosure of the potential risks should be discussed with the patient, preferably beforehand in an office setting during the preoperative visit.

2.223 Adjunctive Pain Management

The suprascapular nerve supplies 70% of the sensation to the shoulder joint. Instillation of 20 cc of 0.25% bupivacaine adjacent to the suprascapular nerve may result in up to a 30% reduction in postoperative narcotic usage and a 5-fold reduction of nausea. [127, 128] This block carries a low risk when performed with a blunt-tipped needle, and may be repeated as necessary, even in an office setting on the first postoperative day. [129] In addition, local infiltration of the portal sites with 0.5% bupivacaine leads to further reduction in pain. Pain pumps remain controversial, but have been safely used in the subacromial space provided that the glenohumeral joint is not exposed to the catheter and infiltrate. Cooling jackets using circulating ice water may also substantially improve patient comfort.

2.23 Portals

When arthroscopic portals are properly placed, they will provide the necessary field of view and instrument access to desired locations within the glenohumeral joint, acromioclavicular joint, and subacromial space. [130-134] A thorough knowledge of the regional anatomy, particularly the palpable bony landmarks, will improve safety during placement and ensure accuracy in establishing the desired portals. Compared to the glenohumeral joint, there is a greater margin of safety in creating access to the subacromial space where the use of various accessory portals is routine.

2.231 General Technique

Bony landmarks are identified by careful palpation and mapped at the beginning of the case prior to soft tissue distortion from fluid extravasation. Anticipated portal sites are referenced from the landmarks and identified using a surgical marker. All anatomical references and diagrams provided in this document are for a right shoulder with the patient in the lateral decubitus orientation. Minor adjustments to the recommended distances from anatomic landmarks may be necessary if the patient is supported in the beach chair orientation or for particularly large or small patients. As experience is gained, surgeon preference may also lead to subtle adjustments in the skin entry site for various portals. The posterior glenohumeral portal is typically established first. It is recommended that all subsequent portals be made in an outside-in manner under direct vision after first establishing the desired tract with a spinal needle. A small skin incision is made at the chosen entry site and a trocar and cannula directed along the path identical to the spinal needle and into the glenohumeral joint or subacromial space.

2.232 Glenohumeral Portals (Figure 2.23)

Posterior (P) serves as the primary intra-articular viewing portal and provides instrument access to the posterior glenoid labrum and rim, posterior capsule, and articular surface of the infraspinatus. The field of view includes the glenoid, posterosuperior humeral head, anterior capsule, biceps, superior subscapularis, glenohumeral ligaments, and articular surface of the supraspinatus and superior subscapularis tendons (Figure 2.24A, B). The entry site is 1.0 to 1.5 cm inferior and 1.0 cm medial to the posterolateral (PL) corner of the acromion.

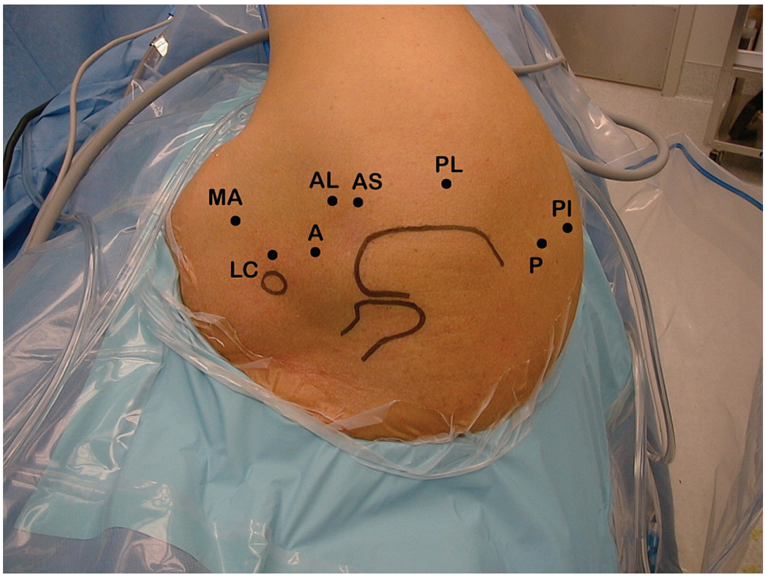
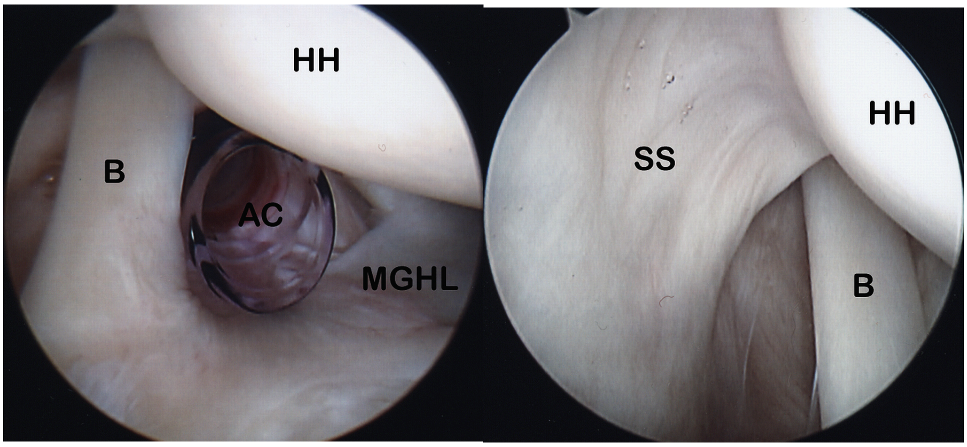


Figure 2.23. Right shoulder in the lateral decubitus orientation viewed from superior (anterior is to the left); bony landmarks are mapped out and the common glenohumeral portals are identified; P, posterior; A, anterior; PI, posteroinferior; PL, posterolateral ("Port of Wilmington"); AS, anterosuperior; AL, anterolateral; MA, midanterior; LC, lateral coracoid.

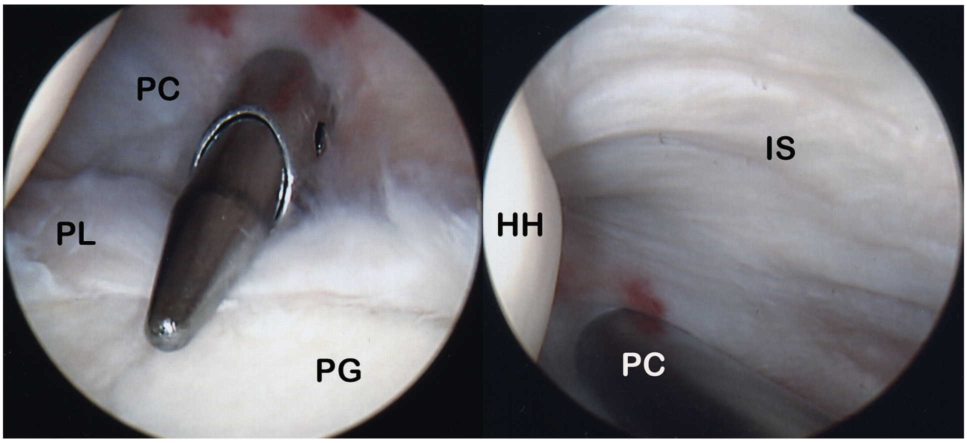
After creating a small skin incision, the cannula is introduced and directed toward the coracoid tip. If it is anticipated that this portal will be employed to drill or insert anchors along the posterior glenoid rim, the entry site must be adjusted 1 cm further lateral to account for the anterior glenoid version. This modification will enable the approach to be approximately 45° to the glenoid in the transverse plane. If this lateral modification is not made, the portal will be too "shallow" and create a risk that instruments will either skive off the articular cartilage or be directed too far medial along the glenoid neck.



A. B.

Figure 2.24. A: All arthroscopic photos are of a right shoulder with the patient in the lateral decubitus position; scope is in the posterior portal viewing anteriorly; HH, humeral head; B, biceps; MGHL, middle glenohumeral ligament; AC, anterior cannula. **B:** Scope is in the posterior portal viewing anteriorly; HH, humeral head; B, biceps; SS, capsule overlying the articular surface of the supraspinatus just posterior to the biceps.

Anterior (A) enters through the middle of the rotator interval and provides instrument access to the biceps, anterior labrum, glenoid rim, anterior and superior capsule, articular surfaces of the supraspinatus, infraspinatus, and the superior aspect of the subscapularis tendons. The field of view includes the posterior glenoid and labrum, anterosuperior (AS) humeral head, articular surface of the infraspinatus, posterior capsule, and the biceps origin (Figure 2.25A, B). The entry site is midway between the coracoid tip and the anterolateral (AL) corner of the acromion. The cannula is directed toward the center of the glenohumeral joint while viewing from the posterior portal.

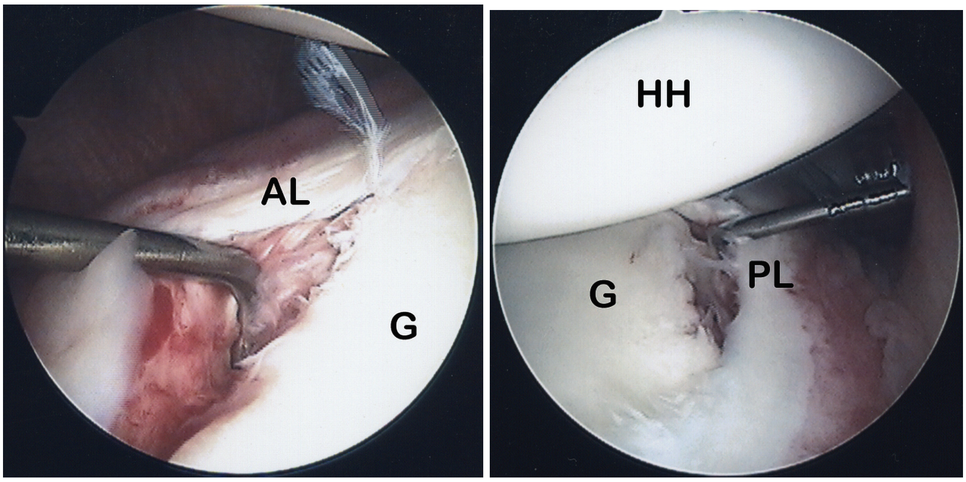


A. B.

Figure 2.25 A: Scope is in the anterior portal viewing posteriorly; PC, posterior capsule; PL, posterior labrum; PG, posterior glenoid. **B:** Scope is in the anterior portal viewing posterosuperiorly; HH, humeral head; IS, capsule underlying the infraspinatus tendon; PC, posterior cannula.

Midanterior (MA) is the preferred portal to instrument the anterior glenoid rim with drills and anchors in preparing the neck for a Bankart repair. In addition, it affords access to the anterior and inferior capsule for suture-passing instruments. The entry site is 1.5 cm lateral and 1.5 cm inferior to the coracoid tip. A spinal needle identifies the appropriate track, which, after penetrating the skin, is directed somewhat superiorly over the superior boarder of the subscapularis. A small superficial skin incision is made, and an obturator and cannula are initially directed superiorly, then over the top of the subscapularis, and finally inferiorly to enable ready access to the inferior glenoid. Instruments passing through this portal should be able to approach the glenoid at a 45° angle in the transverse plane.

AS provides a tangential view to the anterior glenoid rim and neck (for Bankart repairs), the superior insertion of the subscapularis onto the lesser tuberosity, the superior and posterior capsule, labrum, and glenoid rim (Figure 2.26A, B). The entry site is 1.0 cm directly lateral to the AL corner of the acromion, and the cannula is directed immediately anterior to the anterior boarder of the supraspinatus and then either anterior or posterior to the biceps tendon, depending on the intended primary use.



A **B.**

Figure 2.26. A: Scope is in the AS portal viewing anteroinferiorly; probe is demonstrating a Bankart lesion; G, glenoid; AL, anterior labrum. **B:** Scope is in the AS portal viewing posteroinferiorly; probe is inside a posterior labral tear; HH, humeral head; G, glenoid; PL, posterior labrum.

AL serves to enable instrument access to the posterior aspect of the coracoid, the anterior, superior, and posterior aspects of the subscapularis for release, and to the lateral boarder of the subscapularis (e.g., for use with antegrade suture-passing instruments). The entry site is 1.0 cm anterior and 1.0 to 1.5 cm lateral to the AL corner of the acromion. The cannula or instrument is directed toward the posterior aspect of the tip of the coracoid or somewhat more inferiorly toward the biceps groove.

Lateral coracoid (LC) enables instrument access to the lesser tuberosity for subscapularis repair from an intra-articular view. The entry site is 1.0 to 1.5 cm directly lateral to the middle of the coracoid tip and the instrument is then directed somewhat laterally toward the lesser tuberosity.

PL (or Port of Wilmington) facilitates placement of anchors at the posterosuperior glenoid rim for labral repair. The portal may penetrate the infraspinatus tendon. Concern has been raised regarding the defect in the tendinous portion of the rotator cuff and it is advisable to limit this portal to the smallest diameter practical for a given anchor and its preparation. The entry site is 1.5 cm anterior and 1.5 cm lateral to the PL corner of the acromion. Viewing from an anterior portal, a spinal needle is directed approximately 45° from lateral to medial to establish the proper track.

Posteroinferior (PI) provides instrument access to the posterior capsule and axillary recess for capsular excoriation and suture plication. The entry site is 2.0 cm inferior and 1 cm lateral to the posterior portal. A spinal needle is used to establish the proper track while viewing from the AS portal. Care must be taken not to err too far inferior and risk injury to the axillary nerve.

2.233 Subacromial Portals (Figure 2.27)

Posterior subacromial (PSA) is a primary viewing portal and offers instrument access to the posterior bursa, cuff, the acromion, and the greater tuberosity. The field of view includes the entire subacromial space, acromioclavicular joint, extra-articular biceps and sheath, the coracoclavicular ligaments, and suprascapular notch (Figure 2.28A, B). The entry site is the same as the posterior glenohumeral portal. The trocar is directed anterosuperiorly, immediately inferior to the inferior surface of the acromion.

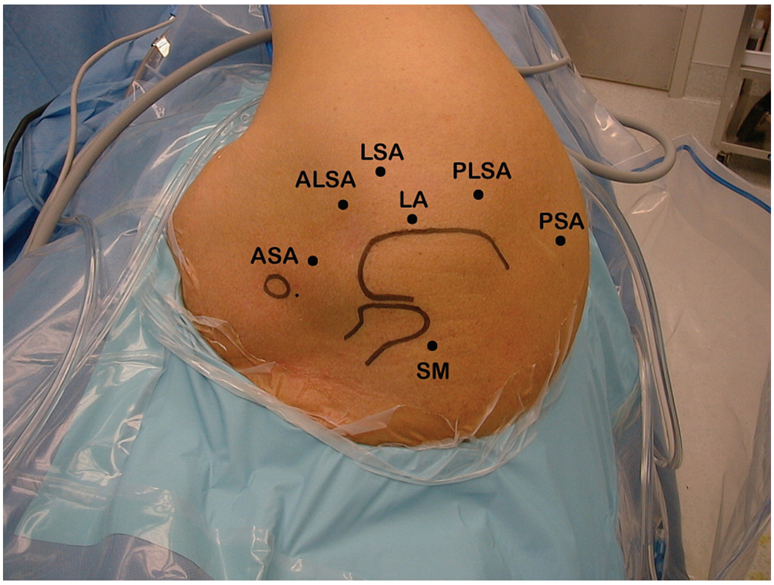
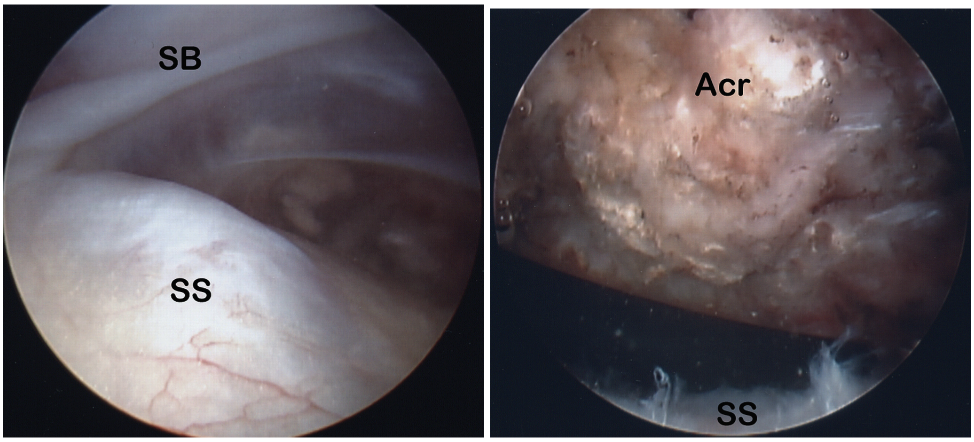


Figure 2.27. Right shoulder in the lateral decubitus orientation viewed from superior (anterior is to the left); bony landmarks are mapped out and the common glenohumeral portals are identified; PSA, posterior subacromial; PLSA, posterolateral subacromial; LSA, lateral subacromial; LA, lateral acromial; ALSA, anterolateral subacromial; ASA, anterior subacromial; SM, Neviaser portal.

LSA provides a "50-yd line" view of the supraspinatus–infraspinatus insertion onto the greater tuberosity and a lateral view of the acromioclavicular joint, the anterior acromion, and the posterior bursal curtain. Instruments are able to approach the rotator cuff, greater tuberosity, and acromion. The entry site is 2.5 to 3.0 cm lateral and 0 to 1.0 cm posterior to the AL corner of the acromion. Instruments roughly parallel the inferior surface of the acromion.

Anterolateral subacromial (ALSA) portal is the same as AL glenohumeral portal, but is placed into the subacromial space. When in the anterior subacromial (ASA) space, it provides a view of the extra-articular biceps, the inter-tubercular groove, the bursal surface of the subscapularis, and the lesser tuberosity (once the clavipectoral fascia has been excised). Instruments can approach the subscapularis tendon for release and suture passage as well as to perform a coracoplasty. The entry site is 1.0 cm anterior and 1.0 to 1.5 cm lateral to the AL corner of the acromion.



A B.

Figure 2.28. A: Scope is in the PSA portal viewing anteriorly (camera is rotated to place the acromion horizontal in all subacromial photos); normal subacromial bursal region; SS, normal supraspinatus with vascular pattern; SB, ASA bursal fold. **B:** Scope is in the PSA portal viewing anteriorly; SS, bursal surface of supraspinatus; Acr, large anterior acromial spur.

ASA is the same as the anterior glenohumeral portal, but enters the subacromial space. It offers a view of most of the subacromial space, but is commonly used for suture management. Instruments can be introduced into the anterior aspect of the rotator cuff for a side-to-side repair. Once through the skin, the trocar is directed immediately beneath the anterior margin of the acromion. When instrument access to the biceps groove is intended, the optimal portal entry site is identified with a spinal needle. While viewing from the AL portal, the needle is directed toward the biceps groove with the humerus internally rotated approximately 20°.

Posterolateral subacromial (PLSA) serves as a primary viewing portal to address rotator cuff pathology. Once established, a 30° scope offers a "50-yd line" view of the rotator cuff and subacromial space (Figure 2.29). The entry site is approximately 1.0 cm anterior and 1.0 cm lateral to the PL corner of the acromion. An arthroscope in the PL portal may interfere with instruments introduced through the LSA portal if a minimum of 3 cm is not maintained between the two sites.

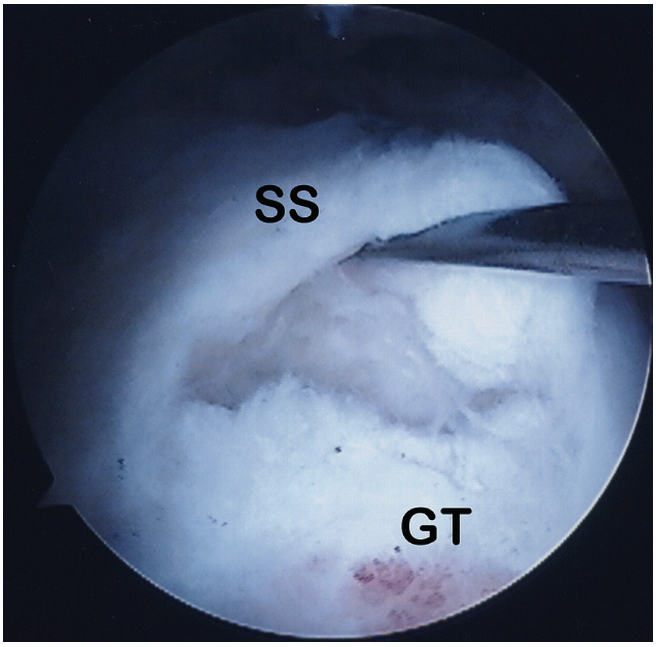


Figure 2.29 Scope is in the PLSA portal viewing anteromedially; probe demonstrates a bursal-sided rotator cuff tear; SS, supraspinatus; GT, greater tuberosity.

Lateral acromial (LA) is primarily used for instrument approach to the greater tuberosity (e.g., drill, tap, and anchor insertion for rotator cuff repair). The entry site is immediately lateral to the lateral border of the acromion. The optimal anteroposterior location is identified using a spinal needle. Access to the entire greater tuberosity is possible with internal and external rotation of the humerus. When attempting to place anchors into the medial aspect of the greater tuberosity adjacent to the articular cartilage, it is essential to nearly completely adduct the humerus to avoid approaching the tuberosity at too shallow an angle and potentially violating the articular surface of the humeral head.

Superomedial (SM—Neviaser) is employed to introduce suture-passing and retrieving instruments toward the rotator cuff. The entry site is 1.0 cm medial to the posterior aspect of the acromioclavicular joint. While the arthroscope in the subacromial space and the arm abducted <45°, a spinal needle is directed from medial to lateral at approximately 60°in the frontal plane. If the portal is introduced too close to the acromioclavicular joint, the mobility of the instrument is significantly restricted.

Anterior acromioclavicular (AAC) affords an anterior approach for resection of the distal clavicle. The entry site is 2.0 cm anteroinferior to and in line with the acromioclavicular joint. The optimal path is identified with a spinal needle. Alternatively, when approaching the acromioclavicular joint in direct fashion, two small portals can be established. One is directly AS and a second posterosuperior to the AC joint. A small-diameter arthroscope and shaver are used initially until a greater space can be established.

2.24 Suture Management

Suture management is one of the most challenging aspects of accurately completing an effective arthroscopic shoulder procedure. By employing a systematic routine, suture can be passed, manipulated, and tied in an efficient manner. Simplifying the steps involved results in time saved and frustration avoided. Suture must be handled carefully to avoid fraying and nicking with the possibility of eventual breakage. Loop rather than jaw-type graspers help maintain this suture integrity. It is optimal to isolate the suture being manipulated whenever possible by placing all other nonworking sutures in a separate portal. Tangling and mistaking various limbs and suture mates can thus be avoided. Once all sutures for a given anchor have been passed, the working cannula is withdrawn and then reinserted placing the sutures outside the cannula, which can then be used to manage a new set of sutures.

In order for sutures to securely re-approximate tissue, they must be optimally placed. When manipulating tissues and suture-passing instruments, efficiency can be gained by having an assistant hold the arthroscope to maintain an acceptable field of view. The surgeon is then able to secure the tissue with graspers in one hand while controlling the suture-passing device with the other, similar to using forceps and a needle driver in an open technique. Antegrade devices, which often simplify suture passage by minimizing the number of steps involved, can be made more efficient by using a counterforce traction suture to control the tissue and prevent it from being pushed away during instrument delivery. When using a penetrating device in retrograde fashion, its mobility can be restricted significantly once it has passed through the tissue. Rather than attempt to “chase” the desired suture with the open jaw, deliver the selected anchor suture to the penetrator with a loop grasper or knot pusher. Various cannulated instruments, with or without an attached suture retrieval loop, do not require the use of a cannula and are able to be introduced through a very small skin nick such as the SM (Neviaser) portal.

Managing sutures in an orderly fashion avoid entanglements. When passing sutures through the rotator cuff, it is helpful to pass them from “far to near.” Those sutures that are to be passed furthest from the viewing arthroscope are introduced first (Figure 2.210). Consequently, as subsequent sutures are delivered closer to the arthroscope, the field of view remains unobstructed by previously passed sutures. The suture pairs are then progressively tied in the opposite sequence, that is, those closest to the scope tied first and those furthest tied last. This method permits adherence to the principle of working with sutures in isolation as much as possible.

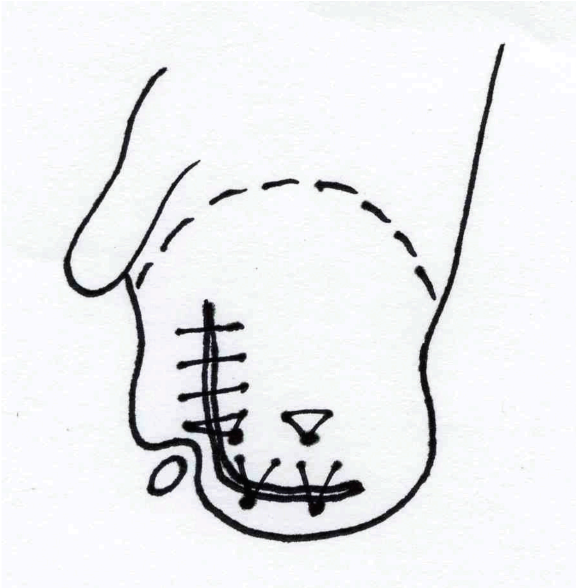


Figure 2.210. Diagram of a right shoulder viewed from superiorly depicting a large "L" shaped rotator cuff tear; consider placing most medial (farthest from scope) sutures first and working progressively laterally (closest to scope); consider tying the most lateral sutures first and then progressing medial with suture pairs.

When working with sutures that pass through anchors, care must be taken to avoid "offloading" the suture from the anchor. The location of the involved anchor should be kept in view while a limb is being retrieved to verify that the suture is being pulled through the anchor. Stop, reorient yourself, and select the proper limb if the suture is moving through the anchor.

Many methods exist for tying knots. When using a sliding knot, the post limb must pass through the tissue being repaired so that the knot is delivered toward the tissue and away from the anchor. Otherwise, as the knot is introduced, it can become "bound up" at the entry site for the anchor and fail to slide further, compromising loop security. In addition, prominent knots near articular surfaces may generate significant chondral scuffing and abrasion. As half hitches are introduced to secure the knot, the post should be alternated, the throw reversed, and each half hitch seen to “lay down” without inappropriate twists.

2.25 Tips / Tricks / Pearls

Accurate portal placement is essential. If the initial portal placement is malpositioned or misdirected, time, frustration, and potential complications can be avoided by establishing a new portal in the optimal location. Using sharp trocars or excessive force to penetrate the capsule can lead to inadvertent damage to the articular cartilage. Once established, screw-in or lock-in cannulas are more secure, particularly when instruments are passed through them frequently. A relatively tight portal of entry through the skin will also help prevent inadvertent withdrawal of the cannula. Clear cannulas improve the visibility of instruments and sutures that are within the tip of the cannula.

It is essential to obtain a clear field of view. Relative hypotensive anesthesia, a hydrodynamic balance of inflow and outflow pressures, irrigation containing epinephrine, and selective radiofrequency cauterization can lead to improved visibility. Repositioning of the joint often improves the view, especially in relatively tight regions (e.g., posterior displacement of the humeral head to improve access to the anterior glenoid or improve working space when addressing subscapularis lesions; adducting the shoulder to safely approach the medial aspect of the lesser tuberosity). Anatomic relationships should be verified prior to resecting or altering any tissue. Motorized instruments and sharp tools must be kept in view to prevent iatrogenic tissue damage.

If a suture is inadvertently offloaded from an anchor with a suture loop eyelet, a new suture can be reintroduced into the anchor (Figure 2.211A–E). Reposition the suture remaining in the anchor to create asymmetric limbs. Load a free suture in a small atraumatic needle and then pass that needle through the braids of the longer limb of the remaining anchor suture where it exits the working cannula. By placing traction on the short limb of the remaining anchor suture, the free suture can be “shuttled” through the anchor eyelet. Once the two sutures are disengaged, both pass through the suture eyelet.

When a suture is accidently offloaded from an anchor with a rigid eyelet and multiple sutures, a new suture can be secured to the anchor (Figure 2.212A–C). A simple overhand throw is created outside the cannula with the suture, which still passes through the anchor. A second free suture is passed beneath the loop that has been created. As the half hitch is delivered down the cannula, it draws the second (free) suture to the anchor head. Second and third alternating half hitches are introduced to secure the free strand. Once all limbs are passed through the tissue, the suture passing through the anchor eyelet is tied first, which helps further secure the free strand. Non-sliding knots must be used for both pairs of sutures.

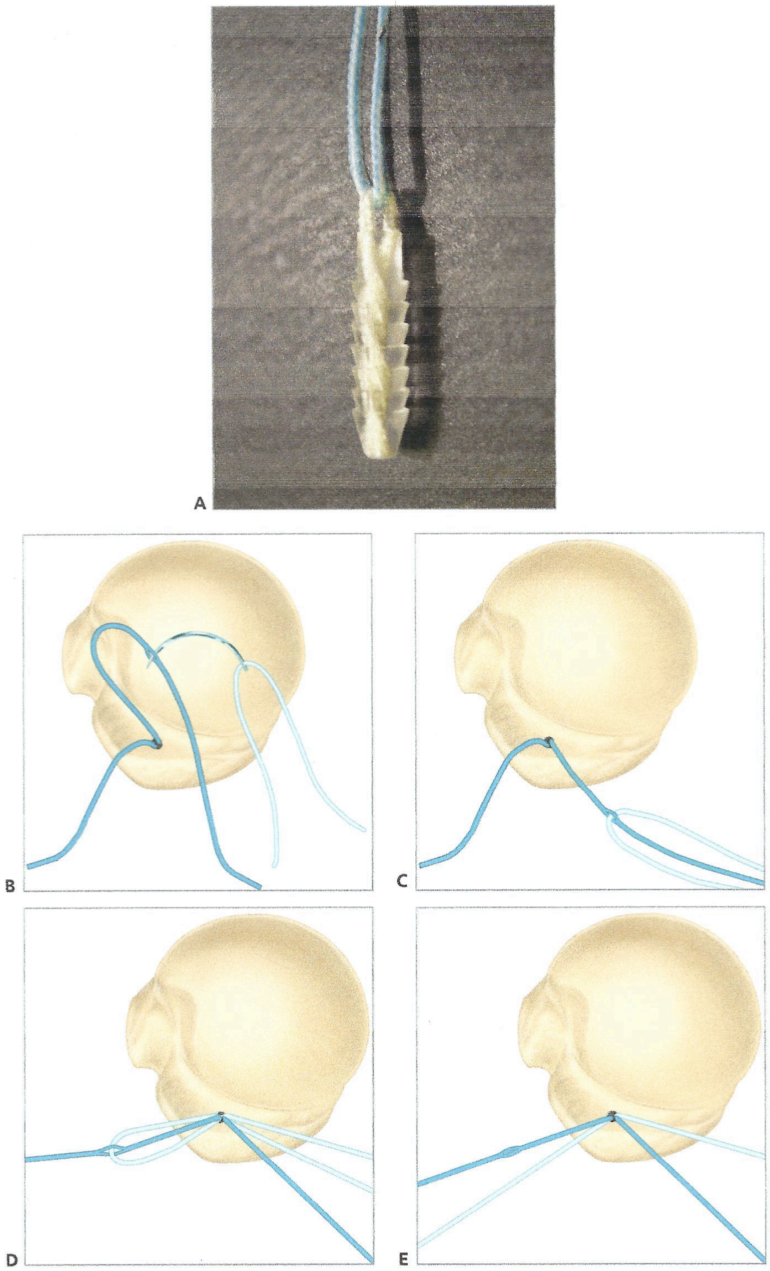


Figure 2.211. A: An anchor with a suture eyelet. **B:** An atraumatic needle delivering a free suture through the braids of the suture, which remains in the anchor. **C:** A

completed pass of the free suture through the anchor suture. **D:** By pulling on the short limb of the anchor suture, it acts as a shuttle to deliver the free limb through the suture eyelet. **E:** Both suture limbs are now through the anchor eyelet in a normal fashion.

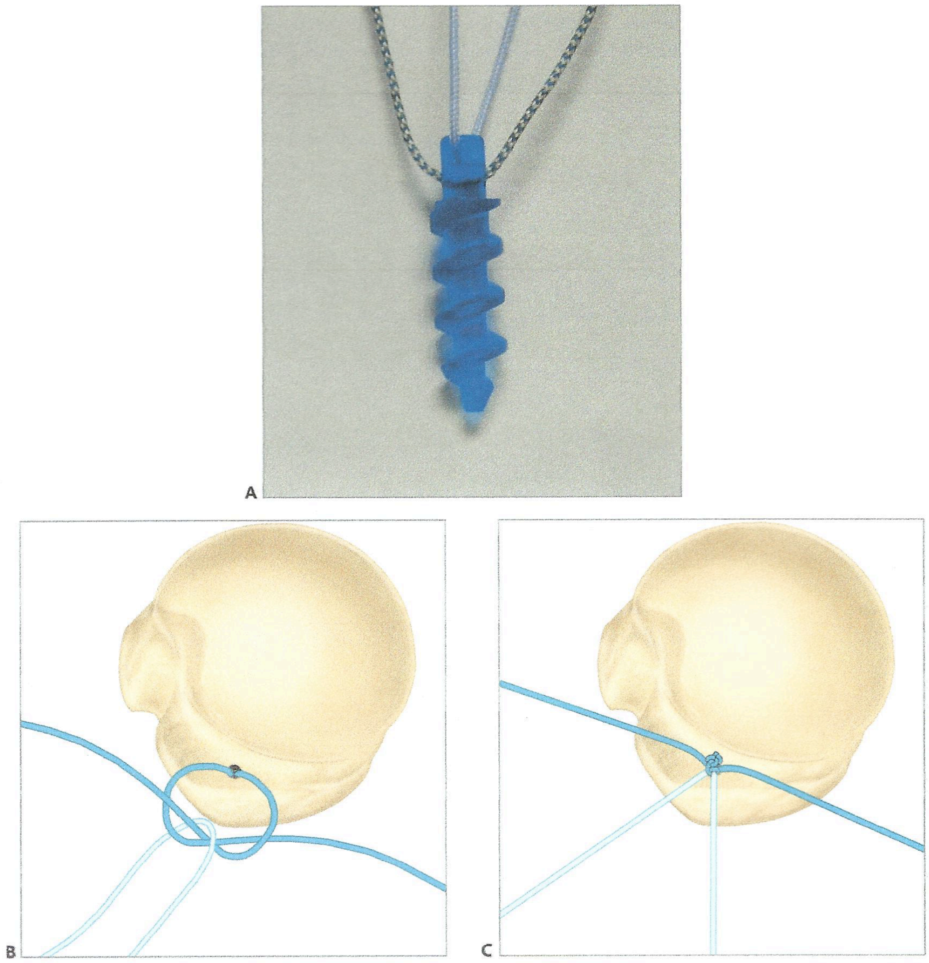


Figure 2.212. A: An anchor with a rigid suture eyelet. **B:** The suture remaining through the eyelet creates an overhand throw and a separate free suture is passed through the loop that is created. **C:** The anchor suture is tied to the anchor head and backed up with two half hitches. Both sutures are now secured to the anchor, but require non-sliding knots to be employed once the sutures are passed through the tissue.

2.26 Conclusions

Either the beach chair or lateral decubitus positions can be used to position patients safely for shoulder arthroscopy. Adequate cerebral blood flow must be maintained when the head and thorax are significantly elevated. General anesthesia is routinely performed and permits greater blood pressure management compared with an ISB. Experienced anesthesiologists with a detailed knowledge of the regional anatomy and an opportunity to perform blocks on a routine basis to maintain their skills should perform ISBs. The use of ultrasound guidance is recommended. Accurate portal placement can either greatly facilitate or hinder the performance of any arthroscopic procedure. An 18-G spinal needle will aid in identifying the optimal entry site and path for specific portals. The choice of camera and view orientation is largely surgeon preference, particularly when working in the subacromial space. Manipulating the position and displacement of the shoulder will aid in optimizing the view and working space. A systematic routine for handling sutures will prevent tangling, suture damage, and insecure knots with poor loop security.

2.3 Evolution of the Arthroscopic Bankart Repair:
"Controversies in Arthroscopic Shoulder Surgery: Arthroscopic Versus Open Bankart Repair, Thermal Treatment of Capsular Tissue, Acromioplasties—Are They Necessary?" Angelo R. Arthroscopy: The Journal of Arthroscopic and Related Surgery, Vol. 19, No 10(October, Suppl. 2), 2003:pp 224-228. Appendix 3
(Candidate is the sole author of this segment of the publication)

A significant debate has involved in the last several years regarding the optimal surgical management of glenohumeral instability. The majority of patients with traumatic anterior-inferior instability sustain a capsulolabral detachment [82, 135] described by Perthes [136] and Bankart.[81] The open Bankart repair described by Rowe [83] has resulted in a high rate of shoulder stability, although the functional outcome has sometimes been suboptimal, especially for higher-caliber athletes. Although suture anchors have simplified the procedure when compared with bone tunnels, [137] the technique has not otherwise changed substantially. In an effort to decrease the morbidity and improve the functional results, Caspari [138] and McIntyre [139] reported on an arthroscopic capsular shift technique, which began a period of intense interest and study in arthroscopic techniques to stabilize the shoulder. As refinements have been made to the arthroscopic methods, the number of proponents has grown. Current arthroscopic techniques using suture anchors [87] mimic the open Bankart repair and comparable results have been reported.[88, 140-142] Rather than attempt to defend a particular position as to whether an arthroscopic or an open approach to shoulder stabilization is best, the focus here is placed on what we have learned regarding the appropriate indications for each of these techniques.

2.31 Arthroscopic Bankart Advantages

A number of advantages have been identified for the arthroscopic Bankart repair, including decreased morbidity. Based on visual analog scales, the pain is typically less than after an open procedure with a decreased need for narcotic pain medications. The arthroscopic approach requires less disruption of normal anatomy, specifically violation of the subscapularis tendon. During revision of failed open Bankart repairs, the subscapularis tendon, if previously divided and repaired, is often quite thin and atrophic. In a small number of cases, a complete failure of the subscapularis repair occurs. Typically, there is less anterior scarring after an arthroscopic repair. It is also possible, although unproven, that there could be a lower incidence of degenerative joint disease of the shoulder after an arthroscopic repair as a result of a greater preservation of normal range of motion.[143, 144]

A clear advantage for the arthroscopic technique is the ability to more thoroughly diagnose the extent of associated pathology within the entire glenohumeral joint, including loose bodies, partial articular surface rotator cuff tears, SLAP tears, biceps tendinopathy, and chondral defects. Provided that the joint remains stable, shoulder function is often more normal after an arthroscopic repair. Throwing athletes are more likely to return to the same or higher level of function after an arthroscopic repair.[145] In addition, there is typically less discomfort with overhead activities, including both work and sport pursuits. If an open Bankart repair requires an overnight hospital stay, then an arthroscopic procedure is likely to be less costly. Depending on the experience and skill level of the arthroscopic surgeon, the total operative time may be less than that for an open procedure.

2.32 Open Bankart Advantages

Open Bankart repairs have been successfully performed for many years and have resulted in a fairly low recurrence rate.[83] Studies of open Bankart repairs, however, often do not include the incidence of subluxations and/or the presence of apprehension, but include only frank dislocations in the reported “failure rate.” After an open dissection, it is possible that the subscapularis and capsular repair could add a “buttress” of additional scar tissue that aids in the prevention of recurrent anterior instability. This is particularly true if the subscapularis and capsule have been separated during the exposure and “laminate” together during the healing process. In addition, the open Bankart repair is a technique familiar to most orthopedic surgeons, requires relatively little special equipment, and results in a reasonably reproducible recurrence rate. It also provides flexibility in addressing large glenoid bone defects with either a coracoid transfer or iliac crests bone graft.

2.33 Considerations in Selecting an Arthroscopic Versus Open Repair

Careful patient selection is an essential component in obtaining an optimal result after any shoulder stabilization procedure. There are a number of key issues that must be considered in choosing whether to perform an arthroscopic or open anterior Bankart repair. It is important to identify the individual patient goals. For one patient with non-dominant shoulder instability, a stable shoulder and modest loss of range of motion would be a successful result. For an overhead athlete, a stable shoulder with modest loss of range of motion in their dominant arm would likely preclude a return to their former level of competition and result in dissatisfaction with the outcome. Throwing athletes tend to have a somewhat better chance at returning to effective participation after an arthroscopic stabilization. On the other hand, the patient routinely involved in contact and collision sports or very heavy lifting and laboring activities could fare better with an open stabilization. Although some authors do not consider contact activities a contraindication to an arthroscopic stabilization,[146] the majority tend to prefer an open stabilization for these patients. Given the relatively high likelihood of recurrence, select high-demand athletes can be considered for arthroscopic stabilization after a first-time dislocation.[147, 148]

The experience of the individual surgeon is also an important consideration. We must each assess our own skill level and ability to arthroscopically treat the entire gamut of glenohumeral pathology contributing to a patient's instability. Usually, experience is gained and skill developed in a "stepwise" fashion. Routinely, different segments of a procedure are mastered until the final surgical procedure is reliable and will withstand careful testing at the completion of the repair. If the arthroscopic procedure is not progressing in a technically sound manner or the time spent becomes excessive, conversion to an open stabilization is prudent. The patient's history will also help in determining the appropriate approach. For those patients less than 20 years old, there tends to be an increased rate of recurrent instability with both open and arthroscopic approaches. An open stabilization in this age group is preferred unless the Bankart lesion is traumatic, the capsulolabral tissue robust, and pathologic tissue laxity is absent. The number of instability episodes that the patient reports is likely to loosely parallel the magnitude of the capsular strain and attenuation. Unless the surgeon is confident that the capsular laxity can be effectively addressed arthroscopically, an open procedure should be considered. In general, the specific age and activity level of the patient is somewhat less important than the specific tissue pathology observed, i.e., robust capsular tissue with an intact labral ring versus a patulous or pathologically lax capsule with labral obliteration.

On physical examination, indicators of tissue hyper-elasticity include elbow hyperextension, excessive patellar laxity, and thumb hyper-abduction as well as broadened scars from previous wounds. An open capsular shift procedure is more likely to result in a successful outcome for these individuals. The most accurate assessment of the magnitude of glenohumeral laxity is often obtained during the examination under anesthesia. Translation in the anterior-inferior and posterior directions is determined with the "load and shift" test and inferior laxity determined by the magnitude of the sulcus sign. If posterior laxity is determined to be greater than normal, i.e., the humeral head rides over the posterior glenoid rim, then an arthroscopic posterior capsular augmentation should be considered. If traction is applied to the adducted arm while in 20° of external rotation and the sulcus sign exceeds approximately 0.5 cm, rotator interval closure or plication should be considered.

During the diagnostic portion of the arthroscopy, the character and extent of soft tissue pathology must be carefully evaluated. A well-defined Bankart lesion is likely to be present after a traumatic anterior inferior dislocation. Alternatively, the capsulolabral tissue can also be significantly scarred back to the glenoid rim and the defect not apparent without careful probing. The only indication that an anterior labral periosteal sleeve avulsion (ALPSA) lesion is present could be that the capsulolabral tissue is attached more medial than normal along the neck of the glenoid. The integrity of the mid-capsule and inferior glenohumeral ligament should also be evaluated. If they appear markedly stretched, attenuated, or incompetent, consideration should be given for an open repair. Intra-capsular tearing must also be sought and may coexist with a labral detachment. In this situation, simply repairing the labrum directly to the glenoid rim will not restore competency of the anterior capsulolabral tissues. A tear of the humeral attachment of the glenohumeral ligaments (HAGL) must always be ruled out during the diagnostic arthroscopy. Although arthroscopic techniques have been described to address this lesion, an open repair is more likely to be reliable.

Bone deficiencies have a significant impact on the selection of the technique. Imaging studies routinely include an anterior-posterior, West Point lateral, and Stryker notch view. If the lateral views suggest a significant anterior glenoid defect, a CAT scan assists in operative planning by quantifying the magnitude of the defect. During arthroscopy, the integrity of the bony glenoid rim must be noted. A fracture that comprises less than approximately 20% is not a contraindication to arthroscopic stabilization. In most instances, the bony fragment should be retained and incorporated in the repair. A significantly higher failure rate after an arthroscopic repair has been reported by Burkhart and DeBeer[11] if >20% to 25% of the glenoid rim is deficient (inverted pear configuration). A coracoid transfer or iliac crest bone graft should be considered in an effort to recreate relatively normal glenoid geometry. If a Hill-Sachs lesion comprises >30% of the articular surface of the humeral head, or engages on the anterior rim of the glenoid with the arm in a position of abduction and external rotation, an open stabilization is recommended. Large humeral head defects may necessitate consideration for an osteochondral graft.

2.34 Technique of an Arthroscopic Bankart Repair

Patient Orientation

Surgeon preference dictates whether to perform the procedure with the patient in the lateral decubitus or beach-chair position. For the lateral decubitus orientation, the involved forearm and hand should be "suspended" distally and the upper arm laterally with accessory support. Distal "traction" should be avoided because it could hinder re-tensioning of the glenohumeral ligaments during repair. A standard posterior arthroscopy viewing portal 1.5 cm inferior and 1 cm medial to the posterolateral corner of the acromion is established with a disposable cannula first (this portal will later be used for instrumentation). While viewing the intra-articular space from this portal, a spinal needle establishes the site for the anterosuperior portal. The entry point lies 1 cm lateral and slightly anterior to the anterolateral corner of the acromion. While viewing from posterior with the arthroscope in the posterior cannula, a switching stick is introduced through a stab incision and penetrates the capsule immediately anterior to the supraspinatus. Again, while viewing from the posterior portal, the mid-anterior portal site is located with a spinal needle 1.5 cm lateral and 1.5 cm inferior to the coracoid tip. The portal enters immediately superior to the subscapularis tendon. This portal should permit instruments to reach the most inferior aspect of the glenoid without an excessively acute angle to the inferior glenoid. In addition, an angle of approximately 45° to the transverse plane is desired to provide proper approach to the glenoid rim. If this portal is "too shallow," the drill bit or anchor can skive across the glenoid surface or enter the glenoid neck too medially. The anterosuperior portal provides an excellent view of the anterior glenoid rim and the most inferior aspect of the Bankart lesion. Some surgeons prefer the view from a posterior portal while performing the repair. In that instance, a 70^0 arthroscope is helpful to optimize the view.

2.35 Glenoid Preparation

Marginal articular cartilage debris is removed with a motorized shaver introduced through the mid-anterior working portal. It is important when working around the capsule or labrum with a motorized instrument that the suction is turned off to avoid inadvertent capture and damage to those tissues. A liberator-elevator or similar tool is then used to release the capsulolabral tissue from its scarred position and thus expose the lateral 1 cm of the glenoid neck. Mobilization is generally inadequate until the subscapularis muscle fibers are seen. It is important that the capsular tissue be released around to the 6-o'clock position on the glenoid. If the release is inadequate, the tissues can be tethered medially and inferiorly, preventing adequate superior and lateral re-tensioning of the tissues. With the suction turned off, a motorized shaver or hooded burr is used to lightly excoriate the glenoid neck surface to prepare for optimal healing.

2.36 Anchor / Suture Delivery

Drill holes for the appropriate anchor are placed 2 to 3 mm onto the glenoid surface at the 5:00, 3:30, and 2:00 positions. The anchor of choice is delivered into the inferior hole and tested for security by placing traction on the sutures. When the anchor is inserted, the eyelet should be oriented to permit the sutures to exit anteroinferior and posterosuperior to allow the repair suture to remain collinear with the eyelet and slide easily during knot tying. Using a loop grasper or crochet hook, the anteroinferior suture is retrieved out the posterior cannula. A cannulated suture hook or similar suture-passing device is then introduced through the midanterior portal. A pass is made through the capsule beginning 1 cm inferior and 1 cm lateral to the capsular rim. The instrument is then brought up under the labrum adjacent to the rim of the glenoid. Superior tension is applied to the instrument - if the exiting tip of the suture delivery device is superior to the anchor drill hole, the tissues will not be adequately re-tensioned and a lax inferior pouch will result. The instrument should be removed and another pass made through the capsule more inferiorly. A No. 0 or 1 polydioxanone suture (PDS) or other suture shuttle is then delivered through the tissues and retrieved out the posterior cannula. Outside the posterior cannula, the PDS is tied near the end of the previously retrieved anteroinferior suture with a simple overhand loop. The PDS shuttles the permanent suture limb through the capsulolabral tissue from posterior to anterior. As this newly passed limb of suture is tensioned, the rim of tissue should be observed to ride up onto the glenoid rim and tighten the inferior glenohumeral ligament. Verify that significant distal traction on the humeral head is not present before knot tying. If tying a sliding not, the limb that passes through the tissue must be the post to deliver the knot away from the articular surface. As the knot is thrown, delivered, and secured with alternating half hitches, the "pursed" tissue creates a pseudolabrum. These steps are repeated for the second and third anchors. Occasionally, four anchors are necessary and are evenly spaced along the glenoid rim from 1:30 to 5:30.

2.37 Augmentation

In the majority of cases exhibiting a Bankart lesion, posterior capsular augmentation is unnecessary. If, however, excessive posterior capsular laxity was determined during the examination under anesthesia and the appearance at the diagnostic arthroscopy, it should be addressed with either suture plication or thermal capsulorrhaphy. Suture plication is simple, low risk, and affords the opportunity to be somewhat quantitative in the amount of capsular tightening that is created.[149-151] While viewing from the anterosuperior portal, a rasp or whisker blade shaver is introduced through the posterior portal. The posterior and inferior capsule is slightly excoriated for 2 cm adjacent to the posterior glenoid labrum to stimulate healing. A cannulated suture hook is then used to create a small pleat in the capsular tissue. The suture hook tip enters the capsule 1 to 1.2 cm lateral to the intact labral rim and exits approximately 5 or 6 mm lateral to the rim. The tip of the instrument then enters and passes beneath the intact labrum so as to create a small pleat of tissue. The PDS suture is delivered and using the most lateral limb of the suture as a post, a sliding knot is securely tied and backed up with half hitches. Two or 3 additional sutures can be placed to complete the augmentation. Care must be exercised so that an increasingly larger pleat of tissue is result.

An alternative is to perform a thermal capsulorrhaphy. The amount of shrinkage and its effect on the mechanical properties of the capsule are both time and temperature dependent. The thermal tip should slightly indent the capsule and must be constantly moving. It is advisable to create a "striped" or grid pattern with the wand.[152] It is advisable to leave as much normal tissue as that which is thermally shrunk. The magnitude of visible capsular shrinkage is quite variable and should not be used as the sole end point for thermal treatment, otherwise overheating of the tissues can occur.

A rotator interval closure may be considered depending on the magnitude of the sulcus sign and the posterior inferior laxity determined on examination. Two methods have been described. An outside-in approach involves passing the sutures and tying them in the subacromial bursa, which involves imbricating the coracohumeral ligament. To begin, a spinal needle is introduced through the anterosuperior skin incision, transgresses the subacromial space, and penetrates the superior glenohumeral ligament as it enters the joint. A sharp-tipped suture retriever is then passed through the mid-anterior cannula (which has been withdrawn outside the capsule), and passes through the middle glenohumeral ligament to retrieve the PDS suture limb within the joint. The limbs are retrieved and the knot tied in the subacromial space. If this method is chosen, the arm should be placed in external rotation during knot tying to avoid over-tightening the shoulder. A second technique involves closing the middle to the superior glenohumeral ligaments. A simple method to accomplish this interval closure involves introducing a sharp-tipped suture grasper/retriever through the mid-anterior portal while viewing from the posterior portal. The instrument is used to grasp the PDS suture 3 cm from the end and deliver it through the middle glenohumeral ligament from outside in. The limb is then retrieved out through the anterosuperior cannula. The end of the remaining limb, which exits the mid-anterior cannula, is then delivered from outside in through the superior glenohumeral ligament and retrieved out the anterosuperior cannula. Arthroscopic knots can then be delivered through the anterosuperior portal to affect closure of the rotator interval. Additional sutures can be placed as needed.

2.38 Test the Repair

It is important to verify that the repair is secure. While viewing arthroscopically, the entire repair is carefully palpated. In addition, the arm should be removed from any traction or suspension and tested in the jeopardy position of abduction and external rotation. If the repair is inadequate, further measures must be taken to address the residual laxity.

2.4 Arthroscopic Bankart Repair – Principles and Technique

"Arthroscopic Bankart Repair for Unidirectional Shoulder Instability" Angelo R. Instr. Course Lect. 2009;58:305-313. Appendix 4

(Candidate is the sole author of this publication)

Abstract

A successful arthroscopic Bankart repair for unidirectional shoulder instability requires careful patient selection and, to the extent possible, the restoration of normal anatomy. The patient's goals and anticipated demands are important considerations. A patient who participates in an overhead sport requires not only a stable shoulder, but also a full range of shoulder motion. An athlete who engages in a contact or collision sport, however, may tolerate a mild loss of motion provided the shoulder is stable. Compared to an open procedure, an arthroscopic repair provides the opportunity to retain the most normal postoperative range of motion and function. Other considerations include patient age, which often relates to overall tissue laxity, and the number of previous instability episodes, which correlates with the severity of pathology (in particular, capsulolabral strain, glenoid chondromalacia and bony deficiency of the glenoid or posterior humeral head). The magnitude of bone loss, particularly for the anterior glenoid, may make an arthroscopic repair inadvisable. Accurate portal placement, glenoid preparation, anchor insertion, and suture passage are key components of the arthroscopic technique, but the most important overall goal is the secure restoration of capsulolabral tissue tension. Secondary posteroinferior laxity, partial rotator cuff tears, labral disorders, and articular cartilage pathology may also require treatment.

2.41 Considerations in Decision Making

2.411 Patient Goals

Shoulder instability recurs in 7% to 10% of patients after an open or arthroscopic suture anchor repair.[153-156] The choice of surgical procedure to correct instability is determined in part by the patient's goals, including anticipated shoulder demands. Although an open Bankart repair is a reliable method for eliminating clinical instability, range of motion and overall function may be unacceptably compromised in some patients who engage in high-demand activities.[157] In particular, overhead athletes, including throwers, are unable to tolerate significant restrictions in flexion or external rotation. The patient's range of motion usually returns more rapidly and completely after an arthroscopic, compared with than an open repair,[158] and the patient is more likely to be able to return to competitive throwing. However, overhead athletes have a lower overall rate of functional success (70%) than other athletes (90%) following an arthroscopic repair.[159]

A comparison of 30 open and 30 arthroscopic Bankart repairs found that muscle strength for forward elevation was markedly weaker after open repair for as long as 3 months; the difference, however, was only 5% after 6 months.[158] Muscle strength for external and internal rotation was significantly weaker 6 weeks after open repair but also approached 5% after 6 months. In a biomechanical investigation of an arthroscopic anterior repair, a traumatic dislocation was created in 12 cadaver specimens, and an arthroscopic suture anchor repair of the Bankart lesion was performed. Glenohumeral translation and rotation were found to approach normal pre-dislocation values.[160]

Arthroscopic Bankart repair remains controversial for athletes who participate in contact and collision sports. Several recent studies have concluded that there is no increased risk of recurrent instability for these athletes after an arthroscopic procedure.[154, 159, 161] The overall recurrence rate was 10% in 85 patients who had undergone an anterior arthroscopic suture anchor repair.[159] Two patients had a recurrence in a subset of 18 collision sport athletes (a similar 11% recurrence rate). In a review of contact or collision sport athletes who were younger than 20 years,[161] two of 18 patients (11%) had a recurrent dislocation at a minimum 2-year follow-up but did not require further treatment. In another evaluation of suture anchor repairs, the 9.5% recurrence rate for contact sport athletes (2 of 21 patients) was not significantly different from the 6% overall failure rate.[154]

A review of 48 shoulders in 46 collision sport athletes reached a different conclusion. Sixteen of the shoulders were arthroscopically stabilized (4 using Suretacs and 16 using suture anchors) and 32 underwent open repair. Instability recurred in 25% of the arthroscopic repairs (1 using Suretacs and 3 using suture anchors) compared to 12.5% of the open repairs. The authors concluded that open stabilization is a more reliable method of repairing anterior shoulder instability in contact athletes.[158]

2.412 Patient History

Patient age is believed to affect the probability of failure after an open or arthroscopic anterior stabilization procedure. Patients in their teens generally have greater tissue and collagen elasticity, which may predispose them to a higher likelihood of repair failure. In addition, younger patients are more likely to be attracted to high-risk activities, including so-called "extreme" sports such as snowboard jumping and aggressive mountain biking. Few studies have specifically evaluated pediatric patients with shoulder instability. In a review of 32 arthroscopic Bankart repairs in patients age 11 to 18 years, 16 shoulders were repaired after unsuccessful nonsurgical treatment and 16 were repaired after the initial instability episode.[162] At an average 25-month follow-up, three re-dislocations had occurred in two patients from the first group (18.5%), and two re-dislocations had occurred in two patients (12.5%) from the second group. The small size of the study, however, does not permit a conclusion as to the optimal treatment of pediatric patients with shoulder instability. The choice of an open or arthroscopic procedure should be based on tissue quality and capsulolabral integrity rather than solely the patient's age. If traumatic pathology is identified and the patient's soft tissues are reasonably robust, an arthroscopic suture anchor technique can provide a reliable repair.

The number of previous instability episodes should also be considered because capsular strain, labral tearing, and glenoid erosion tend to increase with each occurrence of instability.[163] The severity of the accrued pathology must be evaluated during the diagnostic arthroscopy and helps determine whether an arthroscopic repair is suitable.

2.413 Physical Examination

The findings of the physical examination must support the clinical impression of unidirectional shoulder instability. Once the dislocated shoulder is reduced, any deformity usually disappears. There may be diffuse tenderness over the anterior capsular tissues and, less frequently, along the posterior glenohumeral joint line. The patient's range of motion is often restricted following an instability event secondary to pain. In those with chronic instability, acquired anterior capsular laxity can result in an increase in external rotation compared to their normal shoulder. Excessive anterior translation typically appears on the load-and-shift test unless involuntary muscular guarding is present. The magnitude of laxity in the posterior and inferior directions can help determine the need for accessory posterior plication and closure of the rotator interval, respectively. Most patients exhibit apprehension when the shoulder is placed in a position of abduction and external rotation. The relocation test is positive if the apprehension sign is minimized or eliminated when the examiner's hand is placed over the anterior aspect of the proximal humerus to prevent anterior subluxation of the humeral head. With the arm in full adduction and 30° of external rotation, a sulcus sign of more than 1 cm suggests that significant multidirectional laxity is present. In patients with true multidirectional instability, however, clinical symptoms must also be present in more than one direction. The findings of hyperelasticity (the ability of the thumb to be passively placed against the forearm, elbow hyperextension, and marked medial/lateral translation of the patellas) suggest that the collagen tissue is pathologic, which is a known risk factor for failure after arthroscopic stabilization.[164]

2.414 Imaging

Routine radiographs should be obtained: the AP view may show a fracture of the inferior glenoid rim; a West Point axillary lateral view is more sensitive for anterior and inferior rim fracture fragments; the Stryker notch view identifies the presence and size of a Hill-Sachs defect of the posterior humeral head. A CT scan, especially with three-dimensional reconstruction and humeral subtraction, is useful for assessing the size of a glenoid rim deficiency or fracture fragment, and a Hill-Sachs lesion of the humerus.

2.415 Diagnostic Arthroscopy

A thorough arthroscopic assessment of the instability pathology is imperative. The extent and nature of an acute Bankart lesion are usually apparent (Figure 2.41). The humeral head must often be displaced posteriorly to detect the inferior extent of capsulolabral detachment from the glenoid. The true capsular margin may be difficult to identify if the labrum has been obliterated. An anterior labroligamentous periosteal sleeve avulsion (ALPSA) may be difficult to detect in the chronically unstable shoulder. The most reliable clue to its presence is that the capsulolabral tissue appears to be attached too far medial along the glenoid neck (3 to 4 mm medial to the rim). Chondral damage may also have occurred in a shoulder with chronic instability (Figure 2.42).

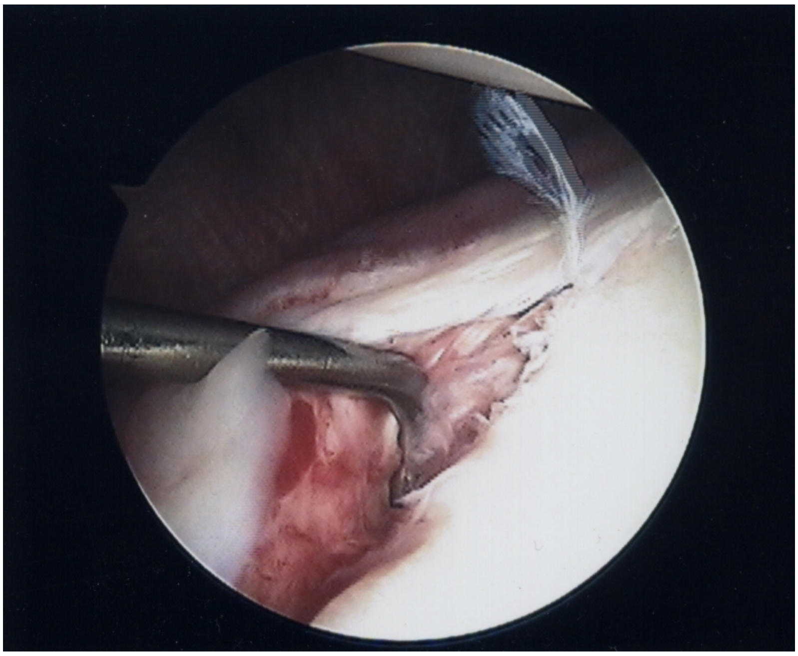


Figure 2.41 Probe entering an acute Bankart lesion (Right shoulder; Scope view fromanterosuperior portal; anterior is left.

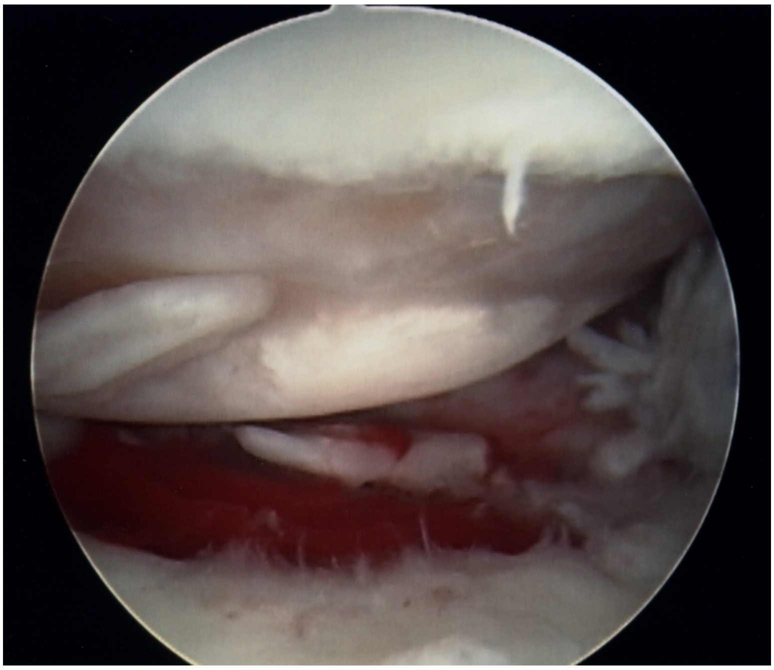


Figure 2.42 Chronic Bankart with chondral lesion of the humeral head and loose bodies (Right shoulder; Scope view from posterior portal; superior is left.

Loss of bone along the anteroinferior glenoid can either result from a fracture or progressive erosion. Thin glenoid rim fracture fragments often remain securely affixed to the capsule and can be detected only by palpation of the tissues with a hook probe. These small wafer fragments can be repaired or simply excised. Evidence suggests that rim fracture fragments larger than 10% of the glenoid diameter should be preserved. In a review of 42 shoulders with posttraumatic recurrent anterior instability, CT was used to estimate the glenoid defect size, which ranged from 11.4% to 38.6%.[165] The bony rim fragment was incorporated during arthroscopic Bankart repair with 39 of the shoulders rated good or excellent on the University of California Los Angeles Shoulder Scale at an average 39-month follow-up. Two re-injuries were reported. In another study, 21 patients with a bony deficiency of the glenoid, including 11 with a traumatic rim fracture and 10 without an identifiable fragment but with attritional bone loss, had a suture anchor arthroscopic Bankart repair.[166] At a mean 34-month follow-up, 2 of the 21 patients had recurrent subluxation, and 1 had a recurrent dislocation. None of the patients with repair of a rim fracture fragment had an episode of postsurgical instability. In a separate study of 65 patients (41 with acute instability and 24 with chronic instability) who had undergone an arthroscopic suture anchor repair of a bony Bankart lesion, two patients (one with acute and one with chronic instability) experienced a re-dislocation at a minimum 4-year follow-up.[167] The average Rowe score of the patients with acute instability improved from 59 to 92 and the score of those with chronic instability improved from 43 to 61. Glenoid rim erosion can increase with recurrent episodes of instability. Arthroscopic Bankart repair has an unacceptably high failure rate if there is significant anteroinferior glenoid bone loss. In a retrospective review of 194 consecutive arthroscopic Bankart repairs using suture anchors,[140] two groups of patients were identified based on whether or not significant glenoid or humeral bone loss was present. Glenoid loss was considered significant if the normal pear-shaped configuration of the glenoid had changed to an “inverted pear” shape, in which bone loss resulted in the inferior one half of the glenoid being narrower than the superior half. A 67% failure rate following arthroscopic Bankart repair was found in patients with an inverted pear glenoid or a significant Hill-Sachs lesion, compared with 4% for patients without a significant bony defect. The size of glenoid defects can be estimated using the central bare spot as a reference. The normal radius of the glenoid (inferior two thirds) is the distance from the central bare spot to the intact posterior glenoid rim. The difference of the normal radius and the distance from the bare spot to the remaining anteroinferior glenoid margin can be used to estimate the percentage defect (for example, a defect of half the length of the radius is approximately 25%).

Hill-Sachs lesions are common after shoulder instability, especially if the episodes are recurrent (Figure 2.43). Most of these lesions are relatively small and can be ignored without compromising the success of the repair. However, a specific subgroup is associated with a higher failure rate after arthroscopic stabilization.[168, 169] In the study cited above regarding glenoid rim deficiency,[140] significant Hill-Sach’s lesions were also defined. Three of 21 failures in that study were deemed to have been caused by an “engaging” Hill-Sachs lesion wherein the posterior humeral defect engaged on the anterior glenoid rim with the arm in a functional position of abduction and external rotation. If the bone loss is greater than 30% to 35% of the articulating surface of the humeral head, an open osteoarticular allograft[168] or arthroscopically-assisted transhumeral impaction grafting[169, 170] may need to be considered.

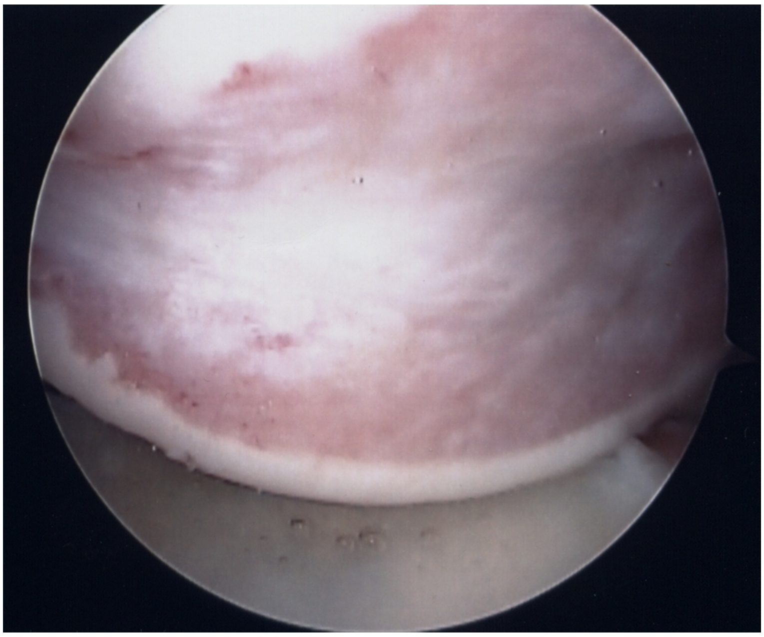


Figure 2.43 Hill-Sachs lesion (Right Shoulder; Scope view from posterior portal; superior is left)

Translation of the humeral head can be difficult to quantify. However, a qualitative estimation of anterior capsular laxity or strain is useful in determining how much to plicate the capsule with each anchor suture. In addition, the magnitude of posterior laxity helps determine whether or not to augment the repair with several posterior "pinch-tuck" capsular plication sutures. A traumatic mid-capsular rent or tear can exist, even in the presence of a distinctly separate Bankart lesion. In a prevalence study, 12 of 303 shoulders undergoing stabilization (4%) had a mid-capsular tear in addition to a Bankart lesion.[167] Eleven of the 12 tears were repaired arthroscopically, with the average Rowe score of those patients improving from 30.4 to 90.4 at 31-month follow-up. In a review of 21 patients with a mid-capsular tear, 7 tears were isolated and 14 were accompanied by a Bankart lesion.[171] More than 90% of the patients had a good or excellent Rowe score after an open or arthroscopic capsular repair along with a Bankart repair when indicated. The average loss of external rotation was 8° for patients with an isolated capsular closure and 16° for those who also had a Bankart procedure.

Humeral avulsion of the glenohumeral ligaments (HAGL) has been recognized as a cause of recurrent shoulder instability. The lesion may not be readily apparent and must be carefully sought during diagnostic arthroscopy by examining the anterior and posterior capsular insertions onto the humeral neck. These avulsions can be repaired arthroscopically, although the procedure is technically demanding.[172] Considerably less morbidity is generated with a posterior arthroscopic repair than an open posterior approach. For an anterior avulsion, however, a standard open deltopectoral approach provides ready access to the anterior neck of the humerus with only partial subscapularis detachment. Any associated lesion (superior or posterior labral lesion, chondral injury, or partial-thickness rotator cuff tear) should be identified and treated.

2.42 Technique for a Suture Anchor Arthroscopic Bankart Repair

A successful arthroscopic Bankart repair requires careful patient selection, a thorough understanding of normal and pathologic anatomy, skill in using arthroscopic tools and implants to approximate normal anatomy, and discernment in guiding the postsurgical rehabilitation program. The technique is not inordinately difficult, but mastery requires study and practice. The necessary skills can be honed in dry model and cadaver laboratories. Thorough planning and the ability to mentally rehearse the procedure are invaluable preparations. The operating room staff must be oriented to the sequence of steps and instruments to minimize miscues and optimize efficiency.

2.421 Optimal Visualization

A clear arthroscopic view of the intra-articular structures is essential and is improved by using mildly hypotensive anesthesia (approximately 100 mm Hg systolic pressure). Epinephrine can be introduced into the inflow solution to help control bleeding (1 cc of 1/1000 epinephrine per 3 L). Although the procedure can be satisfactorily performed using gravity inflow, a pump allows blood pressure spikes to be offset with a temporary increase in inflow pressure. The pump pressure must be carefully monitored to prevent excessive fluid extravasation.

2.422 Patient Positioning

The patient's position must allow access to all areas of pathology. In the lateral decubitus position, the pelvis and thorax are supported by a vacuum pack (bean bag) with a 20° posterior tilt of the chest to facilitate access and orient the glenoid approximately parallel to the floor. A 5- to 10-lb weight attached to the arm sleeve permits suspension rather than traction of the arm. Excessive traction may compromise the ability to adequately retention the soft tissues in a superior direction and can cause undue tension on the brachial plexus during a prolonged procedure. A 5-10# accessory lateral traction pull can be oriented perpendicular to the humerus to aid in separating the humeral head from the glenoid and provide additional working space. Alternatively, an assistant can manually displace the humeral head posteriorly as the need arises. The standard beach chair position may be preferred because of the normal, upright orientation of the shoulder anatomy. However, the posterior aspect of the shoulder is relatively difficult to access arthroscopically with the patient sitting. For a patient in the beach chair position, a higher systolic blood pressure is necessary to maintain adequate cerebral perfusion as the normal compensatory mechanisms for cerebral blood flow may be compromised with the patient in a sitting position under general anesthesia. Conversion to an open procedure is simplest if the patient is in the beach chair orientation but can also be accomplished relatively easily from the lateral decubitus position.

2.423 Portals

Accurate portal placement facilitates identification and treatment of all intra-articular pathology. Poorly placed portals create difficulty in preparing the tissues for repair, placing sutures and anchors, and tying knots. The entry site for the posterior portal is 1.5 cm inferior and 1.0 cm medial to the posterolateral corner of the acromion. The cannula and trocar are directed toward the coracoid tip anteriorly (Figure 2.44). If drills and anchors may be required for the posterior glenoid, the entry site should be adjusted 1.0 cm further lateral to provide an acceptable approach to the narrow posterior rim of the anteverted glenoid.

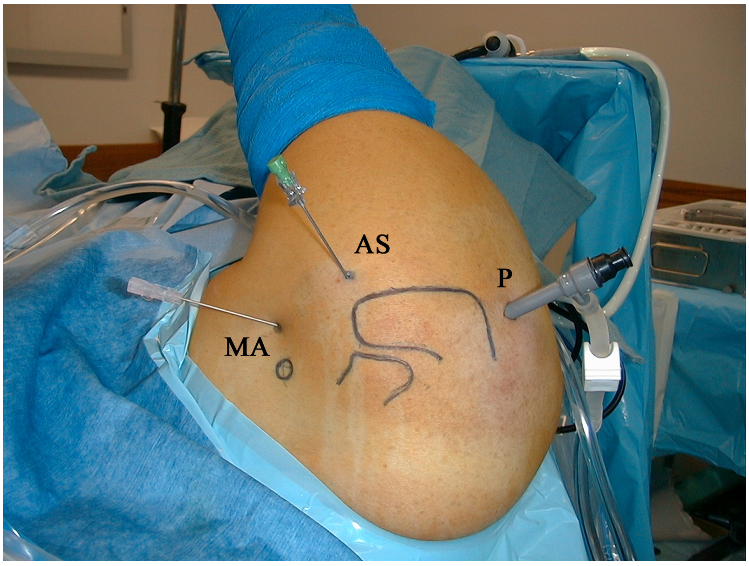


Figure 2.44 Standard arthroscopic portals; P = posterior, AS = anterosuperior, MA = mid-anterior (Right Shoulder; Superior view; Anterior is left, posterior is right)

The mid-anterior portal provides access to the anterior glenoid for instruments. The entry site is located 1.5 cm lateral and 1.5 cm inferior to the coracoid tip. A spinal needle should be used to verify accurate placement; it is initially directed slightly superior, over the superior border of the subscapularis. When the arm is relatively adducted, the subscapularis is relatively lax and can be depressed inferiorly by the incoming needle or cannula and permit ready access to the anteroinferior glenoid. This portal should provide a 30° to 45° approach to the glenoid rim in the transverse plane, which is essential for safe drilling and anchor insertion. An 8.5-mm clear threaded cannula is optimal for this critical working portal.

An anterosuperior portal can be used as a working portal while viewing from the posterior portal. Alternatively, the anterosuperior portal can serve as the primary viewing portal, in which case, only the arthroscopic sheath need be inserted. The optimal entry site is 1.0 cm lateral and slightly anterior to the anterolateral corner of the acromion. The proper path is established using a spinal needle that enters immediately anterior to the supraspinatus tendon and through the rotator interval, either anterior or posterior to the biceps tendon.

2.424 Glenoid Preparation

A full-radius synovial resector is used to debride the articular glenoid margins, removing ragged or unstable articular cartilage (Figure 2.45). To prevent inadvertent damage to the sometimes-fragile adjacent capsulolabral tissue, suction on the shaver should be turned off. The repair must restore normal capsulolabral tissue tension to prevent future glenohumeral instability. Adequate re-tensioning requires mobilization of the capsulolabral tissue, which may have scarred medially along the glenoid neck. A liberator elevator is introduced through the mid-anterior portal and used to free the scarred capsular tissue from bone. When the release is complete, the subscapularis muscle tissue should be visible medial and anterior to the capsule (Figure 2.46). Adequate mobilization allows the capsule to be advanced both superiorly and laterally onto the glenoid rim. An aggressive shaver or a 4.0-mm burr run in reverse lightly excoriates the anterior neck of the glenoid to provide a bed for tissue healing.

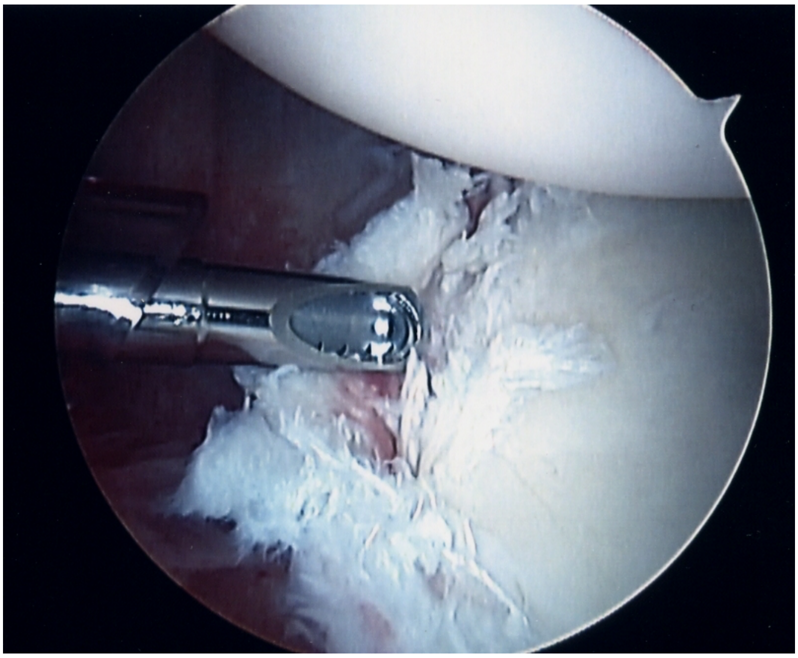


Figure 2.45 Shaver debridement of chondral and labral damage (Right shoulder; Scope view from anterosuperior portal; anterior is left.

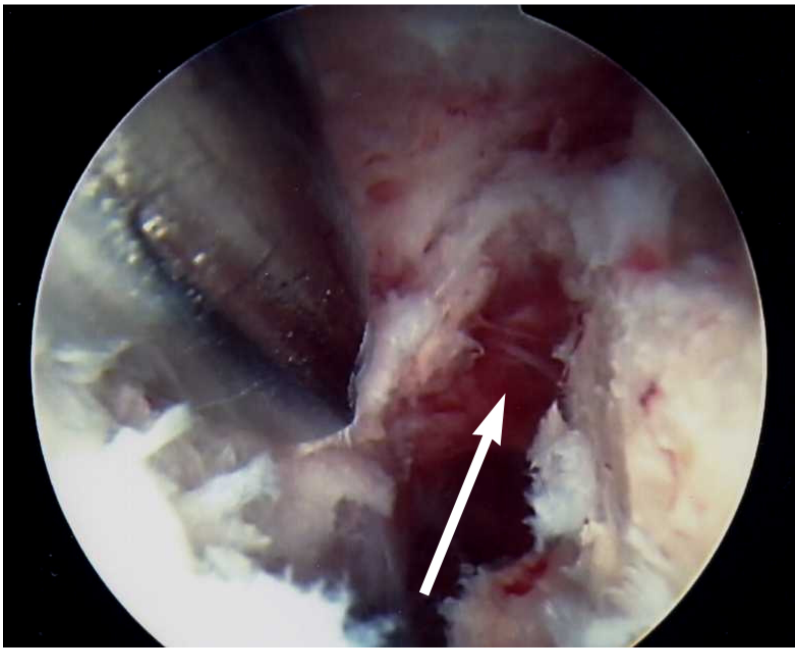


Figure 2.46 White arrow points to exposed subscapularis muscle deep to capsule (Right shoulder; Scope view from anterosuperior portal)

A thin rim fracture fragment (1 to 2 mm wide**)** can be repaired or excised. If the fragment is removed with a burr, the reverse setting should be used to protect the underlying periosteum and thus enhance the integrity of the repair. Reduction and repair should be performed for a fragment larger than 10% of the glenoid diameter. The fragment can be encircled with anchor sutures, or alternatively, sutures can be passed through the fragment by predrilling with a small Kirschner wire. It is often necessary to introduce the Kirschner wire from the posterior portal to safely approach the separate anterior bone segment. Great care must be exercised to avoid inadvertent wire penetration into the soft tissues anterior to the capsule. Over reducing the rim fragment and creating articular incongruity should be avoided.

2.425 Anchor Placement

Anchor holes should be drilled 2 to 3 mm onto the articular surface of the glenoid (Figures 2.47 and 2.48). The anterior wall of the completed drill hole must have sufficient integrity to prevent the strong repair sutures from cutting out anteriorly and medially during the healing period, which will render the repair ineffective. The drill bit should approach the glenoid at approximately 45° in the transverse plane. If the approach angle is too shallow, there is a risk that the bit will skive onto the articular cartilage or that the anchor hole will be located too medial along the glenoid neck. Anchors are evenly spaced between the 5-o'clock and 2-o'clock positions on the glenoid. After an anchor is implanted, its security should be tested by firmly pulling on the suture strands. Higher rates of repair failure have been reported when fewer than three or four anchors are used.[164] Whether made of metal or absorbable material, loose or prominent anchors may cause significant articular cartilage damage. Nonmetallic, non-resorbable anchors made of poly-ether-ether-ketone (PEEK) material have been introduced to avoid the cavitary bone cysts sometimes associated with resorbable anchors.

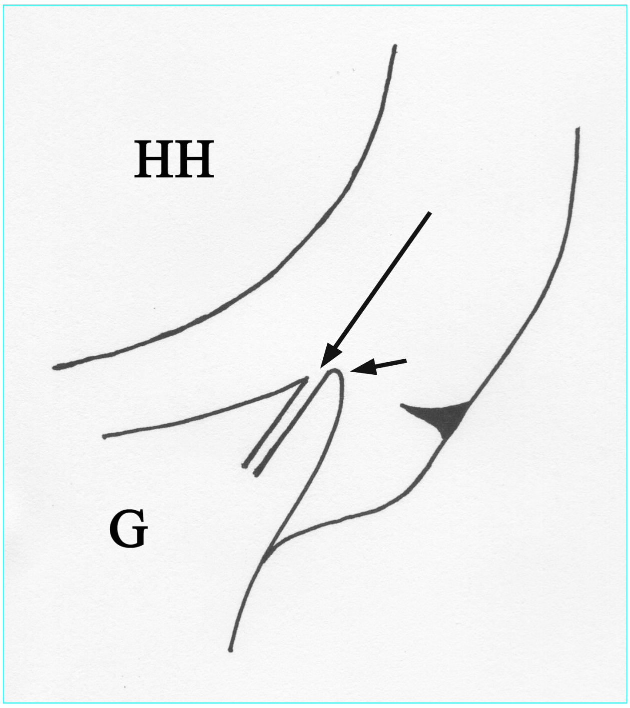


Figure 2.47 Long arrow points to orientation of glenoid drill hole in the transverse plane; short arrow points to intact bone anterior to drill hole (Axillary view; Anterior is to the right)

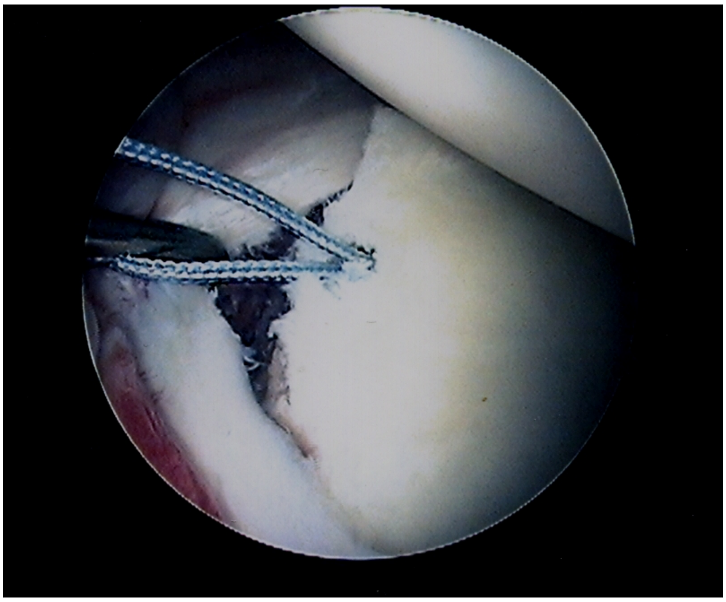


Figure 2.48 Suture anchor located 2 - 3 mm onto face of glenoid with intact anterior bone margin (Right shoulder; Scope view from anterosuperior portal; anterior is left)

Reports on the use of knotless anchors are conflicting. In one study, experienced users documented excellent results at a minimum 2-year follow-up; knotless anchors failed in only 5 of 72 patients (6.9%), all of whom were younger than 22 years.[173] After a similar follow-up period, another retrospective review reported failure of 5 of 21 knotless anchor repairs (23.8%), compared with 3 of 61 repairs using anchors and knot-tying (4.9%).[174]

2.426 Capsulolabral Re-tensioning

A review of 24 patients who underwent revision surgery after an unsuccessful open anterior repair found a persistent or recurrent Bankart lesion in 16, with capsular redundancy in 4.[175] Thus, restoration of normal anatomy including capsular tension during the repair is essential, to the extent possible. We prefer to view from the anterosuperior portal. A serrated drill guide provides a more secure purchase for the guide on the glenoid face than a fish mouth style tip.

After drilling a hole in the appropriate orientation, an anchor is inserted through the mid-anterior portal and the sutures are retrieved out the posterior cannula. A clamp is used to identify the limb that exits the anchor inferiorly so that when this suture is shuttled back through the tissue, a 180° twist of the suture at the anchor eyelet is avoided. Glenohumeral reduction should be maintained during suture placement, and arm traction prevented from causing inferior subluxation of the humeral head. If the capsulolabral tissue is markedly displaced inferiorly, it may be necessary to pass a suture through the capsule and apply superior traction to appropriately tension the capsule while introducing the repair sutures (Figure 2.49).

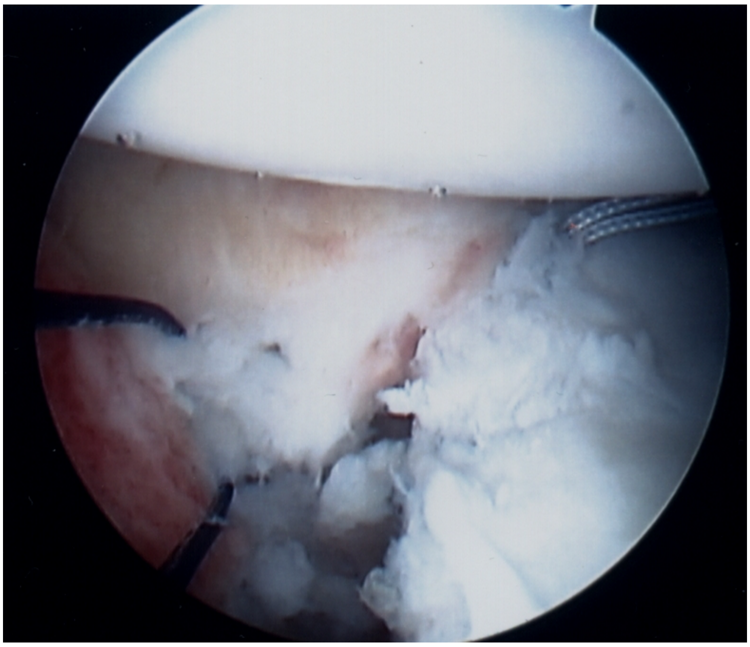


Figure 2.49 Monofilament traction suture re-tensioning capsule superiorly (Right shoulder; Scope view from anterosuperior portal; anterior is left)

The goal is to introduce the inferior limb of the anchor suture approximately 1.0 cm inferior and 1.0 cm lateral to the anchor exit point from the glenoid rim. This placement will permit superior advancement as well as mediolateral plication of the capsule when the suture is tied. A curved, cannulated suture hook is passed down the mid-anterior portal, through the capsule, and then up beneath any remnant of the labrum. Once the hook has been placed through the tissue, it is brought superiorly with moderate tension to check for proper placement. If the hook can be displaced superior to the anchor site, the capsular tissues will not be appropriately tensioned when the suture is tied, and the hook must be replaced further inferior. When the hook is correctly placed, a monofilament shuttle suture is delivered and retrieved out the posterior cannula using a loop grasper (Figure 2.410). Using a simple overhand throw, this posterior limb of monofilament suture is tied around the tail of the inferior anchor suture limb (previously identified with a clamp), which is then shuttled from posterior to anterior through the capsule (Figure 2.411). This limb, which passes through the tissue, becomes the post for a sliding knot that is delivered laterally away from the glenoid as it is secured. Sliding knots are backed up with three or four half hitches. These steps are repeated for each anchor and suture pair (Figure 2.412).

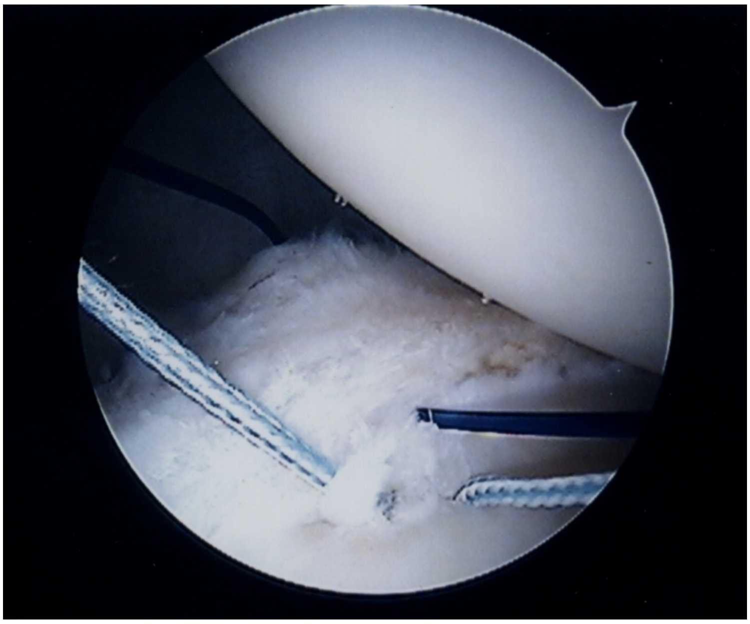


Figure 2.410 Monofilament shuttle suture passed through capsule inferior and lateral to anchor site on glenoid (Right shoulder; Scope view from anterosuperior portal)

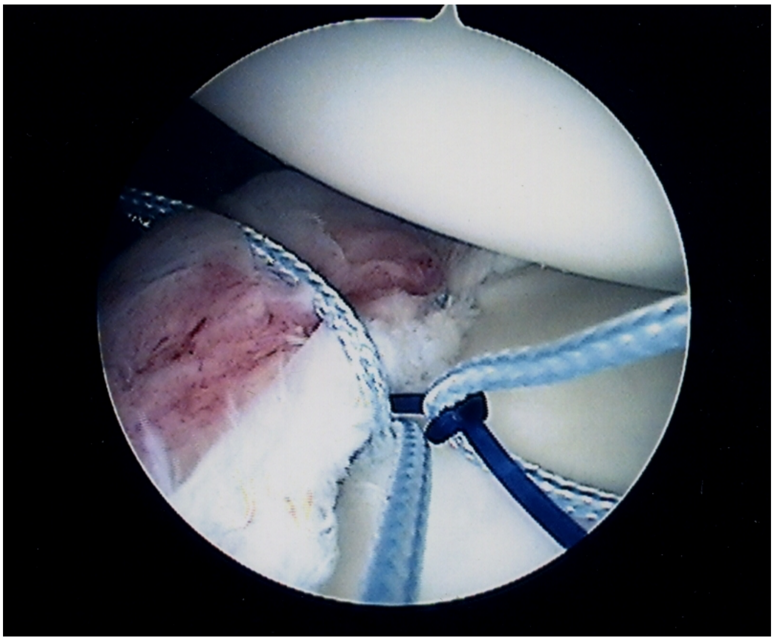


Figure 2.411 Monofilament suture shuttles inferior limb of anchor suture through capsulolabral tissue (Right shoulder; Scope view from anterosuperior portal)

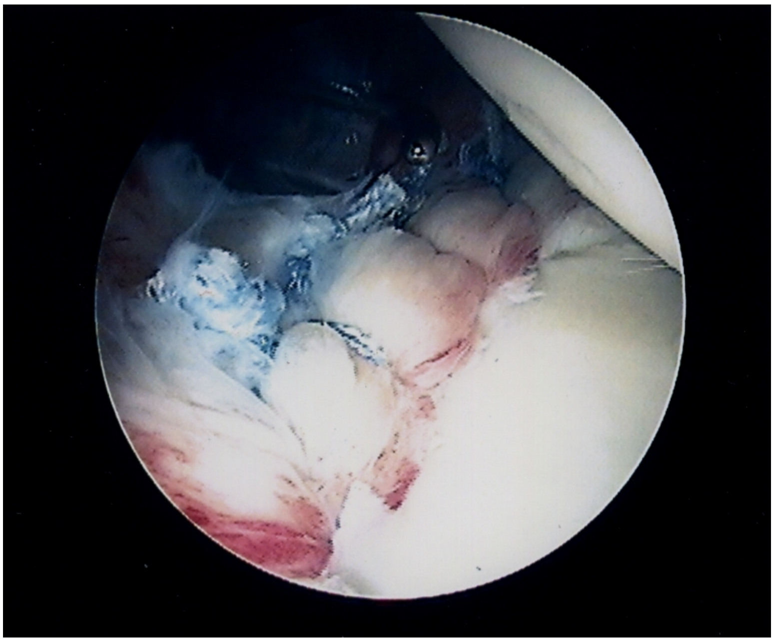


Figure 2.412 Final appearance of Bankart repair with secure reattachment of capsulolabral tissue to glenoid rim (Right shoulder; Scope view from anterosuperior portal)

2.427 Test Repair

The repair should be palpated with a nerve hook to ensure that the sutures are tight, the capsulolabral tissue securely fixed to the glenoid rim, and appropriate capsular tension has been restored. The arm is then removed from suspension and shoulder stability and acceptable range-of-motion confirmed. Absorbable subcutaneous sutures and adhesive strips complete the portal closures.

2.428 Summary

For an appropriately selected patient, an arthroscopic Bankart repair can effectively restore capsulolabral tension and functional shoulder stability while optimizing postsurgical range of motion. The patient's anticipated demands and the surgeon's familiarity and experience with arthroscopic techniques may affect the choice of an open vs. arthroscopic stabilization procedure. _Although youth is not an absolute contraindication to an arthroscopic anterior repair, if significant tissue laxity is present in a pediatric patient, the shoulder should be stabilized using an open approach. If many instability episodes have led to marked capsular attenuation, an open repair is also advisable. Significant bone loss usually requires a bone graft to restore glenoid integrity. A CT scan is particularly valuable in assessing the extent of anterior glenoid bone loss and magnitude of posterolateral humeral head impression defects. An arthroscopic Bankart repair requires adequate mobilization of the capsulolabral tissue, careful glenoid preparation, secure anchor placement, and accurate suture delivery. These steps, when properly performed and followed by appropriate rehabilitation, will lead to a high rate of success for an arthroscopic shoulder stabilization.

3. "Metric Development for an Arthroscopic Bankart Procedure: Assessment of Face and Content Validity" Angelo R L, Ryu R K N, Pedowitz R A, Gallagher A G. Metric. Arthroscopy 2015;31:1430-1440 Appendix 5

(Candidate is the first and primary author)

Purpose: To establish the metrics (operational definitions) necessary to characterize a reference arthroscopic Bankart procedure and to seek consensus from experienced shoulder arthroscopists on the appropriateness of the steps, and errors identified.

Methods: Three experienced arthroscopic shoulder surgeons and an experimental psychologist, (the Metrics Group), deconstructed an arthroscopic Bankart procedure. 14 full-length videos were analyzed to identify the essential steps and potential errors. 'Sentinel' (more serious) errors were defined as either 1) potentially jeopardizing the procedure outcome, or 2) creating iatrogenic damage to the shoulder. The metrics (operational definitions) were stress tested for clarity and the ability to be scored in binary fashion during a video review as either occurring or not occurring. The metrics were subjected to analysis by a panel of 27 experienced arthroscopic shoulder surgeons to obtain face and content validity using a modified Delphi Panel methodology (consensus opinion of experienced surgeons rendered by cyclical deliberations).

Results: 45 steps and 13 phases characterizing an arthroscopic Bankart procedure were identified. 77 potential procedural errors were specified, 20 designated as "sentinel errors". The modified Delphi Panel deliberation created changes: 2 metrics were deleted, 1 was added, and 5 were modified. Consensus on the resulting Bankart metrics was obtained and face and content validity verified.

Conclusions: This study rejects the null hypothesis and confirms that a core group of experienced arthroscopic surgeons are able to perform a task deconstruction of an arthroscopic Bankart repair and create unambiguous step and error definitions (metrics), which accurately characterize the essential components of the procedure. Analysis and revision by a larger panel of experienced arthroscopists were able to validate the Bankart metrics.

Clinical Relevance: The ability to perform task deconstruction and validate the resulting metrics will play a key role in improving surgical skills training and assessing trainee progression toward proficiency.

*see table 3.1 (glossary) for *face validity, content validity, metrics, steps, errors, sentinel errors, operational definitions, Delphi Panel (modified)*

3.1 Introduction

The intent of any surgical training program, both for residents and established surgeons acquiring a new procedural skill, is to enable the trainee to acquire the requisite skill sets necessary to perform the designated surgery well and safely. To accomplish that mission, a clearly defined endpoint or set of skill proficiencies must be identified. Further, it must be verified that mastery of those skill sets can accurately be measured during the trainee's progress. It must also be confirmed that the acquisition of those skills is predictive of the ability to perform an effective surgical procedure. Many experienced surgeons who are *proficient* (table 3.1 – glossary) in the performance of a specific procedure and are able to perform it well, are also able to identify and agree on the essential *steps* (table 3.1) to be completed as well as the *errors* (table 3.1) to be avoided for that procedure. One challenge, however, in identifying those key features is that surgeons rarely think about the procedures they perform in that level of detail. Surgeons who are proficient in the performance of a specific surgery will exhibit many if not all of the important *performance characteristics* (table 3.1) that contribute to actually performing the procedure well. They may, however, have automated to many of these steps and how they are performed and, as a consequence, may be less cognizant of the details and more granular elements of the techniques they use[176-178].

The units of performance that constitute skill can be elucidated with a *task analysis* (table 3.1) or breakdown and detailed description of the steps or actions necessary to perform the procedure. In attempting to characterize specific skills, psychologists have subjected them to a detailed task analysis and then *operationally defined* (table 3.1), rather than simply described the resulting steps. A definition specifies the order, duration, and result of the specific action, and provides precise parameters such that it can be unambiguously determined whether that specific event did or did not occur. A *description* (table 3.1), on the other hand, only offers a general characterization of an event or behavior in qualitative terms. Definitions are the preferred foundations of measurement science. The definitions or *'metrics'* (table 3.1) for a specific procedure provide a quantitative standard of measurement, which can be used to objectively assess performance. These metrics must then be validated with respect to whether their characterization fits with what is known about the skill being analyzed. The task-analysis derived characterizations or *'metric units'* (table 3.1) of skilled performance do not have to capture every aspect of performance, but should at least allow for ordinal differentiation between different levels of performance as described by Dreyfus and Dreyfus[179]. The metrics created from this analysis can serve as a tool to evaluate the effectiveness of different training protocols for a particular surgical procedure.

'*Face validity'* (table 3.1) is verified by the opinion of an experienced panel that review the content of an assessment or tool to determine if it is appropriate and relevant to the concept it purports to measure. *'Content validity'* (table 3.1) of a testing instrument is similarly obtained and based on the opinion of an experienced panel that performs a detailed examination of the contents of the test items. Thus, the face and content validity of tools assessing procedural skill are not verified by statistical analysis, but rather, by the summary opinion of an experienced panel of surgeons. An additional question that relates to establishing the validity of the metric definitions for a particular procedure asks, “Do more skilled individuals perform better on the defined metrics than less skilled or experienced individuals and do the specific metrics identify the quality, ability or trait they were designed to measure (*‘construct validity’* – table 3.1)”? In contrast to face and content validity, the establishment of construct validity requires sufficient data and statistical analysis to prove that it exists.

Task analysis for a particular surgery should be done initially for a *‘reference procedure’* (table 3.1) [180-182] – one that is straightforward with a generally accepted or agreed upon method that is uncomplicated under ideal circumstances. An optimal approach to learning should ensure that trainees are capable of performing a routine procedure before they have to deal with the technique variations necessary to address more complex pathology.

We sought to study the effectiveness of *proficiency-based progression* (PBP) (table 3.1) training plus simulation for the acquisition of surgical skills. The proficiency based progression methodology dictates that the learner must demonstrate the ability to meet specific performance benchmarks before they are permitted to progress in training. That investigation required the development and validation of specific tools to conduct the analysis. The first component needed was a *metric tool* (table 3.1), which could objectively and accurately characterize an arthroscopic Bankart repair. The development of that tool is the focus of this study. Future investigations will report on the establishment and validation of additional tools.

The purpose of this study was to establish the metrics (operational definitions) necessary to characterize a reference arthroscopic Bankart procedure and to seek consensus from experienced shoulder arthroscopists on the appropriateness of the steps, and errors identified. The null hypothesis states that face and content validity for the step and error metrics derived from a task deconstruction of an arthroscopic Bankart procedure would not be demonstrated.

3.2 Methods

3.21 Arthroscopic Bankart metric development

Three experienced arthroscopic shoulder surgeons, each with over 25 years in clinical practice, and an experimental psychologist formed the Metrics Group who characterized an arthroscopic Bankart repair. A detailed task analysis and deconstruction process (described in detail elsewhere)[180] was employed to identify the units of performance that are integral to the skilled performance of the instability repair. The goal was to characterize a "reference" arthroscopic Bankart repair and not one attempting to manage unusual or complex instability pathology. Procedure performance characterization (task deconstruction) was guided by a) decades of practice and teaching experience by the Metrics Group, b) published studies of an arthroscopic Bankart repair [183-185] and c) manufacturer guidelines on device usage. Two, 2 ½ day face-to-face meetings and eight 1 ½ - 2 hour on-line conferences were conducted along with countless email exchanges to craft the procedural metrics. For the online sessions, the use of Skype (www.skype.com) videoconferencing enabled the investigators (who reside in different geographic locations) to simultaneously review arthroscopic videos in real-time with acceptable resolution. One investigator initiated a standard Skype video connection for a group call using a laptop computer. A second computer (desktop) with a high-resolution screen was used to play the arthroscopic video being studied. An independent USB camera (Ipevo; Sunnyvale, Ca.) was connected to the USB port of the laptop and to which the Skype video input was directed instead of the resident camera on the laptop screen ("settings" tab in Skype). Thus, all of the members on the group Skype call viewed the arthroscopic video rather than the call initiator's image.

14 video recordings of a complete in vivo Bankart procedure, performed by surgeons with varying levels of experience (table 3.2), were reviewed by the Metrics Group in detail to assist in the creation and stress testing of the metrics. The videos represented surgeons with practice experience ranging from 3 – 33 years. Both the lateral decubitus (N=10) and beach chair (N=4) orientations for the patients were represented. All metrics were constructed to be applicable and able to be scored for surgeries performed with patients in both the lateral decubitus and beach chair orientations. During the series of video reviews, each metric unit was identified and the definition refined so that it could be unambiguously scored as either occurring or not occurring with a high degree of reliability by an independent group of raters. Each step was further defined by identifying beginning and end points during the procedure for that metric. The aim was that these detailed metric units would accurately capture the essence of procedure performance as well as serve as a sound and comprehensive training guide for those learning the procedure. The metrics included the specific operative steps, general order in which they should be accomplished, and the instruments and the manner in which they should be used. '*Procedural phases'* (table 3.1) were specified for groups of related steps. In addition to specifying each procedural step, metrics were also created to identify potential *errors* (table 3.1) or actions that deviate from optimal performance and should <u>not</u> be done.[186] The intent again, was to create unambiguous operational definitions (rather than descriptions) for each metric error. A special designation was made for more serious or 'sentinel' errors defined by events that, by themselves could either, 1) jeopardize the outcome of the procedure, or 2) lead to significant iatrogenic damage to the shoulder joint. An additional error characterization was termed, 'damage to non-target tissue' (DNTT). This occurrence defined an event, which was injurious to tissues not intentionally being addressed during the defined task, i.e. 'scuffing of articular cartilage by an instrument', or 'lacerating the intact labrum'.

By agreed upon convention, an event (step or error) must be observed on the video to be scored. Thus, inference that an event was "likely to have occurred" was eliminated. For example, if comparable views of the anterior humeral head showed relatively healthy or pristine articular cartilage early in the procedure, with scuffing and abrasion later during the repair, but the injurious event was not observed on the video, it was not scored as an error (or damage to non-target tissue).

3.22 Metric stress testing and reliability of identification

After the 4 members of the Metrics Group were satisfied that the entirety of the procedure had been well characterized, they *'stress tested'* (table 3.1) the metrics by subjecting them to a robust assessment of how reliably they could be independently scored in blinded fashion. Eight video recordings of complete arthroscopic Bankart procedures that were performed by surgeons possessing a wide range of technical skill were independently reviewed and scored. Both the lateral decubitus and beach chair orientations were represented by the videos studied. Each metric was scored in binary fashion as either a "yes" or "no" (occurring or not occurring). After each video review, differences in the scoring of each metric by the reviewers were compared and discussed. Where necessary, operational definitions were clarified, modified, or dropped and new ones added to optimize the functionality of the characterizations as a whole. This process of independent viewing, scoring, and revising the step and error metrics was continued until the Metrics Group was satisfied that the metrics accurately and unambiguously characterized the specifics of an arthroscopic Bankart procedure and could be *'reliably identified'* (table 3.1) by independent reviewers. The extent of agreement between two raters for the entire group of step and error Bankart metrics could potentially range between 0 = no agreement, to 1.0 = complete agreement.

3.23 Face and Content Validation of the Bankart Metrics by a modified Delphi Panel

The *"Delphi Panel method"*[187] (table 3.1) is a process that provides an interactive communication structure between researchers (i.e., the Metric Group authors) and an experienced panel (see below) in a field or discipline in order to provide systematic feedback on a given topic (i.e., the accuracy of the metrics developed for a reference approach to a Bankart procedure). The Delphi method employs an '*iterative' process'* (table 1) for progressing toward a desired result by means of repeated cycles of deliberations. The iterative process should be convergent, i.e., it should come closer to the desired result as the number of iterations or cycles of review increases. For the Bankart characterization, the desired result (consensus on the appropriateness of a particular metric) was obtained by means of repeated cycles of questioning, deliberation, metric modification, and voting on the appropriateness of each refined metric definition. The methodology assumes that good quality knowledge evolves from the process. The Delphi method was modified to the extent that the voting cycles, with each new iteration, were not anonymous.

The determination of face and content validity for the Bankart characterization was made by subjecting each metric to an appraisal by a group of surgeons who were very experienced in the performance of an arthroscopic Bankart repair. 27 board certified orthopedic surgeons (the three Metrics Group surgeons and 24 additional Arthroscopy Association of North America [AANA] shoulder faculty instructors) with an average of over 23 years in clinical practice involving shoulder arthroscopy participated in a modified Delphi Panel. Four of the panelists are full-time academicians, 9 are in private group practice and have direct involvement in teaching fellows, and 14 are in private practice with a clinical affiliation with a University Orthopedic Department. Each member of the "Delphi Panel" is a Master or

Associate Master faculty member for AANA and has taught the technique for an arthroscopic Bankart repair during shoulder courses conducted at the Orthopedic Learning Center (Rosemont, Il.). An experimental psychologist (AGG) facilitated the meeting.

3.24 Delphi Panel Procedure

An overview of the project and meeting objectives was presented. Background information regarding proficiency based progression training, prior literature demonstrating the validity of that training approach for procedural specialties, and the specific objectives of the current Delphi Panel[187] were reviewed. It was explained to the panel that the Bankart metrics had been developed by the Metrics Group for a reference approach to arthroscopic anterior shoulder stabilization for unidirectional anterior glenohumeral instability[180, 181]. It was acknowledged that the designated reference procedure might not reflect the exact techniques employed by individual panelists, but that the operative steps which were presented, accurately embodied the essential and key components of the procedure. An affirmative vote by a panel member indicated that the metric definition presented was accurate and acceptable as written, but not necessarily that it was the manner in which that particular panelist might have chosen to complete the step. "Consensus" meant that there was unanimity in voting among the panelists, and that a particular metric definition was "not wrong or inappropriate". Each of the procedural steps and potential errors were evaluated individually. After each metric definition was presented, panel members voted on whether or not the metric was acceptable as written. If the panel could not achieve consensus due to lack of clarity or differences in opinion, the metric definition was revised accordingly and a new vote conducted on the acceptability of the modified metric. This process was repeated until the metric was accepted. If consensus could not be achieved through a series of modifications, the metric was deleted. When it was deemed necessary, a new metric was defined and added.

3.3 Results

3.31 Bankart Procedure Metrics

The step metrics resulting from the task deconstruction were grouped into 13 separate phases of the procedure (in Roman numerals). Each phase (i.e. “Arthroscopic Instability Assessment”, “Inferior Anchor Preparation / Insertion”) contains a series of related, unambiguously defined, observable procedure events (steps) with specific beginning and ending points. All potential errors identified had been noted to occur during the stress testing of the metrics. Some of the identical errors and sentinel errors could occur during different phases of the procedure that recurred during the 3 anchor repair, i.e. “uncorrected entanglement of shuttling device or suture”.

3.32 Modified Delphi Panel

All phases of the procedure were accepted as identified. Only a minority of procedure phases and their associated metrics were accepted without discussion. At the conclusion of the deliberations, consensus amongst the Delphi Panel was reached for 45 steps, 77 errors (29 unique) and 20 sentinel errors (8 unique) (tables 3.3 and 3.4). During the panel deliberations, 2 metrics were deleted, 1 added, and 5 modified before consensus was achieved (Table 3.5). The comments and recommendations for each of the phases, steps, and errors, with the associated Delphi Panel vote, are presented in Table 3.6.

3.4 Discussion

The principle findings of this study are; 1) an arthroscopic Bankart procedure can be deconstructed into the essential steps necessary for the effective completion of the repair, 2) the potential errors related to the procedure are able to be identified and characterized, and 3) face and content validity for the resulting step and error metrics can be obtained through employment of the modified Delphi Panel technique.

Traditionally, surgeons have been trained using the 'apprenticeship' model, which is "process", or time based (i.e.- a certain variety of rotations, exposure to numbers of specific cases, etc.). A paradigm shift toward 'proficiency-based progression' training, which is "outcomes" based is occurring and mandates that the trainee be able to demonstrate the ability to meet specific skills benchmarks in order to progress in training. Those benchmarks must have specific, clear, objective, and fair standards of performance. Validated metrics will be essential in defining those standards. In addition, as the move toward including surgical skills credentialing and procedural competency occurs for licensing, the same validated standards will be needed. The methodology employed in this study provides a framework for the development of those metrics and standards.

An arthroscopic Bankart (index procedure) was selected as the reference surgical procedure to study for several reasons. For the patient with unidirectional anterior instability due primarily to a Bankart lesion without significant bone loss, a suture anchor repair employing 3 implants is a commonly accepted method employed to obtain a successful patient outcome.[188-191] In addition, the essential components of the procedure are well outlined regardless of whether the patient is placed in the lateral decubitus or beach chair orientation. [89, 155, 192-194]

The task analysis stage of metric development is crucial as metrics are the fundamental building blocks of a good training program. Metrics, thus, not only define how the training should be characterized and the procedure performed by the trainee, but must also afford the opportunity for meaningful assessment of the trainee's performance and progress. The entire process of metric development should be as transparent, objective, and unambiguous as possible. Metric definitions should be characterized in such a way that they are sufficiently complete and detailed for an individual, not associated with the initial development process, to use them to score performance reliably. Metric definitions should include behavioral markers that indicate the beginning and endpoints of the performance characteristics (steps) to be assessed. These parameters will become particularly important in the future as the procedural metrics are employed with higher fidelity simulators. The details of the metric definitions will be necessary for the simulator to be appropriately programmed to provide the trainee with performance assessments and accurate feedback.

Other approaches to the measurement of surgical performance use qualitative descriptions of performance and require the user to rate items on a graduated *'Likert-type scale'* (table 3.1), which ascribes a quantitative value to qualitative data to make it amenable to statistical analysis. Likert scales (often with a range from 1-5, or 1-7) are typically constructed with responses (opinion) around a neutral option (i.e. "suture delivery was: 1=awkward …3=effective…5=highly efficient") and were originally designed to assess a range of attitudes. Because of the component of subjectivity, this method of attempting to rate objective performance can render it difficult to obtain acceptable levels of inter-rater reliability [>80%] in the scoring of events. In contrast, the approach to the assessment of performance used in this study employs precise definitions of performance and simply requires the reviewer to report whether the specific event occurred or not. This binary approach to the measurement of individual events has been shown to facilitate the reliable scoring of metric-based performance units across a variety of functions during skills training[195-198] of individuals with different experience levels.[199, 200] This approach has also been shown to be more reliable than Likert-scale scoring.[63]

Behaviors that deviate from optimal performance (errors) can be characterized including those of a “more serious nature”. The issue of whether those more serious errors should be termed a ‘critical error’ or some alternative label was raised at the outset during the metric definition process. It was agreed upon by the Metrics Group that use of the term ‘critical error’ could imply that the event was ‘life-threatening’ or might infer serious medico-legal implications. It was elected instead, to use the term *‘sentinel’* (table 3.1) to connote an error that should be carefully ‘watched for and to avoid’. Sentinel errors involve a serious deviation from optimal performance during a procedure because they can either jeopardize the success / desired result of the procedure, or create iatrogenic insult to patient’s tissues. A single specific sentinel error may not always lead to a bad outcome but should stringently be avoided.[180] The underlying philosophy of this approach to errors is that suboptimal outcomes do not happen by accident but usually result from the coalescence of deviations from optimal procedure performance.

The face and content validity of the metric-based procedure characterization by subject specialists and can be accomplished using the modified Delphi Panel methodology reported in this study. The metrics developed were informed by research studies, professional guidelines, clinical experience, and manufacturers’ guidelines.[180, 181] Although the surgeons in the Metrics Group are very experienced in the performance of a Bankart procedure, the Delphi process provided an excellent method to ensure that the procedure characterization is appropriate, represents best practice, and is acceptable to a larger group of experienced master and associate master Bankart faculty. As anticipated, many surgeons pointed out that they might perform a specific step in a different manner, but that the approach outlined by the Metrics Group was ‘not incorrect’ or inadvisable. For the majority of the procedure metrics, the members of the panel made very helpful suggestions for improving the definitions.

Assuming that the Bankart metric identifications and definitions represent a 'real-world' surgical procedure, these performance characteristics should be able to distinguish between experienced (skilled) surgeons and novices, i.e., provide construct validity. Future studies regarding construct validity will seek to provide information about which metrics best distinguish between experienced and novice surgeon performance. That information will facilitate the establishment of a benchmark to define the 'proficiency level', which trainees should acquire before progressing to *in vivo* practice.[79, 201]

3.41 Limitations

A limitation of this study resides in the fact that every potential error, regardless of how rare the occurrence, might not have been included. The Delphi Panel, however confirmed that the errors listed were those most likely to occur and that should be avoided in the safe performance of an arthroscopic Bankart repair. Although common errors may be relatively easy to agree on, it is somewhat more challenging to decide on which errors should be designated as "sentinel", without a specific weighting methodology. While the issue of employing that designation for events that cause iatrogenic damage is more straightforward, the concept of also using the term for events which might 'potentially lead to a suboptimal outcome', is more subject to the opinion of the Metrics Group and the Delphi Panel.

Further, data is not available to confirm that the specific steps identified by the Metrics Group and the Delphi Panel directly correlate with a successful surgical outcome for patients with unidirectional shoulder instability. Therefore, the metrics created remain predominantly based on the opinion of experienced surgeons and instructors. An outcomes study will be needed to fully establish the predictive validity for the Bankart metrics as authored. In addition, the Metrics Group and Delphi

Panel were all North American surgeons. International arthroscopists may have created somewhat different metrics for an arthroscopic Bankart repair.

3.5 Conclusions

This study rejects the null hypothesis and confirms that a core group of experienced arthroscopic surgeons are able to perform a task deconstruction of an arthroscopic Bankart repair and create unambiguous step and error definitions (metrics), which accurately characterize the essential components of the procedure. Analysis and revision by a larger panel of experienced arthroscopists were able to validate the Bankart metrics.

Table 3.1 - Glossary

	Definition
Behavioral scientist	A professional who engages in any discipline concerned specifically with the subject of human actions and behavior
Concurrent validity	A type of evidence in which there is a positive relation between the test scores from one instrument and the scores from another instrument purporting to measure the same construct
Construct validity	A type of evidence that supports that specific test items identify the quality, ability, or trait they were designed to measure
Content validity	An estimate (by expert/experienced opinion) of the validity of a testing instrument based on a detailed examination of the contents of the test items
Definition	A definite, distinct, and clear objective characterization providing an accurate and reliable identification of whether an event was or was not observed to have occurred
Delphi Panel (modified)	A structured communication technique originally developed as a systematic, interactive forecasting method that relies on the opinion of an experienced panel; in the modified form, the members of the panel answer queries/vote in 2 or more rounds (cycles) on the appropriateness of the metric-based operational definitions of detailed aspects of procedure performance with the goal of achieving consensus—voting is not anonymous
Description	A qualitative characterization of certain or salient aspects or features of an event
Error	A deviation from optimal performance
Face validity	An estimate by an experienced panel that reviews the content of an assessment or tool to see if it seems appropriate and relevant to the concept it purports to measure
Iterative process	A process for calculating or progressing toward a desired result by means of repeated cycles of operations (deliberations); an iterative process should be convergent, that is, it should come closer to the desired result as the number of iterations increases
Likert-like scale	A method of ascribing a quantitative value to qualitative data to make them amenable to statistical analysis
Metric	A standard of measurement of quantitative assessments used for objective evaluations to make comparisons or to track performance
Metric stress testing	A method for determining how specific metric definitions fare during their application and use in scoring in vivo or video-recorded performances
Metric unit	A method of measurement in which the basic parts or components are discrete performance elements
Operational definition	Terms used to define a variable or event in terms of a process (or set of validation tests) needed to determine its existence, quantity, and duration
Performance characteristics	The features determining the accomplishment of a given task measured against preset known standards of accuracy and completeness
Predictive validity	A type of evidence that determines the extent to which the scores on a test are predictive of actual performance
Procedure phase	A group or series of integrally related events or actions that, when combined with other phases, make up or constitute a complete operative procedure
Proficiency/proficient	A specific level of performance defined by a quantitative score (benchmark) or scores on a standardized test or other form of assessment
Reference procedure	A straightforward operative procedure; an agreed on/accepted approach to the performance of an uncomplicated surgical procedure
Reliability of identification (inter-rater reliability)	The extent of agreement between 2 raters on the occurrence of a series of observed events; it ranges between 0, no agreement, and 1.0, complete agreement
Sentinel error	An event or occurrence involving a serious deviation from optimal performance during a procedure that either (1) jeopardizes the success/desired result of the procedure or (2) creates iatrogenic insult to a patient's tissues
Step	A component task, the series aggregate of which constitutes the completion of a specific procedure
Task analysis	An assessment of how a procedure is accomplished, including a detailed (functional) description of the manual activities or tasks along with their duration, frequency, and complexity and any other unique and distinguishing factors
Task deconstruction	To break down a procedure into constituent tasks, steps, or components

Table 3.2

Source videos for Bankart metric creation and stress testing.

	Surgeon - Years in Practice	Patient Orientation
1	25	LD
2	17	LD
3	25	LD
4	26	LD
5	25	LD
6	17	BC
7	18	BC
8	26	LD
9	28	LD
10	21	LD
11	24	LD
12	25	LD
13	3	BC
14	4	BC
		LD = lateral decub. BC = beach chair

Table 3.3. The 13 phases of the Bankart procedure (in Roman numerals) and a brief summary of the 45 steps of the procedure. (© AANA, 2013)

I. Portals
1. Posterior portal established
2. View posterior humeral head and extent of the Hill-Sachs when present
3. Introduce mid-anterior spinal needle immediately superior to the subscapularis and direct it toward the anteroinferior glenoid and labrum
4. Establish a cannula that abuts the superior border of the subscapularis near the lateral subscapularis insertion
5. Demonstrate instrument access to the anteroinferior glenoid/labrum
6. Introduce anterosuperior spinal needle at the superolateral aspect of the rotator interval and direct it toward the anterior glenoid
7. Establish an anterosuperior cannula, arthroscopic sheath, or switching stick
II. Arthroscopic instability assessment
View from posterior portal
8. View or probe the superior labral attachment onto the glenoid
9. View or probe articular surface of the rotator cuff
10. Probe anteroinferior glenoid/Bankart pathology including rim fracture, articular defect
View from anterosuperior portal
11. View or probe the midsubstance of the anterior-inferior glenohumeral ligaments
12. View or probe the insertion of the anterior glenohumeral ligaments onto the anterior humeral neck
III. Capsulolabral mobilization/glenoid preparation
13. Elevate the capsulolabral tissue from the glenoid neck and articular margin
14. View the subscapularis muscle superficial to the mobilized capsule
15. With an instrument, grasp and perform an inferior to superior shift of the capsulolabral tissue (to show tension is restored)
16. Obtain a view of the anterior glenoid neck
17. Mechanically abrade the glenoid neck
IV. Inferior anchor preparation/insertion
18. Seat the guide for the most inferior anchor hole at the inferior region of the anteroinferior quadrant
19. Drill anchor hole oblique to the glenoid articular face
20. Insert anchor
21. Test anchor security by pulling on suture tails
V. Suture delivery/management
22. Pass a cannulated suture hook or suture retriever through the capsular tissue inferior to the anchor
23. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula
VI. Knot tying
24. Deliver an arthroscopic sliding knot
25. Back up with 3 or 4 half-hitches
26. Cut suture tails
VII. Second anchor preparation/insertion
27. Seat the drill guide for the second anchor superior to the first anchor and inferior to the equator of the glenoid
28. Drill anchor hole oblique to the glenoid articular face
29. Insert suture anchor
30. Test anchor security by pulling on suture tails
VIII. Suture delivery/management
31. Pass a cannulated suture hook or suture retriever through the capsular tissue inferior to the suture anchor
32. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula
IX. Knot tying
33. Deliver an arthroscopic sliding knot
34. Back up with 3 or 4 half-hitches
35. Cut suture tails
X. Third anchor preparation/insertion
36. Seat the drill guide for the third anchor at or superior to the equator
37. Drill anchor hole oblique to the glenoid articular face
38. Insert suture anchor
39. Test anchor security by pulling on suture tails
XI. Suture delivery/management
40. Pass a cannulated suture hook or suture retriever through the capsular at or inferior to the suture anchor
41. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula
XII. Knot tying
42. Deliver an arthroscopic sliding knot
43. Back up with 3 or 4 half-hitches
44. Cut suture tails
XIII. Procedure review
45. View and/or probe final completed repair

Table 3.4. A summary of the 29 different Bankart procedure metric errors. Metric errors can be associated with multiple phases and steps of the procedure. (© AANA, 2013)

1. Failure to maintain intra-articular position of the posterior cannula
2. Failure to maintain intra-articular position of the mid-anterior cannula
3. Failure to maintain intra-articular position of the anterosuperior cannula
4. Damage to the superior border of the subscapularis during creation of the midanterior portal
5. Damage to the anterior border of the supraspinatus during creation of the anterosuperior portal
6. Loss of intra-articular position of arthroscope/sheath or operating cannula (loss of each portal is scored only once for each Roman numeral, i.e., up to a total of 3 for arthroscope + 2 portals)
7. Laceration of intact capsulolabral tissue (sentinel error)
8. Failure to maintain control of a working instrument (sentinel error)
9. Guide is not located in the inferior region of the anteroinferior quadrant of the glenoid for the most inferior anchor
10. Entry of the completed tunnel lies outside safe zone of 0 to 3 mm from the bony glenoid rim (sentinel error)
11. Shallow undermining and deformation of articular cartilage (sentinel error)
12. Failure to maintain secure seating of the drill guide during anchor insertion
13. Breakage of the implant
14. Implant remains visibly proud (sentinel error)
15. Failure to insert the anchor with the inserter laser line (when present) to or beyond the laser line on the drill guide
16. Anchor fails to remain securely fixed within bone at the appropriate depth
17. Capsular penetration is at or superior to anchor hole (sentinel error)
18. Capsular penetration is not at or peripheral to the capsulolabral junction
19. Instrument breakage
20. Tearing of capsulolabral tissue
21. Uncorrected entanglement of shuttling device or suture
22. Off-loading of suture anchor
23. Breakage of suturing device
24. Failure to create and maintain indentation of the capsule or labral tissue on knot completion (sentinel error)
25. Visible void is present between throws of the completed primary knot (sentinel error)
26. Completed knot abuts articular cartilage
27. Visible void is present between throws of the completed half-hitches
28. Suture breakage
29. Guide is inferior to the equator of the glenoid for the third anchor position

NOTE. Metric errors can be associated with multiple phases and steps of the procedure.

Table 3.5 – Delphi Panel metric changes

Modification	Issue	Deliberation	Resolution
Deleted (2)	Should failure of anchor purchase remain an error?	Cadaveric bone variability too great to score accurately	Delete error
	Alternating posts for knot tying	Arthroscope views not consistent enough to score reliably	Delete error
Added (1)	Completed knot position	May cause iatrogenic damage if it abuts articular cartilage	Add error to each of the 3 knots tied
Modified (5)	Diagnostic steps—probe or view (how long?)	Does "looking" equal "ascertaining"?	View must be held long enough to determine pathology
	Adequacy of capsular mobilization	Should demonstrate effort at capsular mobility	Must take instrument (grasper) to demonstrate
	Whether to ascribe "critical" to "laceration of labrum"	Labral variability too great in cadavers; can still see violation of "hoop" integrity	Retain as a sentinel error
	Consider deleting the term "sliding" from knot description	Sliding knot would be acceptable for a reference procedure	Retain the term "sliding"
	Should diagnostic steps be included in the procedure metrics?	May skew results if there is excessive influence of diagnostic steps in procedure	Include only steps directly related to instability assessment

Table 3.6 -A summary of the points raised / voting outcomes of the Bankart Delphi Panel. Minutes of Metric Validation Meeting: "Copernicus Study / Delphi Panel" (Nov. 18, 2011)

Rick Angelo, MD, Meeting Chair (Recorded by Robert Pedowitz, MD, PhD)

Attendees: R. Angelo, R. Ryu, R. Pedowitz, J. Tokish, R. Bell, R. Hunter, K. Nord, V. Goradia, A. Barber, S. Snyder, B. Beach, J. Abrams, B. Shaffer, J. Tauro, L. Higgins, S. Weber, S. Koo, D. Richards, J. Esch, J. Dodds, J. Randle, J. Richmond, A. Curtis, J. Burns, N. Sgaglione, J. Kelly, S. Powell (27 voting attendees), T. Gallagher (meeting facilitator)

Meeting Overview

A) Dr. Angelo presented a brief overview of the project and meeting objectives.

B) Dr. Gallagher presented the background of proficiency based training, some prior literature demonstrating the validity of this training approach for procedural specialties, and he explained the specific objectives of the current "Delphi" meeting.

C) Dr. Angelo presented each procedural step and explained the associated metrics that have been developed by the Project Leadership Team for a reference approach to anterior shoulder stabilization for glenohumeral instability (version 18, 11/16/11, attached). The comments and recommendations for each of the steps, with associated vote, are presented in the Table, below.

Comments and Recommendations Regarding Procedural Steps and Metrics	Vote on Steps and Metrics
I. Portals (steps 1-7)	Unanimous affirmative
Agreement that this is an outside-in reference approach for portal placement, though some surgeons use an inside-out approach	
Importance of pre-surgery setup, though assessment of this phase would be very difficult using arthroscopic videotapes	
II. Diagnostic arthroscopy (steps 8-12)	Unanimous affirmative
Clarified metric "view or probe," not "view and probe"	
Discussion about whether we should include metrics for view or probe of posterior labrum, superior labrum, biceps, and rotator cuff	
Recommendation: Limit diagnostic elements for the current procedural assessment, and consider creation of a diagnostic arthroscopy reference procedure (vote to drop diagnostic elements failed—because one component of the metrics is teaching essential components of the procedure)	
III. Capsulolabral mobilization (steps 13-15)	Unanimous affirmative
Should a laceration of the labrum be defined as a "critical error"?	
Should the pre-existing tissue quality of the labrum be assessed so that laceration of poor tissue does not qualify as a critical error?	
Consider adjustment of metric definition to describe grasping of the anatomic structure of the anterior inferior glenohumeral ligament	
IV. Glenoid neck preparation (steps 16-17)	Unanimous affirmative
V. Insertion of first anchor (steps 18-21)	Unanimous affirmative
Failure of anchor from bone should not be considered a critical error because loss of fixation could be related to bone quality	
VI. Suture management—first anchor (steps 22-23)	Unanimous affirmative
Should we include retrieval of broken suturing device (generally thought this would be quite rare, so not a useful metric)?	
The definition for adequacy of capsulolabral tissue capture seems adequate	
VII. Knot tying—first anchor (steps 24-26)	Unanimous affirmative
Consider deleting the term "sliding" from the knot description	
Add error of a knot completed and left on the articular surface	
Need to drop "alternating posts" metric	
VIII. Insertion of second anchor (steps 27-30)	Unanimous affirmative
Failure of anchor from bone should not be considered a critical error because loss of fixation could be related to bone quality	
IX. Suture management—second anchor (steps 31-32)	Unanimous affirmative
X. Knot tying—second anchor (steps 33-35)	Unanimous affirmative
Consider deleting the term "sliding" from the knot description	
Add error of a knot completed and left on the articular surface	
Probably need to drop "alternating posts" metric	
XI. Insertion of third anchor (steps 36-39)	Unanimous affirmative
Failure of anchor from bone should not be considered a critical error because loss of fixation could be related to bone quality	
XII. Suture management—third anchor (steps 40-41)	Unanimous affirmative
XIII. Knot tying—third anchor (steps 42-44)	Unanimous affirmative
Consider deleting the term "sliding" from the knot description	
Add error of a knot completed and left on the articular surface	
Need to drop "alternating posts" metric	
XIV. Final assessment (step 45)	Unanimous affirmative

4. "The Bankart Performance Metrics Combined With a Shoulder Model Simulator Create a Precise and Accurate Training Tool for Measuring Surgeon Skill" Angelo R L, Pedowitz R A, Ryu R K N, Gallagher A G. Arthroscopy 2015;32:1639-1654. Appendix 6

(Candidate is the first and primary author)

Abstract

Purpose: To determine if a dry shoulder model simulator coupled with previously validated performance metrics for an arthroscopic Bankart repair (ABR) is a valid tool with the ability to discriminate between the performance of experienced and novice surgeons; To establish a proficiency benchmark for an ABR using a model simulator.

Methods: We compare an experienced group (N = 12) of arthroscopic shoulder surgeons (Arthroscopy Association of North America Faculty) to a novice group (N = 7) (postgraduate year 4 or 5 orthopedic residents). All surgeons were instructed to perform a diagnostic arthroscopy and a 3-suture anchor Bankart repair on a dry shoulder model. Each procedure was videotaped in its entirety and scored in blinded fashion independently by 2 trained reviewers (N = 10 reviewers). Scoring employed previously validated metrics for an ABR and included steps, errors, and 'sentinel' (more serious) errors.

Results: The inter-rater reliability among pairs of raters averaged 0.93. The experienced group made 63% fewer errors, committed 79% fewer sentinel errors, and performed the procedure in 42% less time than the novice group (all significant differences). The greatest difference in errors between the groups involved anchor preparation and insertion, suture delivery and management, and knot tying.

Conclusions: The tool comprised of validated arthroscopic Bankart repair metrics coupled with a dry shoulder model simulator is able to accurately distinguish between the performance of experienced and novice orthopedic surgeons. A performance benchmark based on the

experienced group includes: completion of a 3 anchor Bankart repair, and enacting no more than 4 total errors and 1 sentinel error.

Clinical Relevance: The combination of performance metrics and an arthroscopic shoulder model simulator can be used to improve the effectiveness of surgical skills training for an ABR. The methodology employed may serve as a template for outcomes based procedural skills training in general.

4.1 Introduction

Some authors from professional bodies and health care training organizations around the world argue that the surgical trainee should acquire basic procedural skills outside of the surgical theater before operating on real patients.[202-204] Furthermore, evidence now clearly indicates that when performed to a quantitatively defined level, skills practiced and acquired outside the operating room are superior to skills acquired in a traditional apprenticeship manner primarily in the operating room.[205, 206] Satava first introduced the concept of simulation-based training in the early 1990s[26] with quantitative evidence from prospective, randomized, double blinded clinical studies showing that simulation-based training is a powerful tool for the acquisition of surgical skills.[22, 198, 207, 208] The "simulator" can either be a physical model or computer-generated video images,[22, 198] as both are equally effective if used as part of a *metric-based training curriculum*[180] (table 3.1 – glossary).

An implicit assumption in a simulator-based training process is the use of validated metrics, which appropriately characterize the procedure to be trained. Previously, Angelo et al.[65] reported on the development of a tool defining *performance metrics* (*steps* and *errors* – table 4.1) for a standard *reference approach* (table 3.1) to performing an arthroscopic Bankart repair.[185, 190, 191, 193, 209] That tool was derived from a careful *task deconstruction* (table 3.1) using videos of complete Bankart procedures performed with patients in either the lateral decubitus or beach chair orientation. The metrics were constructed so that they could be scored in an identical manner with the patient in either orientation. *Face and content validity* (table 3.1) of the metrics were verified using a *modified Delphi Panel* methodology (table 3.1).

The purpose of this study was to determine if a dry shoulder model simulator coupled with previously validated performance metrics for an arthroscopic Bankart repair is a valid tool with the ability to discriminate between the performance of experienced and novice surgeons. We also sought to establish a *proficiency* (table 3.1) benchmark for that procedure using the model simulator. The null hypothesis states that when using a shoulder model simulator, the Bankart metrics would fail to discriminate between experienced and novice surgeon performance.

4.2 Methods

No IRB was obtained for this study investigating the validity of the Bankart metrics coupled with the model simulator. An IRB was sought for the final Copernicus Study proper, which will compare 3 different training protocols. The Western Institutional Review Board (WIRB) (#1-776362-1) opined that, as an educational curriculum study, it was "Exempt" from the need for full IRB approval (criteria: 45 CFR 46.101(b)(1)). The final study comparing the 3 training protocols was registered with the NIH (ClinicalTrials.gov: #NCT01921621).

4.21 Study Groups

Two groups were compared in their performance of an arthroscopic Bankart procedure on a shoulder model simulator. The experienced group consisted of all faculty members who served as Master and Associate Master instructors for a standard 3 day Arthroscopy

Association of North America (AANA) Resident Course conducted at the Orthopedic Learning Center (Chicago, Il.). The novice group was limited to PGY 4 and PGY 5 orthopedic residents who had registered for a Resident's Course and who volunteered to participate in the investigation.

4.22 Arthroscopic Bankart Repair Metrics

Metrics have been previously defined for a standard reference arthroscopic Bankart repair.[65] Forty five essential steps in 13 *phases* (Table 3.1) (Roman numerals) were defined with beginning and end points (Table 3.3). 29 potential unique errors were specified (Table 3.4), 8 of which were designated as "*sentinel*" (table 3.1). The more serious sentinel errors were defined as those expected to either; 1) substantially compromise the outcome of the shoulder stabilization (i.e. – 'capsular penetration of the suture passing instrument is superior to the anchor hole', resulting in failure to retention the capsule and inferior glenohumeral ligaments), or 2) potentially lead to iatrogenic damage to the shoulder (i.e. 'laceration of the intact labrum'). Some of the same errors could be enacted more than once during different phases of the procedure. Thus, a total of 77 potential errors, 20 of which were sentinel errors, were specified for the complete procedure. In addition, events that led to less consequential "*damage to non-target tissues*" (DNTT) (table 3.1) were recorded as a standard error (i.e. scuffing of the articular cartilage). A perfect score would indicate that all 45 steps were completed satisfactorily without committing any errors.

4.23 Dry Shoulder Model Simulator

The shoulder simulator employed is a physical model comprised of a dense foam plastic endoskeleton including a humerus, scapula, glenoid, coracoid, acromial spine and acromion with proportions appropriate to the human skeleton (Sawbones, Vashon Is, Washington) (Figure 3.1). The articulating surfaces of the humerus and glenoid are laminated with a softer, white layer designed to mimic articular cartilage. A Hill-Sachs lesion measuring 1 cm by 3.5 cm is oriented vertically on the posterior aspect of the humeral head and is represented by a red impaction trough. A rim of off-white, rubber-like material encircles and lightly adheres to the glenoid neck, simulating the labrum. Red staining in the region where the labrum is joined to the anteroinferior glenoid represented the Bankart lesion. The adhesive attachment of the labrum requires the operator to intentionally "liberate" the labrum from the glenoid to demonstrate mobilization of the capsulolabral tissues. A more medial and superficial pink layer of soft foam represents the subscapularis muscle. A tubular strand of rubber simulates the long head biceps tendon and courses from its anatomic attachment to the superior labrum, out of the shoulder joint into the bicipital groove of the humerus. The capsule is replicated by a pliable, rubberized material containing the glenohumeral joint and has a molded imprint of the inferior glenohumeral ligaments on the articular surface. A separate band represents the superior boarder of the subscapularis tendon. Holes measuring 8 mm in diameter were created in the capsule during molding and enable cannulas for the posterior, mid-anterior, and anterosuperior portals to pass through the relatively tough capsular material. Beige-colored, soft, moderately dense foam represents the skin and soft tissues exterior to the glenohumeral joint and possesses a contour and bulk that mimics the shape of the human shoulder. The acromion, acromial spine, and coracoid landmarks are readily palpable through the "soft tissues" and assist in locating proper portal placement.

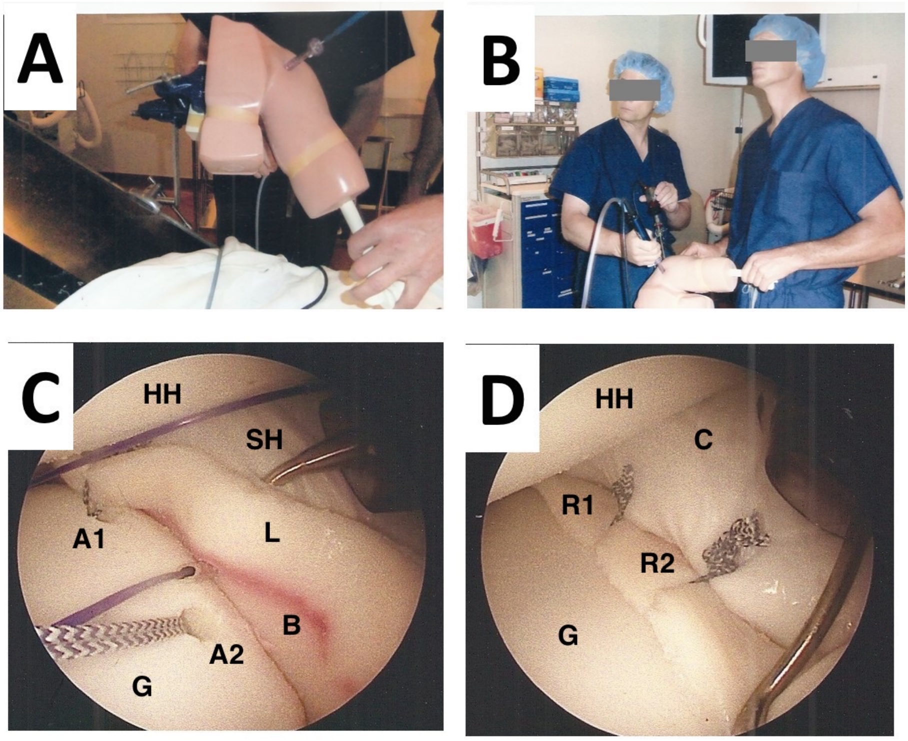


Figure 1A - D

Figure 4.1 A – D. The dry shoulder simulator model used in the current study (left shoulder):

A – Anterior view of the shoulder simulator oriented in the beach chair position

B – An operator and assistant performing arthroscopic surgery on the simulator model oriented in the lateral decubitus position

C – Arthroscopic view from an anterosuperior portal; anterior is right (HH = humeral head, G = glenoid, A1,2 = anchor position 1 and 2, L = labrum, B = Bankart lesion, SH = suture hook) Shows the inferior-most anchor and sutures in place and completed. The second anchor has been inserted and the sutures retrieved out the posterior cannula. A cannulated suture hook enters through the mid-anterior portal and has been passed through the capsule and labrum inferior to the exit of the suture anchor hole; a monofilament shuttle suture is then delivered.

D - (HH = humeral head, G = glenoid, R2,3 = repair position 2 and 3, C = capsule) demonstrates a hook probe examining the completed repair (with the 3rd anchor just out of view beneath the hook probe); the capsule has been re-tensioned and the labrum secured to the glenoid rim.

4.24 Arthroscopic Bankart Repair

During a single weekend AANA resident arthroscopy course, the surgeons from both groups were instructed to establish portals (posterior, anterosuperior, and mid-anterior), complete a thorough diagnostic arthroscopy, and perform a 3 anchor arthroscopic Bankart repair on the simulator model. Further, they were instructed to demonstrate / complete all of the steps for the Bankart repair that they would normally perform in clinical practice on a real patient. The model was secured in either the lateral decubitus or beach chair orientation according to surgeon preference. Equipment representatives from multiple different vendors served as surgical assistants and were randomly assigned to participant surgeons. The assistants were instructed to act only at the specific direction of the operating surgeon. Prompting and coaching (of technique) were prohibited (the procedures were proctored by staff from the Orthopedic Learning Center). A standard equipment tower with a 30^0 arthroscope was provided along with all instruments necessary to complete an arthroscopic Bankart repair (table 4.1).

The surgeon created the required portals based on the palpable “bony” landmarks of the shoulder and then progressed to complete the diagnostic arthroscopy and Bankart repair. A continuous video recording was made beginning with the first arthroscopic view of the joint from the posterior portal and ending with the withdrawal of the arthroscope after the surgeon’s examination of the completed repair with a hook probe. No time limit was imposed on the performance of the procedure on the simulator model.

4.25 Video Reviewer Training

Once the construction of the metrics for an arthroscopic Bankart repair was completed and face and content validity verified[65], a final version of a score sheet was formatted. 10 AANA Master / Associate Master faculty surgeons (none belonging to the experienced group from this study) formed the panel of reviewers designated to score the videos. This group included the 3 members who developed the arthroscopic Bankart *metric definitions* (table 3.1) in conjunction with a consultant experimental. The ten reviewers were assigned by the AANA research coordinator to one of five fixed pairs, which remained constant throughout the scoring of all videos. Assignments were made based on similar time zones of their residence / practice. Reviewer training was initiated with an 8-hour in-person meeting during which time each metric was studied in detail. Multiple video examples of live patient cases were shown to illustrate each particular metric. Videos of the patients in both the lateral decubitus and beach chair orientations were represented. Discussion helped to clarify how each step and error was to be scored, including the nuances and conventions to be used. Several weeks later, full-length practice videos # 1 and #2 (one each in the lateral decubitus and beach chair orientation) were sent to and independently scored by each of the 10 reviewers, and the scores then tabulated. In two subsequent 2-hour group phone conferences, the differences and discrepancies amongst all reviewers were compared and discussed seeking conformity in scoring. Each of the designated pair of reviewers also conducted one to three additional phone conferences to analyze the specific instances, in which the two of them scored particular events differently. Subsequently, all reviewers scored practice videos #3 and #4 and the results were tabulated (each patient orientation again represented). The scores for each of the 5 designated pairs of reviewers were compared for the second set of practice videos. In only 1 of 10 comparisons (2 videos X 5 reviewer pairs) did the *inter-rater reliability* (table 3.1) calculation (see below) fall below an acceptable level of 0.8[210] at 0.76.

4.26 Video Scoring

The AANA research coordinator randomly assigned the 19 full-length study videos of experienced and novice surgeons performing an arthroscopic Bankart procedure on the shoulder simulator model to a single pair of reviewers. Other than the research coordinator and the study consultant, all reviewers remained blinded to the source of all videos. Each of the 19 videos were independently reviewed and scored by the two members of an assigned pair of reviewers. All scores were tabulated for each of 13 phases of the procedure (Appendix 4.1 A and B). Each step and error metric was scored as either a “yes” or “no”, designating whether the specific event was or was not observed to occur by the reviewer. In addition to scoring steps and errors, each event characterized as “damage to non-target tissue” (DNTT) was scored. There was no limit to the number of individual instances DNTT could be scored with each occurrence simply tallied as a single error event. The score sheet also contained a box for specific reviewer comments for each metric. The 2 individual scores from a pair of reviewers were averaged to obtain the overall score for each step, error, or DNTT event. In addition, the score agreement or disagreement between the specific pair of reviewers was tabulated for each individual event (step, errors, and DNTT events) and used to calculate inter-rater reliability correlations (see below in Statistical Methods).

The total time in minutes was documented for each video beginning with the first view of the arthroscope from the posterior portal to withdrawal of the arthroscope after examination of the completed repair.

4.27 Performance Benchmark

Prior research has used the metric based mean performance of a group of experienced or expert operators to objectively define “proficiency”.[22, 79, 198, 201, 208] Prior to initiating this study, the four primary investigators specified that any subject from the experienced group who was performing > 2 standard deviations better or worse than the group as a whole would be deemed an 'outlier' and not representative of the experienced group. Any such performance by a participant of this group would have their scores removed from the analysis so as not to skew the establishment of the reference benchmark.

4.28 Statistical Methods

For each of the 13 separate phases of the procedure, the # of 'uncompleted steps' (steps which were not performed) and 'errors made' were tabulated and the scores for the 2 reviewers averaged (appendix 4.1A and B). These data were used to determine which of the procedural phases demonstrated the greatest differences in performance when comparing the experienced and novice surgeons (one factor -ANOVA analysis; IBM SPSS statistical software program). Further, for the entire procedure, the total number of steps 'completed', errors made, and sentinel errors enacted were also averaged and tabulated for the pair of reviewers.

The two raw score sheets were compared for each of the individual steps (N = 45) and the number of 'agreements' tabulated (either both reviewers documented that a step was performed, or both scored the step as not being completed). In addition, the number of 'disagreements' in scoring steps was tabulated (one of the reviewers indicated that the step had and other scored that the step had not been completed). The inter-rater reliability for the steps was calculated according to the following formula:

$$\frac{\textbf{\# Agreements}}{\textbf{\# Agreements + \# Disagreements}}$$

In a similar manner, there was either agreement or disagreement in the 2 scores for each of the potential errors (N = 77). The IRR for error scoring was calculated in the same manner as that for the steps. Finally, the IRR for scoring the metrics for the complete procedure was calculated using both the step and error agreements / disagreements for the entire procedure (N = 122). Acceptable IRR is = or > 0.80.[210]

4.3 Results

4.31 Participants

The entire group of 12 Master or Associate instructors serving as faculty for an Arthroscopy Association of North America Resident's course chose to participate and comprised the experienced group. The faculty, all fellowship trained in arthroscopy or sports medicine, averaged over 17 years in clinical practice with each having routine experience in arthroscopic shoulder techniques. All faculty members have been recognized nationally by AANA for their talent and ability to teach and communicate shoulder arthroscopy skills to trainees. The novice group was made up of 7 volunteers (from a total of 46 registered orthopedic residents) who elected to participate in the study and perform an arthroscopic Bankart repair on a shoulder model simulator.

4.32 Inter-rater Reliability (IRR) Assessments

The IRR calculations across each of the assessments were strong. The mean IRR for the paired scoring of all 19 videos for procedural steps was 0.92 (range = 0.84 – 0.98) and for errors was 0.94 (range = 0.84 – 0.99) (includes DNTT events). The mean IRR for the total of steps and errors was 0.93 (range = 0.85 – 0.97). In no instance did any of the 3 IRR calculations for the 19 scored videos fall below 0.80.

4.33 Outlier Experienced Surgeon Performance

Before analyzing the data, score profiles of those in the experienced group were examined for atypical performance. One subject in the experienced group was found to have enacted 3 sentinel errors in comparison to a mean experienced group sentinel error rate = 0.71 (standard deviation = 0.98). The mean and standard deviation was used to convert the data to Z-scores with a mean = 0 and a standard deviation = 1. The outlier subject's Z-score for sentinel errors = 2.31 (> 2 SD from the mean of the experienced group) (p = 0.01). When this subject's score is removed from the sentinel error data, the new mean = 0.5 (SD = 0.71). Figure 4.2 (A – D) show that the exclusion of the one experienced surgeon had little impact on the overall experienced group scores. In subsequent statistical analysis, all of this outlier's data were excluded.

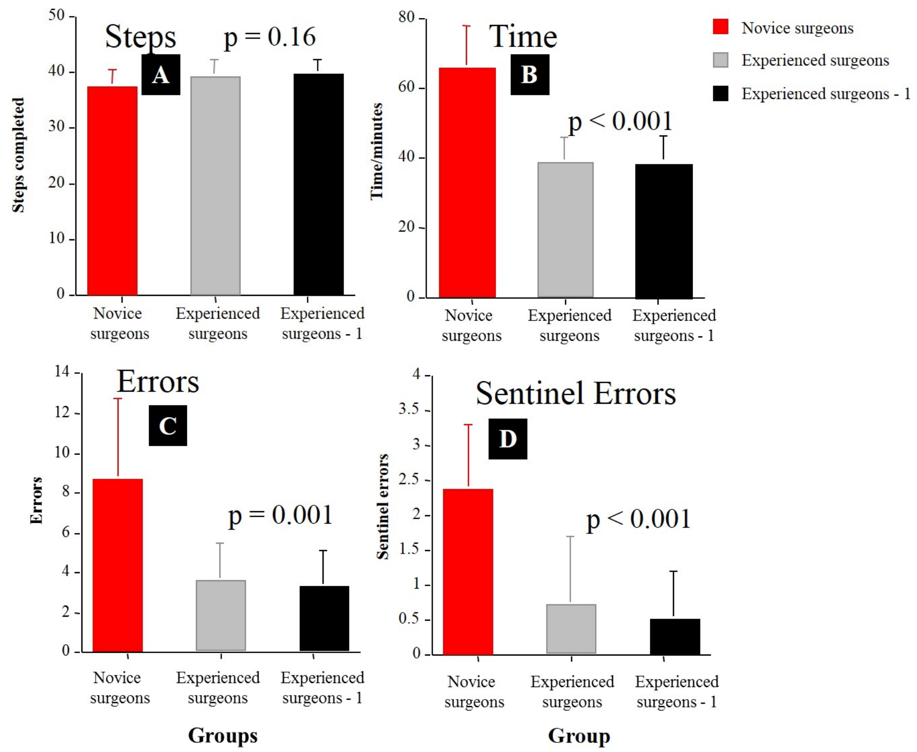


Figure 4.2A - D. Summary performance data for the shoulder model simulator showing the mean and standard deviation scores for A) steps completed, B) time taken, C) errors enacted and D) sentinel errors made by experienced and novice operators. Also shown are the mean scores of the experienced group with one outlier data set excluded.

4.34 Overall Performance Comparisons

Figure 4.2A - D shows that, on average, experienced surgeons completed more steps (figure 4.2A) than those in the novice group (39.54 vs. 37.36) but this difference was not statistically significant. The experienced group took significantly less time (Figure 4.2B) to perform the procedure on the shoulder model in comparison to the novice surgeons (39 vs. 66 minutes; $p < 0.001$). They also made significantly fewer errors (Figure 4.2C) than the novice surgeons (3.23 vs. 8.64; $p = 0.001$) and significantly fewer sentinel errors (0.5 vs. 2.36; $p < 0.001$) as shown in Figure 4.2D. Overall, experienced arthroscopists made 63% fewer errors, 79% fewer sentinel errors, and performed the procedure in 41% less time than novice surgeons. The procedural phases, which exhibited the greatest differences in performance between the groups included: anchor preparation and insertion, suture delivery and management, and knot tying. The experienced group also performed better than the novices on the phases of portal placement, arthroscopic instability assessment, and capsulolabral mobilization / gleniod preparation, but the differences were not significant.

4.35 Anchor Preparation and Insertion Steps

Figure 4.3A shows the mean number of steps not completed by both groups during the anchor preparation / insertion phase of the procedure. Few steps were not completed or omitted by either group. Experienced surgeons performed better on anchor 1 but those in the novice group performed marginally better on anchors 2 and 3. These differences were not statistically significant.

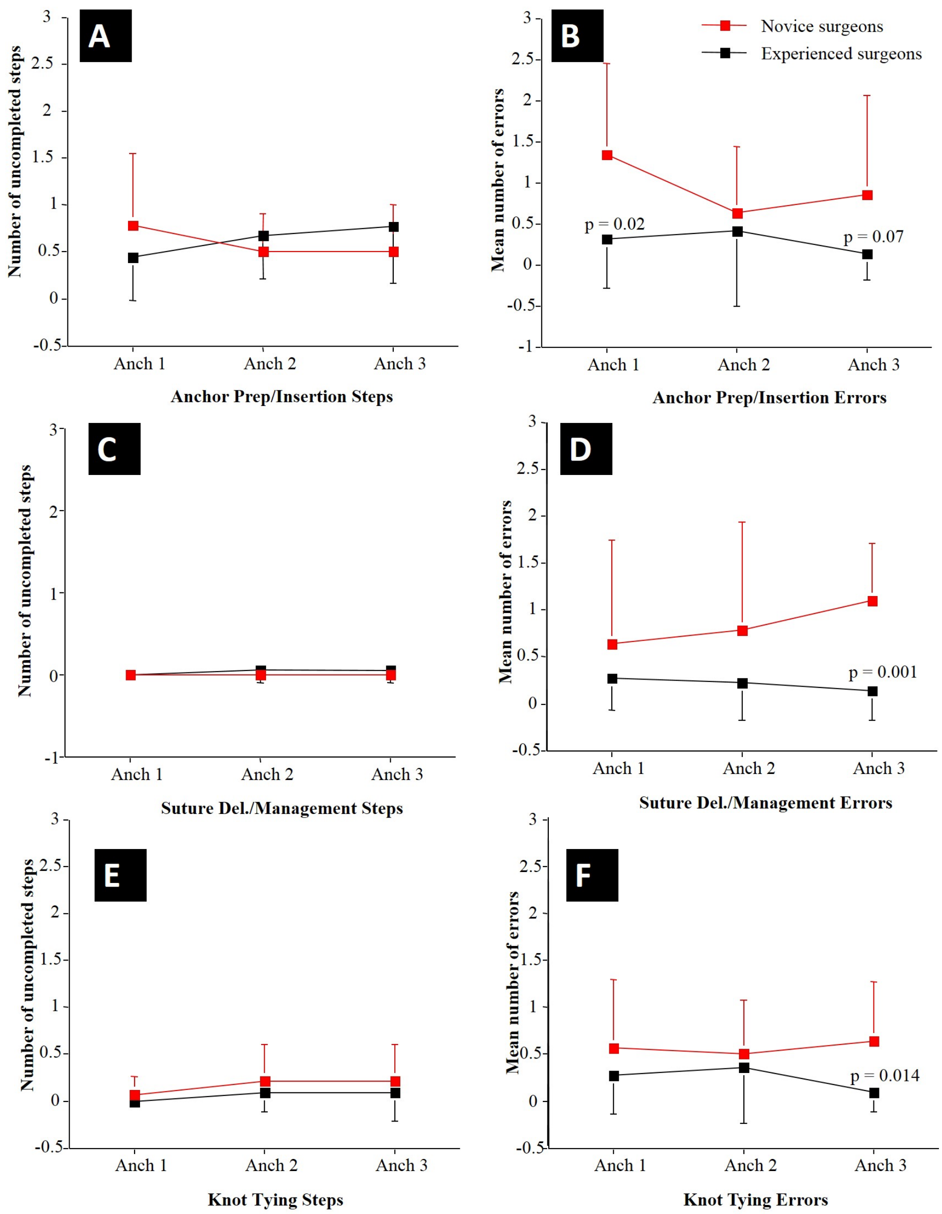


Figure 3 (A – F)

Figure 4.3A – F. Mean and standard deviation data for experienced and novice surgeons for anchor preparation and insertion A) steps, B) errors (phases IV, VIII, X); suture delivery and management C) steps, and D) errors (phases V,VIII, IX); and knot tying E) steps, and F) errors (phases VI, IX, XII) during the Bankart procedure.

4.36 Anchor Preparation and Insertion Errors

Figure 4.3B shows the mean (and standard deviation) number of errors made by the experienced group and the novice group during the anchor preparation and insertion phase of the procedure. Across all three anchors the experienced group made fewer errors than the novices and also showed more consistent performance as indicated by the smaller standard deviation scores. Experienced arthroscopists made significantly fewer errors than novices on the preparation and insertion of anchor 1 (0.32 vs. 1.36; $p = 0.02$). Although experienced surgeons also made considerably fewer errors than novices on preparation and insertion of anchor 3, this did not reach statistical significance (0.14 vs. 0.86; $p = 0.07$).

4.37 Suture Delivery and Management Steps

The number of uncompleted steps during suture delivery and management for anchors 1 – 3 was few and there were almost no differences between the groups (Figure 4.3C).

4.38 Suture Delivery and Management Errors

Figure 4.3D shows the mean (and standard deviation) number of errors made by experienced and novice surgeons on the suturing steps of the procedure. The number of errors made by experienced arthroscopists was small across all three anchors and showed substantial consistency as indicated by the small standard deviation scores. Also, their performance showed slight improvement across the anchors. In contrast, the novices showed considerable performance variability and performance deterioration across the three anchors. Only the differences in suture management and delivery on anchor 3, however, were found to be statistically significant (0.14 vs. 1.1; $p = 0.001$).

4.39 Knot Tying Steps

Similar to the suturing steps results, only a small number of uncompleted or omitted steps during knot tying for anchors 1 – 3

(Figure 4.3E) was observed for both groups and the small differences that did exists were not statistically significant.

4.310 Knot Tying Errors

The experienced group consistently made fewer errors during the knot-tying phase of the model procedure (Figure 4.3F). They performed best on anchor 3 as indicated by their mean score and very small standard deviation. On average, the novice group made less than one error per anchor. A significant difference between the groups was, however, observed for knot tying errors on anchor 3 (0.09 vs. 0.64; $p = 0.014$).

4.311 Performance Summary Assessments

The performances of both groups across the six measures presented in figures 4.3 A-F were summed to give an indicator of each group's overall performance on the three anchor repair. These data are presented in figure 4.4. Although the experienced group completed more of the procedure steps than the novice group for the anchor preparation and insertion, suture management and knot tying phases, none of these differences in steps completed were statistically significant. In contrast, all of the error variables did show large and statistically significant differences between the groups. The experienced group made significantly fewer (i.e., 70%) anchor preparation and insertion and suture management errors (0.86 vs. 2.9; $p = 0.012$). The experienced group also made 74% fewer suture delivery and management errors in comparison to the novice group, which was found to be statistically significant (0.6 vs. 2.5; $p = 0.041$). The smallest difference between the error performances of the two groups was in the knot tying phases. The experienced group still made 57% fewer knot tying errors than the novice group, which was also found to be statistically significant (0.73 vs. 1.7;, $p = 0.023$).

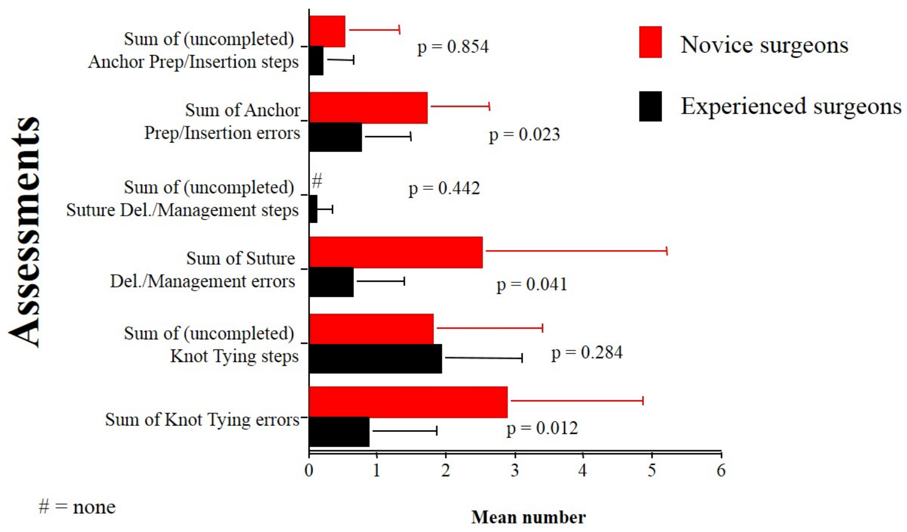


Figure 4.4 The mean and standard deviation of novices and experienced surgeons on six groups of metrics; anchor preparation and insertion steps, errors (phases IV, VIII, X); suture delivery and management steps, errors (phases V,VIII, IX); knot tying steps, errors (phases VI, IX, XII).

4.312 Procedure Review

Both groups completed the final examination of the repair and there were no errors committed during this phase of the procedure by either group.

4.313 Proficiency Benchmark

The experienced group had a mean total error rate of 3.23. A surgeon could not make a portion of an error, so for practical purposes, the error benchmark was rounded to the next greater whole number, 4. The experienced group also created a mean number of 0.5 sentinel errors (standard deviation = 0.71), rounded to 1. Thus, the overall benchmark is set at completing a 3 anchor Bankart repair with no more than 4 total errors and no more than 1 sentinel error.

4.4 Discussion

The most important findings of this study include verification that the arthroscopic Bankart metrics coupled with a shoulder model simulator are able to accurately discriminate between the performance of experienced and novice orthopedic surgeons and demonstrate *construct validity* (table 3.1) In addition, a performance benchmark was able to be established based on the mean performance of the experienced group and included: completion of a 3 anchor Bankart repair, enacting no more than 4 total errors and 1 sentinel error.

4.41 Bankart Metrics

The primary intent of the study was to determine whether construct validity could be demonstrated for the previously established arthroscopic Bankart metrics coupled with the use of a shoulder model simulator. For construct validity to be demonstrated, the combination of metrics and simulator tools must be able to discriminate between the performance of experienced and novice surgeons. The differences between the two groups were significant and those that best distinguished between experienced and novice surgeons in the performance of an arthroscopic Bankart procedure included: 1) the errors enacted in the performance of the procedure, 2) the number of sentinel errors made, and 3) the time it took to perform the procedure.

In this study, error scores were a very powerful and accurate[211] discriminator between the groups with the novices making more than twice as many errors as the experienced group with a difference in the standard deviation scores of a similar magnitude. The goal of surgical education should be to help trainees perform well with as few errors as possible. The trainee, however, should be afforded the opportunity to create errors in an inconsequential manner (i.e. on a simulator without associated patient morbidity) and learn from them. Effective progression in training should be demonstrable with a concomitant reduction in errors. For the individual trainee, the identification of specific errors facilitates a focused correction of deficiencies. Performance errors can also be used as a powerful metric tool to shape and configure the related educational curriculums and remedial training as well as to establish benchmarks, which trainees must meet and demonstrate before progressing.[79, 201, 212] Lastly, defined errors can serve to guide the development of simulators, i.e., not only what they should emulate, but also, what they should measure.

Within surgical and procedural disciplines, it is unanimously agreed that certain types of technique errors are so egregious and pose such a threat to either the success of the procedure or patient safety that they should constitute their own performance category. We have elected to term these more serious deviations from optimal performance, 'sentinel errors'. Similar designations have not been formally made in other studies that have objectively assessed surgical performance and evaluated the construct validity of specific metrics.[22, 70, 198, 207, 208, 213] The use of such a special metric classification could have profound implications for "high stakes" assessments (i.e. - determining whether a trainee is allowed to progress in the specific educational / residency program) and proficiency based progression approaches to training

Although the # steps completed did not distinguish between the groups in this study, it is none-the-less important to include all of the essential steps in a training program. The procedure cannot be completed without knowing and performing all of the correct steps in the proper order.[79, 201, 212] The fact that this performance unit is being assessed also increases the probability that the steps will be learned.[214] However, the proper steps and sequences should be communicated and learned outside of the skills training proper, i.e., in an online educational module, as its inclusion is sensible rather than essential. A thorough diagnostic evaluation of pathology potentially related to shoulder instability is necessary for the comprehensive and appropriate treatment of the unstable shoulder, and thus, its inclusion in the steps of the procedure.

4.42 Simulator Model

Gallagher and O'Sullivan[79] have proposed that to be effective, a simulation model should provide the learner with the span of appropriate sensory responses to physical actions that are behaviorally consistent with what would be experienced in the real situation (including the opportunity to enact both appropriate (steps) and inappropriate actions (errors). The simulator should also afford the opportunity to execute the procedure in the same order and with the same tools and devices with which the procedure would normally be performed.[212] The simulator model used in this study was sufficiently realistic to provide the learner the opportunity to perform each of the 45 steps in a realistic fashion using the same tools, implants and techniques employed for an anterior stabilization on a real patient.

4.43 Benchmarking

The second purpose of this study was to establish a performance benchmark for the arthroscopic Bankart metrics coupled with a shoulder model simulator. The definition of 'proficiency', in distinction to qualitatively described 'competency', is based on objectively defined performance metrics. Proficiency based progression (PBP) training requires the establishment of a benchmark to which trainees must perform to be able to progress. We sought to establish an objective, reliable, transparent, and fair performance benchmark for an anterior Bankart repair on the shoulder model simulator. As the benchmark is established based on the mean performance of the group of experienced surgeons, it was important that the performance of the members be representative of that group. Based on a pre-study stipulation, the scores of one member of the experienced group were removed from the analysis, as their performance was 2.3 SD worse than their peers. This policy was established so as not to skew the creation of the reference benchmark and was demonstrated to have little impact on the overall scores of the experienced group.

An IRR > or greater than 0.80 is considered acceptable.[210] The very high IRR for the scores from reviewer pairs for the entire group of metrics (0.93) is reflective of the clarity and precision of the arthroscopic Bankart metrics drafted, and the thorough training of the 10 reviewers. The ability to score the steps and errors consistently is essential in obtaining a reliable measure of the surgeon's performance and skill level for a particular procedure.

4.44 Limitations

A limitation of this study is that there was no confirmation that those serving as Master / Associate Master surgeons and representative of the experienced group possessed a specified level of expert skill in performing an arthroscopic Bankart procedure. Nevertheless, the individual surgeons so identified, have been recognized by the Arthroscopy Association of North America as skilled and effective educators either from lecture presentations with video exhibiting skilled shoulder arthroscopic techniques or from their performance in an arthroscopic lab setting in which they demonstrated the ability to teach each of the key components of a Bankart procedure. Therefore, this group was defined as 'experienced' rather than 'expert'. Similarly, other than identifying the year in training, no information was obtained to determine the extent of the novice group's (resident) experience with arthroscopic shoulder surgery, i.e., number of arthroscopy / sports medicine rotations previously completed or the number of shoulder arthroscopic procedures in which they served as an assistant surgeon. Even with those pieces of information, it would not be possible to gain any reliable measure of an individual resident's level of first hand 'experience' or skill with arthroscopic shoulder surgery. Furthermore, the structure of residency rotations and level of independence permitted varies a great deal among training programs. As a result, the residents' knowledge and skill sets are unlikely to be uniform but the PGY 4 and 5 levels provide a general measure of their training experience.

The numbers of surgeons in both groups were small but one of the strengths of the detailed metric-based procedure characterization method that we used is the sensitivity to detecting differences when in fact they exist. One hundred twenty three data points (metrics) including the duration of the procedure in minutes were obtained for each scored video. Thus, small numbers of subjects can still produce statistically powerful differences, assuming that performance has been reliably measured. An 'a priori' power analysis was not performed, as we had no previous studies in arthroscopic shoulder repair that used a similar detailed assessment methodology. This is the first study of its kind in this field and our results will afford other researchers the opportunity to develop their sample sizes based on reported mean and standard deviation scores. The only published scientific reports we could draw on were similar type studies published in the laparoscopic surgical literature. Those reports could only give an indication of the possible sample sizes required.

Although not specifically a limitation of the study or design, the option to use either the lateral decubitus or beach chair position could potentially introduce some variability. Both patient positions are in common use among practicing surgeons. The metrics were carefully constructed to facilitate unbiased scoring for the model simulator / patient in either orientation with no penalty. Several metrics require that the arthroscope be placed in the anterosuperior portal to adequately complete the step, however, this is true for both orientations (i.e. step #12 – “view or probe the insertion of the anterior glenohumeral ligaments onto the anterior humeral neck”, and step #16 – “obtain a view of the anterior glenoid neck”). The challenge in completing these steps relates more to the position of the arthroscope (posterior vs. anterosuperior portal) than the patient orientation. While the authors believe that the lateral decubitus orientation makes some steps easier to perform, (i.e. appropriate seating of the drill guide in the anterior /posterior dimension relative to the bony rim, and accurate passage of the suturing device through the capsulolabral tissue inferior to the anchor site), no inherent bias is introduced in scoring the metrics for procedures performed in either patient orientation.

Based on the data from this study, the null hypothesis is rejected. The shoulder model employed, coupled with previously validated arthroscopic Bankart metrics, is able to accurately distinguish between experienced and novice operators. Construct validity is demonstrated for the simulator model plus performance.

4.45 Conclusions

The tool comprised of validated arthroscopic Bankart repair metrics coupled with a dry shoulder model simulator is able to accurately distinguish between the performance of experienced and novice orthopedic surgeons. A performance benchmark based on the experienced group includes: completion of a 3 anchor Bankart repair, and enacting no more than 4 total errors and 1 sentinel error.

Table 4.1 The Arthroscopic Instruments used to perform the Bankart procedure on the cadaver shoulder.

5.5- and 8.5-mm obturator cannulas
Switching sticks
Hook probe
Regular and looped graspers
Liberator/elevator
Shaver
Drill guide/drill
Push-in anchor loaded with single suture
Mallet
Cannulated suture hook
Penetrator
Monofilament suture
Knot pusher
Arthroscopic scissors

Appendix 4A – Model Metric Validation: Novice

	64A	64B	64 Ave	94A	94B	94 Ave	24A	24B	24 Ave	34A	34B	34 Ave	14A	14B	14 Ave	44A	44B	44 Ave	84A	84B	84 Ave
Video																					
I. Portals																					
Steps uncompleted	0	0	0	0	0	0	0	0	0	3	2	2.5	2	3	2.5	1	2	1.5	1	2	1.5
Errors made	0	2	1	3	3	3	0	1	0.5	3	0	1.5	1	1	1	1	1	1	2	0	1
II. Instabl assess																					
Steps uncompleted	2	1	1.5	1	0	0.5	1	1	1	1	1	1	3	3	3	0	1	0.5	2	1	1.5
Errors made	0	0	0	1	1	1	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0
III. Caps/gen prep																					
Steps uncompleted	2	2	2	3	3	3	1	1	1	3	4	3.5	5	5	5	3	3	3	3	4	3.5
Errors made	1	0	0.5	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IV. First inf anch prep																					
Steps uncompleted	0	0	0	0	1	0.5	0	1	0.5	0	0	0	1	2	1.5	1	3	2	1	1	1
Errors made	1	0	0.5	2	1	1.5	3	3	3	1	4	2.5	1	2	1.5	0	1	0.5	0	0	0
V. First sut del/mgmt																					
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0	1	1	1	3	3	3	0	0	0	0	0	0	1	0	0.5
VI. First knot tying																					
Steps uncompleted	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	1	0.5	0	0	0	0	0	0	2	2	2	1	1	1	1	0	0.5
VII. Second anch prep																					
Steps uncompleted	0	0	0	0	1	0.5	1	0	0.5	0	0	0	1	1	1	1	1	1	1	0	0.5
Errors made	0	0	0	2	1	1.5	0	1	0.5	2	2	2	1	0	0.5	0	0	0	0	0	0
VIII. Second sut del/mgmt																					
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	2	1	1.5	0	0	0	1	1	1	3	3	3	0	0	0	0	0	0	0	0	0
IX. Second knot tying																					
Steps uncompleted	1	0	0.5	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
Errors made	2	1	1.5	0	1	0.5	0	0	0	1	1	1	0	0	0	0	0	0	1	0	0.5
X. Third anch prep																					
Steps uncompleted	0	0	0	0	1	0.5	1	1	1	0	0	0	1	1	1	1	1	1	0	0	0
Errors made	0	0	0	0	1	0.5	0	0	0	2	2	2	0	0	0	1	0	0.5	3	3	3
XI. Third sut del/mgmt																					
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	1	2	1.5	0	0	0	1	1	1	2	2	2	1	1	1	1	1	1	2	0	1
XII. Third knot tying																					
Steps uncompleted	0	1	0.5	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
Errors made	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	3	0	1.5
XIII. Eval repair																					
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total time (Port, Dx, and Rx), min			105			93			67			76			62			67			39
Rating pairs																					
Steps completed (45)	40	41	40.5	40	39	39.5	41	41	41	36	36	36	32	31	31.5	38	34	36	37	37	37
Errors made (77)	7	6	6.5	10	11	10.5	8	6	7	17	17	17	7	6	6.5	5	5	5	13	3	8

(continued)

	64A	64B	64 Ave	94A	94B	94 Ave	24A	24B	24 Ave	34A	34B	34 Ave	14A	14B	14 Ave	44A	44B	44 Ave	84A	84B	84 Ave
Sentinel errors	3	2	2.5	1	4	2.5	3	3	3	3	4	3.5	1	3	2	0	1	0.5	4	1	2.5
Steps completed																					
Agreement			42			40			43			43			41			41			39
Disagreement			3			5			2			2			4			4			6
Step IRR			0.93			0.89			0.95			0.95			0.91			0.91			0.87
Errors made																					
Agreement			73			70			72			69			70			73			65
Disagreement			4			7			5			8			7			4			12
Error IRR			0.95			0.91			0.93			0.9			0.91			0.95			0.84
Total score (S + E)																					
Agreement			115			110			115			112			111			113			104
Disagreement			7			12			7			10			11			9			18
Total IRR			0.94			0.9			0.94			0.92			0.91			0.93			0.85

anch, anchor; caps, capsule; Dx, diagnostic; E, errors; eval, evaluate; glen, glenoid; inf, inferior; instab, instability; IRR, inter-rater reliability; mgmt, management; min, minutes; port, portals; Rx, treatment; S, steps; sut, suture; sut del, suture delivery.

Appendix 4B – Model Metric Validation: Experienced

	Excluded																																			
	15A	15B	15 Ave	25A	25B	25 Ave	35A	35B	35 Ave	65A	65B	65 Ave	75A	75B	75 Ave	45A	45B	45 Ave	55A	55B	55 Ave	85A	85B	85 Ave	95A	95B	95 Ave	105A	105B	105 Ave	115A	115B	115 Ave	125A	125 B	125 Ave
Video																																				
I. Portals																																				
Steps uncompleted	3	2	2.5	1	0	0.5	2	1	1.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	1.5	2	1	1.5
Errors made	0	1	1.5	0	0	0	1	0	0.5	1	0	0.5	0	1	0.5	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	3	2	2.5	0	0	0
II. Instabl assess																																				
Steps uncompleted	3	2	2.5	1	1	1	1	2	1.5	1	0	0.5	2	1	1.5	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.5	2	0	1	2	1	1.5
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
III. Caps/gen prep																																				
Steps uncompleted	2	3	2.5	2	1	1.5	4	4	4	1	3	2	2	3	2.5	2	2	2	1	3	2	1	2	1.5	1	3	2	1	3	2	4	3	3.5	3	3	3
Errors made	0	0	0	0	0	0	0	1	0.5	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
IV. First inf anch prep																																				
Steps uncompleted	1	1	1	0	0	0	0	1	0.5	0	1	0.5	0	0	0	1	0	0.5	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1
Errors made	2	2	2	0	1	0.5	2	2	2	1	0	0.5	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	2	2	2	0	0	0	0	0	0
V. First sut del/mgmt																																				
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	2	0	1	1	0	0.5	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	1	0.5	0	1	0.5	1	0	0.5	0	0	0
VI. First knot tying																																				
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	0	1	0.5	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0
VII. Second anch prep																																				
Steps uncompleted	1	1	1	0	0	0	1	1	1	0	1	0.5	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	0	2	1	1	1	1
Errors made	0	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	3	1	0	0.5
VIII. Second sut del/mgmt																																				
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0
Errors made	1	0	0.5	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
IX. Second knot tying																																				
Steps uncompleted	1	0	0.5	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.5	0	0	0	0	0	0
Errors made	0	0	0	0	0	0	1	1	1	0	0	0	0	1	0.5	1	2	1.5	0	0	0	0	0	0	0	0	0	0	0	0	2	1	1.5	0	1	0.5
X. Third anch prep																																				
Steps uncompleted	1	1	1	2	0	1	1	1	1	1	1	1	0	0	0	1	3	2	1	1	1	1	1	1	0	0	0	0	0	0	0	1	0.5	1	1	1
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
XI. Third sut del/mgmt																																				
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0
Errors made	1	0	0.5	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
XII. Third knot tying																																				
Steps uncompleted	1	1	1	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	1	0	0.5	0	0	0	1	1	1	0	0	0	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
XIII. Eval repair																																				
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total time (Port, Dx, and Rx), min			38			47			37			37			29			43			43			52			42			31			35			27
Rating pairs																																				
Steps completed (45)	33	34	33.5	41	43	42	36	33	34.5	42	39	40.5	42	41	41.5	40	42	41	39	42	40.5	41	40	40.5	42	42	42	43	39	41	36	35	35.5	37	37	37
Errors made (77)	5	3	4	3	1	2	7	6	6.5	3	2	2.5	4	6	5	3	4	3.5	1	1	1	3	1	2	3	4	3.5	3	4	3.5	8	7	7.5	1	1	1
Sentinel errors	1	0	0.5	0	0	0	3	3	3	0	0	0	1	1	1	0	3	1.5	0	0	0	0	0	0	0	0	0	0	0	0	2	2	2	1	0	0.5
Steps completed																																				
Agreement			42			39			38			40			43			42			43			44			41			41			38			41
Disagreement			3			6			7			5			2			3			2			1			4			4			7			4
Step IRR			0.93			0.87			0.84			0.89			0.95			0.93			0.95			0.98			0.91			0.91			0.84			0.91
Errors made																																				
Agreement			73			73			73			75			75			73			75			75			76			76			72			75
Disagreement			4			4			4			2			2			4			2			2			1			1			5			2
Error IRR			0.95			0.95			0.95			0.97			0.97			0.95			0.97			0.97			0.99			0.99			0.93			0.97
Total score (S + E)																																				
Agreement			115			112			111			115			118			115			118			119			117			117			111			116
Disagreement			7			10			11			7			4			7			4			3			5			5			11			6
Total IRR			0.94			0.92			0.91			0.94			0.97			0.94			0.97			0.97			0.96			0.96			0.91			0.95

(continued)

anch, anchor; caps, capsule; Dx, diagnostic; E, errors; eval, evaluate; glen, glenoid; inf, inferior; instab, instability; IRR, inter-rater reliability; min, minutes; mgmt, management; port, portals; Rx, treatment; S, steps; sut, suture; sut del, suture delivery.

5. "Bankart Performance Metrics Combined With a Cadaveric Shoulder Create a Precise and Accurate Assessment Tool for Measuring Surgeon Skill" Angelo R L, Ryu R K N, Pedowitz R A, Gallagher A G. Arthroscopy 2015;32:1655-1670. Appendix 7

(Candidate is the first and primary author)

Purpose: To determine if previously validated performance metrics for an arthroscopic Bankart repair (ABR) coupled with a cadaver shoulder are a valid assessment tool with the ability to discriminate between the performances of experienced and novice surgeons; To establish a proficiency benchmark for an ABR using a cadaver shoulder.

Methods: 10 Master / Associate Master Faculty from an Arthroscopy Association of North America Resident course (experienced group) were compared with 12 PGY 4 and PGY 5 orthopedic residents (novice group). Each group was instructed to perform a diagnostic arthroscopy and a 3 suture anchor Bankart repair on a cadaver shoulder. The procedure was videotaped in its entirety and independently scored in blinded fashion by a pair of trained reviewers. Scoring was based on defined and previously validated metrics for an arthroscopic Bankart and included steps, errors, 'sentinel' (more serious) errors, and time.

Results: The inter-rater reliability = 0.92. Novice surgeons made 54% more errors (5.68 vs. 2.61; $p = 0.026$), showed more performance variability ($SD = 3.5$ vs. 1.6) and took longer to perform the procedure (45.5 vs. 29.5 minutes; $p < 0.001$). The greatest difference in errors related to suture delivery and management (exclusive of knot tying)(1.95 vs. 0.33; $p = 0.02$).

Conclusions: The assessment tool comprised of validated arthroscopic Bankart metrics coupled with a cadaver shoulder accurately distinguishes the performance of experienced from novice orthopedic surgeons. A benchmark based on the mean performance of the experienced group includes: completion of a 3 anchor Bankart repair, and enacting no more than 3 total errors and 1 sentinel error.

Clinical Relevance: Validated procedural metrics combined with the use of a cadaver shoulder can be used to assess the performance of an arthroscopic Bankart repair. The methodology employed may serve as a template for outcomes based procedural skills training in general.

5.1 Introduction

The traditional manner in which surgical trainees have acquired their operative skills is under considerable pressure. Concerns about patient safety,[215, 216] pressures on operating room efficiency[7] and the reduced availability of work hours[217, 218] have resulted in fewer opportunities for in vivo operative experience. As a consequence, trainees are graduating from residency programs with considerably less operative experience and almost certainly less technical skill than residents graduating in the past who were exposed to greater surgical volumes. For example, Bell et al.,[219] found that of the 121 surgical procedures that general surgery residency program directors believed residents should be competent in by the time of graduation, only 18 of them had been performed with sufficient frequency by the resident for them to acquire competence during their training. They also found that the mode frequency with which the 121 procedures were performed was 0. The implications of these findings for surgical training are considerable and concerning. At a more practical level it means that surgical skills training must be optimized and preparation for a surgical practice maximized.

Traditionally, surgical residents have been trained using the 'apprenticeship' model, dependent in part, on exposure to surgical cases, variable graduated participation in surgery, and time spent on specific clinical rotations. At the outset, we sought to determine if a *'proficiency based progression'* (PBP)(table 3.1 – glossary) method was potentially a more effective manner in which to train surgical skills than the apprenticeship model. A PBP training program dictates that the trainee demonstrates that skill performance, to a pre-determined benchmark level, before advancing to more complex techniques. This method relies on a comprehensive and quantitative characterization of the skills to be learned. These performance characteristics or *metrics* (table 3.1) and their *operational definitions* (table 3.1) (rather than descriptions) offer very specific goals and guidelines for the training curriculum. Previously, we reported[65] on the development of *performance metrics* (*steps* and *errors*)(table 3.1) for a standard reference approach to performing an arthroscopic Bankart repair (ABR).[185, 190, 191, 193, 209] Those metrics were derived from a careful *task analysis* and *deconstruction* (table 3.1) using videos of complete Bankart procedures performed with patients in either the lateral decubitus or beach chair orientation. The metrics were constructed in such a manner that they could be scored in an identical manner with the patient in either orientation. *Face* and *content validity* (table 3.1) of the metrics were verified using a modified *Delphi Panel* (table 3.1) methodology. The Delphi Panel was composed of 27 experienced shoulder arthroscopists who have all served as Master or Associate Master faculty for Arthroscopy Association of North America (AANA) shoulder courses at the Orthopedic Learning Center (Rosemont, Il). The Delphi Panel obtained excellent consensus on the metric-based characterization of the Bankart procedure.

In a subsequent report, we verified *construct validity* (the ability to discriminate between the performance of experienced and novice groups of surgeons)(table 3.1) for the use of the ABR metrics with a shoulder model simulator as a training tool.[93] In the present study we evaluate the construct validity of these exact same metrics on a much higher fidelity platform, the human cadaver shoulder. At present, an arthroscopic Bankart repair in a cadaver shoulder provides the closest approximation to a similar surgical repair in a live patient. Full physics, high fidelity virtual reality simulators with haptic feedback are likely to play a greater role in the future, but are expensive to develop and not currently available. Even with those simulators, validated metrics will be needed to substantiate their effectiveness.

The purpose of this study was to determine if previously validated performance metrics for an arthroscopic Bankart repair (ABR) coupled with a cadaver shoulder are a valid assessment tool with the ability to discriminate between the performances of experienced and novice surgeons. We also sought to establish a *proficiency* (table 3.1) benchmark for that procedure using the cadaver shoulder. The null hypothesis states that when using a cadaver shoulder, the Bankart metrics would fail to discriminate between experienced and novice surgeon performance.

5.2 Methods

No IRB was obtained for this study investigating the validity of the Bankart metrics coupled with the cadaver shoulder. An IRB was sought for the final Copernicus Study proper, which will compare 3 different training protocols evaluating surgical simulation and proficiency based training methods. The Western Institutional Review Board (WIRB) (#1-776362-1) opined that, as an educational curriculum study, it was "Exempt" from the need for full IRB approval (criteria: 45 CFR 46.101(b)(1)). The final study comparing the 3 training protocols was registered with the NIH (ClinicalTrials.gov: #NCT01921621).

5.21 Study Groups

Two groups were compared in their performance of an arthroscopic Bankart procedure on a cadaver shoulder. The experienced group consisted of all faculty members who served as Master and Associate Master instructors for a standard 3 day Arthroscopy Association of North America (AANA) Resident Course conducted at the Orthopedic

Learning Center (OLC, Rosemont, Il.). 'Experienced' meant that they performed the procedure consistently in practice and taught the principles at the OLC during shoulder courses. An 'expert' would not be possible to define without the surgeon meeting objective performance criteria (metrics) that achieved a specific benchmark (that some group or body determined, meant 'expert' performance). The novice group was limited to PGY 4 and PGY 5 orthopedic residents who had registered for a Resident's Course and who volunteered to participate in the investigation.

5.22 Arthroscopic Bankart Repair Metrics

Metrics have been previously defined for a standard reference arthroscopic Bankart repair[65][7]. 45 essential steps in 13 *procedural phases* (Roman numerals) were defined with beginning and end points (Table 3.3). 29 potential unique errors were specified, 8 of which were designated as *'sentinel'* (Table 3.4). The more serious sentinel errors were defined as those expected to either; 1) substantially compromise the outcome of the shoulder stabilization (i.e. – 'capsular penetration of the suture passing instrument is superior to the anchor hole' resulting in failure to retention the capsule and inferior glenohumeral ligaments), or 2) potentially lead to iatrogenic damage to the shoulder (i.e. 'laceration of the intact labrum'). Some of the same errors could be enacted more than once during separate but similar phases of the procedure (i.e. suture delivery and management for each of 3 anchors). Thus, a total of 77 potential errors, 20 of which were sentinel errors, were specified for the complete procedure. In addition, events that led to less consequential *"damage to non-target tissues"* (DNTT) (table 3.1) were recorded as a standard error (i.e. scuffing of the articular cartilage). A perfect score would indicate that all 45 steps were completed satisfactorily without committing any errors.

There are many ways to perform a Bankart procedure, but the task deconstruction was designed for a 'reference' procedure (a routine Bankart repair) breaking it down into essential components. Each of the metrics were specifically crafted to accommodate different methods that can be used to accomplish the steps, i.e. suture passage could be performed with a number of different instruments and techniques, but to accomplish re-tensioning of the capsulolabral tissue, the capsule must be purchased inferior to the anchor site (one of the metrics). The modified Delphi panel procedure used to obtain face and content validity, asks the question of each panel member, "is this metric (step or error) acceptable as written”, i.e. 'it is not incorrect' (although a particular panel member might perform the step in a different manner). The 45 steps and 77 errors were drafted, revised, and stress tested by the core group of primary investigators and then submitted to the Delphi Panel for comment, modification, and revision. The panel then obtained consensus for all of the 122 metrics, which were found to be acceptable in their final form.

5.23 Cadaver Shoulder Study Specimens

Fresh frozen cadaveric specimens with a complete shoulder girdle from the scapula and associated soft tissues to the mid-humerus were used. After appropriate thawing, the scapula was mounted with a clamp in the subject surgeon’s orientation of preference (lateral decubitus vs. beach chair). The cadaver specimens were considered acceptable if: 1) arthroscopic visibility of the target tissues was obtainable; 2) the specimen (flexibility) permitted adequate access to the target tissues; and 3) the integrity of the capsulolabral tissues was sufficient to permit mobilization, suture delivery and knot-tying. One of 3 designated AANA shoulder arthroscopy Master Instructors (the surgeon members of the group who created the arthroscopic Bankart metrics) determined the acceptability of the cadaver specimens.

5.24 Arthroscopic Bankart Repair

During a single weekend AANA resident arthroscopy course, the surgeons from both groups were instructed to establish portals (posterior, anterosuperior, and mid-anterior), complete a thorough diagnostic arthroscopy, and perform a 3 anchor arthroscopic Bankart repair on the cadaver shoulder. Further, they were instructed to demonstrate / complete all of the steps for the Bankart repair that they would normally perform in clinical practice on a real patient. Equipment representatives from multiple different vendors served as surgical assistants and were randomly assigned to participating surgeons. The assistants were instructed to act only at the specific direction of the operating surgeon. Prompting and coaching (of technique) were prohibited (the procedures were proctored by staff from the Orthopedic Learning Center). A standard equipment tower with a 30^0 arthroscope was provided along with all instruments necessary to complete an arthroscopic Bankart repair (Table 4.1).

The subject surgeon identified bony landmarks with a marking pen, established their desired portals and performed a diagnostic examination. The arthroscope was then withdrawn from the shoulder joint. One of the 3 Master surgeons who evaluated the cadaver acceptability according to the criteria noted above, then reintroduced the arthroscope and created a standardized Bankart lesion from 2 – 6 o'clock, and 6 – 9 mm deep (medial). While the lesion was clearly delineated, the capsulolabral tissues were not mobilized. Prior to the study, we attempted to employ different techniques to create an anteroinferior capsular detachment from the glenoid (Bankart lesion) in a cadaver shoulder specimen (i.e. a long-handled scalpel blade, a hook tip cautery, and manual dissection with an elevator). It was determined that the most consistent lesion could be created using a liberator/elevator along with a mallet to provide a gentle, controlled impact force to the elevator to carefully "sculpt" the Bankart pathology. This method optimized preservation of the integrity of the capsulolabral tissues for subsequent repair.

Once the Bankart lesion was completed, the arthroscope was withdrawn and reintroduced by the subject surgeon who operated for the duration of the procedure. A continuous video recording was made beginning with the first arthroscopic view of the joint from the posterior portal, ending with the final examination of the completed procedure by the surgeon. In calculating the total operating time of the Bankart repair procedure for the subject surgeon, the segment of time required by the master faculty surgeon to create the Bankart pathology was subtracted from the total absolute running time. No time limit was imposed on the performance of the procedure in the cadaver specimen.

5.25 Video Reviewer Training

Once the construction of the metrics for an arthroscopic Bankart repair was completed and face and content validity verified[65], a final version of a score sheet was formatted. 10 AANA Master / Associate Master faculty surgeons (none belonging to the experienced group from this study) formed the panel of reviewers designated to score the videos. This group included the 3 members from the group who, in conjunction with a consultant experimental psychologist, created the arthroscopic Bankart metric *definitions* (table 3.1). The ten reviewers were randomly assigned to form five fixed pairs, which remained constant throughout the scoring of all videos. Reviewer training was initiated with an 8-hour in-person meeting during which time each metric was studied in detail. Multiple video examples of live patient cases were shown to illustrate each particular metric. Videos of the patients in both the lateral decubitus and beach chair orientations were represented. Discussion helped to clarify how each step and error was to be scored, including the nuances and conventions to be used. Several weeks later, full-length practice videos # 1 and 2 (one each in the lateral decubitus and beach chair orientation) were sent to and independently scored by each of the 10 reviewers, and the scores tabulated. In two subsequent 2-hour group phone conferences, the differences and discrepancies amongst all reviewers were compared and discussed seeking conformity in scoring. In addition, each designated pair of reviewers conducted one to three additional phone conferences to analyze the specific instances in which the two of them scored particular events differently. Subsequently, all reviewers scored practice videos #3 and #4 and the results were again tabulated (each patient orientation represented). The scores for each of the 5 designated pairs of reviewers were compared for the second set of practice videos. In only 1 of 10 comparisons (2 videos X 5 reviewer pairs) did the *inter-rater reliability (IRR)*(table 3.1) calculation (see below) fall below an acceptable level of 0.8[210] at 0.76.

5.26 Video Scoring

The AANA research coordinator randomly assigned each of the 22 full-length study videos of the experienced and novice surgeons performing an arthroscopic Bankart repair on a cadaver shoulder to a single pair of reviewers. Other than the research coordinator and the study consultant, the primary investigators and all video reviewers remained blinded to the source of the video being reviewed. Each of the 22 videos were independently reviewed and scored by the two members of an assigned pair of reviewers. All scores were tabulated for each of 13 phases of the procedure (Appendix 5 A and B). Each step and error metric was scored as either a "yes" or "no", designating whether the specific event was or was not observed to occur by the reviewer. In addition to scoring steps and errors, each event characterized as "damage to non-target tissue" (DNTT) was scored. There was no limit to the number of individual instances DNTT could be scored with each occurrence simply tallied as a single error event. The score sheet also contained a box for specific reviewer comments for each metric. The 2 individual scores from a pair of reviewers were averaged to obtain the overall score for each step, error, or DNTT event. In addition, the score agreement or disagreement between the specific pair of reviewers was tabulated for each individual event (step, errors, and DNTT events) and used to calculate inter-rater reliability correlations (see below in Statistical Methods).

5.27 Performance Benchmark

Prior research has used the mean performance (metric based) of a group of experienced or expert operators to objectively define proficiency.[22, 79, 198, 201, 208] To assess and ensure performance homogeneity amongst the experienced surgeon group for establishment of an accurate benchmark, their performances were converted to Z-scores. The standard score (more commonly referred to as a Z-score) is a very useful statistic because it, a) creates the ability to calculate the probability of a score occurring within a normal distribution and b) enables one to compare more precisely the scores from two individuals on a standardized scale, i.e. to objectively and transparently determine whether and precisely how much a given subject's score is above or below the mean of their peers. In the pre-study design phase of the project, a stipulation was made by the primary investigators to remove the data from analysis for an experienced surgeon performing > 2 standard deviations from the mean of the experienced group for any of the 4 assessments: steps, errors, sentinel errors, and time. Any such performance by a participant of experienced group would be deemed an 'outlier' and the scores would be removed from the analysis so as not to skew the establishment of the reference benchmark. This stipulation applied to both the previously reported shoulder model[93] construct validity study and the cadaver construct validity study in this report.

5.28 Incomplete Repairs

It was prospectively determined that if a surgeon did not substantially complete the 3 anchor Bankart repair, their partial scores would be removed from the analysis. This policy was established due to the fact that if only a portion of a procedure was performed, it would not have been possible to accurately estimate or extrapolate how many errors, sentinel errors, or instances of damage to non-target tissues they may have enacted had they completed the entire procedure. Further, no estimate of total time for the procedure would be sensible.

5.29 Statistical Methods

For each of the 13 separate phases of the procedure, the # of 'uncompleted steps' and 'errors made' were tabulated and the scores for the 2 reviewers averaged (appendix 5A and B). These data were used to determine which of the procedural phases demonstrated the greatest differences in performance when comparing the experienced and novice surgeons (one factor -ANOVA analysis; IBM SPSS statistical software program). Further, for the entire procedure, the total number of steps completed, errors made, and sentinel errors enacted were also averaged for the pair of reviewers.

The two raw score sheets were compared for each of the individual steps (N = 45) and the number of 'agreements' tabulated (either both reviewers documented that a step was performed, or both scored the step as not being completed). In addition, the number of 'disagreements' in scoring steps was tabulated (one of the reviewers indicated that the step had and other scored that the step had not been completed). The inter-rater reliability for the steps was calculated according to the following formula:

$$\frac{\textbf{Agreements}}{\textbf{Agreements + Disagreements}}$$

In a similar manner, there was either agreement or disagreement in the 2 scores for each of the potential errors (N = 77). The IRR for error scoring was calculated in the same manner as that for the steps. Finally, the IRR for scoring the entire procedure was calculated using both the step and error agreements / disagreements for the entire procedure (N = 122). Acceptable IRR is = or > 0.80.[210]

5.3 Results

5.31 Participants

Two groups were compared in their performance of an arthroscopic Bankart repair on a cadaver shoulder. The entire group of Master or Associate Master instructors, serving as faculty for an Arthroscopy Association of North America Resident's course, chose to participate and comprised the experienced group (N = 10). The faculty, all fellowship trained in arthroscopy or sports medicine, averaged over 16 years in clinical practice with each having routine experience in employing arthroscopic shoulder techniques. All faculty members have been recognized nationally by AANA for their talent and ability to teach and communicate shoulder arthroscopy skills to trainees. The novice group (N = 12) was comprised of 11 PGY 5 and one PGY 4 orthopedic resident volunteers (from a total of 44 orthopedic residents registered for the course) who elected to participate in the study and perform an arthroscopic Bankart repair on a cadaver shoulder. Those volunteers had previously registered for an AANA Resident's Course with no prior knowledge of the Bankart repair assessment protocol. Other than their year in training, no information regarding their arthroscopic experience or surgical skill was obtained.

5.32 Cadaver Specimens

One specimen was rejected and replaced from each of the study groups. A specimen for the novice group was large with very non-compliant tissues substantially restricting the anterior working space and making the use of instruments difficult. One specimen from the experienced group was arthritic which limited the ability to distract the humeral head from the glenoid – visualization and the use of instruments from the posterior portal were unacceptably restricted. A fresh cadaver shoulder replaced each of these two specimens.

5.33 Inter-rater reliability (IRR) assessments

The IRR calculations across each of the assessments were strong. 21 videos could be scored completely. One novice completed only a single anchor repair during the entire duration of the procedure, which provided incomplete data. The mean IRR for the paired scoring of the 21 videos for procedural steps was 0.91 (range = 0.82 – 1.00) and for errors including DNTT events was 0.93 (range = 0.77 – 1.00 [the 0.77 for the error IRR calculation for 1 video was the single instance that fell below 0.80 out of the 63 IRR calculations]) The mean IRR for the total of steps and errors was 0.93 (range = 0.81 – 0.98).

5.34 Outlier Performance

One novice subject completed only 13 of 45 steps and a one anchor repair before failing to progress and electively terminating the procedure. During that time, 4.5 errors and 1 sentinel error were created (the average of the pair of designated reviewers). Inclusion of the data for the relatively small number of errors enacted during the partial repair would bias and understate the average number of errors for the novice group. As a result, all scores for this outlier were removed. The number of steps could theoretically have been used as that number accurately reflected the number of steps that were actually completed, but we elected to use none of the data from this subject's limited repair.

Before analyzing the data for complete repairs, score profiles were examined for significantly atypical performance of the experienced group. One subject in the experienced group took dramatically longer than their colleagues to perform the procedure, primarily due to substantial difficulties with suture delivery and management. This subject required 63 minutes to complete the Bankart repair in comparison to their colleagues' mean of 29.5 minutes (SD = 8 min.). For the experienced surgeon outlier, $Z = 4.6$ with an associated probability value < 0.001 which indicated that the difference in performance from their peers was highly significant. Consistent with the prospectively established policy of removing an experienced subject's scores if their performance was >2 SD from the group mean, this subject's data was removed from further statistical comparisons between the groups.

5.35 Experienced / Novice Group Comparisons

Comparisons were made separately for steps and errors for each of the 13 phases of the Bankart procedure as well as the summary data for steps, errors and sentinel errors (Appendix 5A – novice; 5B - experienced). The phases of anchor preparation / insertion, suture delivery and management, and knot tying were repeated for each of the 3 anchors. The three sets of data for the three similar phases (1 for each of 3 anchors) were combined. Novice surgeons made significantly more objectively scored overall procedure errors (Figure 5.1) than the experienced surgeons (novice surgeons = 5.68 vs. experience surgeons 2.61; $p = 0.026$). Not only did the novice surgeons make more errors but they also showed greater performance variability as demonstrated by the considerably larger standard deviation score (3.5 vs. 1.6). The greatest difference in the mean number of errors made, occurred during the suture delivery and management phases of the procedure which was statistically significant (novice surgeons = 1.95 errors vs. experienced surgeons = 0.33 errors; p = 0.02, Figure 5.1). Novice surgeons also made more sentinel errors than the experienced surgeons (1.5 vs. 0.94), but this difference was not statistically significant.

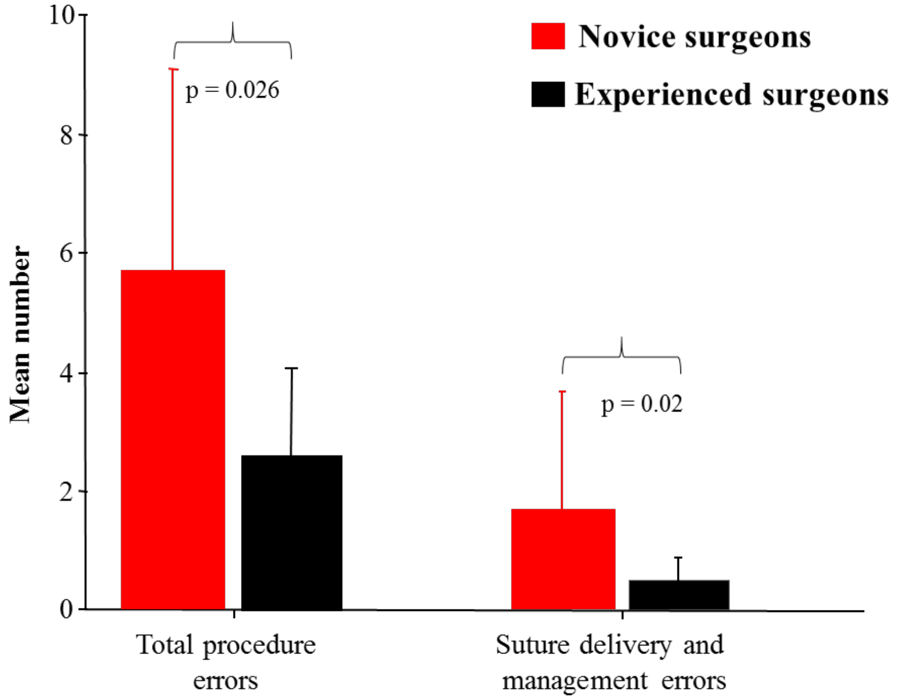


Figure 5.1 Mean total errors enacted by the novice and experienced surgeon groups; p = 0.026; Mean suture and delivery errors enacted by the novice and experienced surgeon groups; p = 0.02

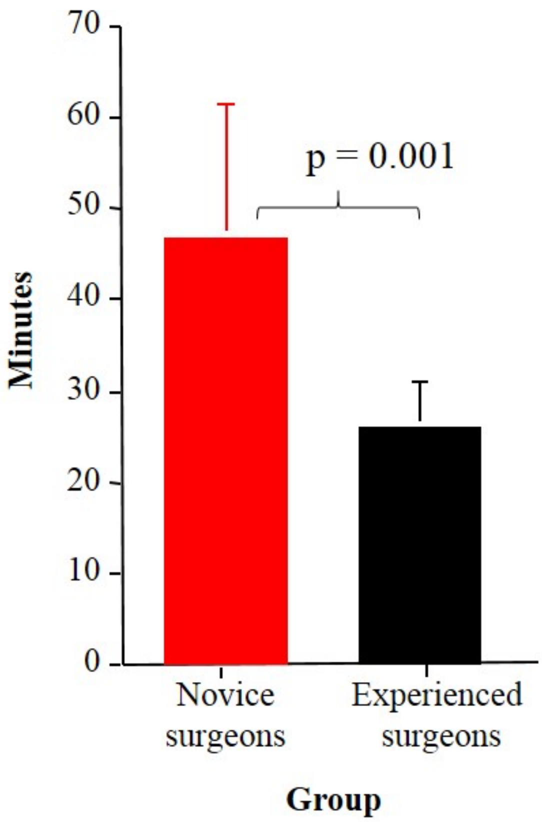


Figure 5.2 Mean time in minutes taken by the novice and experienced surgeon groups to complete a 3 suture anchor Bankart repair; $p < 0.001$.

The most common errors and sentinel errors are shown in table 5.1 with those errors common to all of the 3 anchors being summed. With respect to regular errors, failure to maintain intra-articular position of the cannulas was frequently observed for the novice group. Both groups experienced occasional instances of anchor pullout, the experienced group somewhat more often than the novice group. By far, the most common sentinel error enacted by the novice group was improper introduction of the suture delivery device into the capsule at or above the anchor hole resulting in failure to retention the capsuloligamentous tissues superiorly. Damage (laceration) of the intact labrum during attempts to mobilize the capsulolabral tissues was also notably more common among the novice group compared to the experienced group.

Overall, the novice surgeons also completed fewer steps than the experienced surgeons (35.04 vs. 38.33) but this difference was not statistically significant ($p = 0.186$).

Figure 5.2 shows the mean amount of time both groups of subjects took to perform the procedure. Novice surgeons took significantly more time to perform the repair than the experienced group (novice surgeons = 64 (SD = 19.25) vs. experience surgeons 33 (SD = 7.88), $p < 0.001$).

5.4 Discussion

5.41 Novice vs. Experienced Surgeon Performance

The present study demonstrates robust construct validity for the use of the arthroscopic Bankart procedure metrics with a cadaver shoulder. The Bankart metrics are both precise (high IRR) and accurate (able to distinguish between novice and experienced surgeon performance.[211] Overall, experienced surgeons performed better than novice orthopedic surgeons when evaluated using an objectively assessed and blinded review of video recorded operative performance. Whilst the objectively assessed performance of experienced surgeons was better than the novice surgeons across all of the measures, the metrics which best distinguished the two groups were procedure errors, particularly suture management errors. Operative time was also significantly different. We are unaware of previous similar studies using a detailed metric based assessment of a complete surgical procedure with which to compare and contrast our results.

5.42 Tool Development

At the outset of the series of investigations termed the AANA "Copernicus Initiative" (a paradigm shift from the apprenticeship model to one of proficiency based progression (PBP) training), we sought to study the effectiveness of PBP training for surgical skills. This investigation required the development and validation of 3 separate, specific tools to conduct the analysis. The first component to be created was a metric tool (steps, errors, and sentinel errors) for a specific procedure (an arthroscopic Bankart repair [ABR] was selected). This metric tool was demonstrated to have face and content validity.[65] Secondly, a training tool (a shoulder model simulator coupled with the ABR metrics) was demonstrated to have construct validity for an ABR with the ability to distinguish between experienced and novice surgeon performance.[93] Lastly, the current study demonstrates construct validity for the cadaver shoulder (coupled with the ABR metrics) as a valid assessment tool for comparing the performance of different surgeons.

5.43 Inter-rater Reliability

The very high IRR for the scores from reviewer pairs for the entire group of metrics (0.92) is reflective of the clarity and precision of the arthroscopic Bankart metrics drafted, and the thorough training of the 10 reviewers. The ability to score the steps and errors consistently is essential to obtain a reliable measure of the surgeon's performance and skill level for a particular procedure.

5.44 Shoulder Simulator Model vs. Shoulder Cadaver

For the prior study undertaken to assess construct validity for the shoulder model simulator and metrics, surgeons in the experienced group made 63% fewer errors, committed 79% fewer sentinel errors, and performed the procedure in 42% less time than those in the novice group (all differences being significant). The greatest difference in errors between the groups involved anchor preparation and insertion, suture delivery and management, and knot tying. In the current study employing a cadaver shoulder, experienced surgeons made 54% fewer errors, and performed the procedure in 48% less time (both differences being significant). Sentinel errors were significantly less for the experienced group in the model validation study, and were also committed less frequently by the experienced surgeons in this cadaver evaluation, although the difference was not statistically significant. With respect to specific phases of the procedure, the greatest discriminator in both investigations was for the phases of suture delivery and management. This finding is not surprising as the steps involved in that phase are among the most challenging for the Bankart repair. The number of steps performed did not differentiate in either evaluation. This result is not unexpected as the intent and effort to perform each of the steps predominantly reflects a familiarity and knowledge of the steps necessary to perform the procedure. Overall, in both studies, the experienced group showed less performance variability than the novice surgeons as demonstrated by their smaller standard deviation scores.

The benchmark was established based on the mean performance of the group of experienced surgeons. For this current cadaver study, it included: completion of a 3 anchor Bankart repair with no more than 3 total errors and no more than 1 sentinel error. For the similar previous study on the simulator model using the identical metrics, the one difference in the benchmark was that no more than 4 (instead of 3) total errors were permitted.

5.45 Novice / Experienced Outliers

The partial data from one novice surgeon was removed from the analysis, as the surgeon was not able to complete the suturing and knot tying for the first anchor prior to electively terminating the procedure due to the inability to make progress. A relatively large number of errors were enacted (ave.4.5) (Appendix 5A) for the portion of the procedure performed, but it was not possible to accurately estimate or extrapolate to the total number of errors that might have been created had the entire procedure been completed. The total number of errors enacted would likely have been substantial had the procedure been completed. Thus, the average total number of errors and sentinel errors is likely to have been substantially understated for the novice group. A relatively small number of steps were accomplished (ave.13) (Appendix 5A) and had the data for this novice been included in that analysis, it would have impacted the average total number of steps that the novices completed. Given that the diagnostic portion of the procedure took over 25 minutes for this subject, the overall time for completion of the 3 anchor repair would also likely have been much larger. Because the data for the majority of the analysis was incomplete, it was not possible to include any of that surgeon's performance data in the analysis.

One of the issues that emerged during this study and indeed in one of our previous studies[93] was atypical performance of one experienced subject. Atypical performance is an important issue as it relates to establishing benchmarks, which may have considerable implications for trainee progression. Since proficiency-based progression training was first introduced and validated in 2002[22], the average or mean performance of experienced operators has been used as the performance benchmark, which trainees must meet and demonstrate before being allowed to progress in their training.[22, 79, 198, 201, 208] If an experienced individual's performance score was dramatically worse than their peers, and their scores were to be included the establishment of the benchmark, that reference level would clearly be lowered. The lowering of the performance threshold (benchmark) could have important patient safety implications. For example, in a study on bariatric surgery it was found that surgeons performing at the lower end of the performance range had significantly poorer outcomes than surgeons performing at the upper end.[220] This was one of the first studies to quantitatively link objectively assessed surgeon skill performance with patient outcomes.

The criteria for removing outlier data from the group being used to create a performance benchmark must be established before the data is collected and should be objective, transparent, and fair. At the outset, prior to conducting the study reported here, the authors identified and discussed the possibility of encountering atypical experienced surgeon performance. The core group of 4 primary investigators agreed that the resolution to this potential issue would be to remove all of a subject's scores from subsequent analysis if it could be unambiguously established that the subject's performance was statistically atypical, (greater than 2.0 standard deviations from the group mean). The experienced individual participating in this study was performing considerably worse than that for operative time (2.44 SD from the mean inclusive of the outlier).

5.46 Performance Errors

The enactment of errors is emerging as one of the most important indicators of skill for operative performance.[65] An individual may be able to perform all of the correct steps in an acceptable order with the appropriate instruments and score very well on those parameters, but still perform the steps poorly. Procedural errors are operative behaviors that deviate from optimal performance. These metrics are a reliable measure of performance quality and are likely to be the most sensitive assessment tool in the evaluation of operative performance and safety.[30] Although simulation-based education and the resulting transfer of training will influence other performance parameters as well, such as procedural time, the greatest impact of that training strategy appears to be on limiting performance errors.[30] It is therefore a necessity, at the outset, that the performance characterization of 'deviations from optimal performance' (errors) must be particularly robust and well validated. It also implies that error performance assessed using a less rigid *Likert scale* (table 3.1)(global rating scales) may result in a less focused approach to minimizing errors due to the fact that the deviations from optimal performance have been less clearly defined[63, 64] A 'Likert-type scale' is a method of ascribing a quantitative value to qualitative data to make it amenable to statistical analysis. Likert scales (often with a range from 1-5, or 1-7) are typically constructed with responses (opinion) around a neutral option (i.e. "suture delivery was: 1=awkward …3=effective…5=highly efficient") and were originally designed to assess a range of respondent attitudes.[62] Due to the inherent subjectivity in this method of attempting to rate objective performance, it may be difficult to obtain acceptable levels of inter-rater reliability [> or = 80%] in the scoring of events.[64] In contrast, the approach to the assessment of performance used in this study employs precise definitions of performance and simply requires the video reviewer to determine whether the specific event did or did not occur. This binary approach to the measurement of performance has been shown to facilitate the reliable scoring of metric-based performance units across a variety of functions from skills training[195-198] at different experience levels.[199, 200] It has also been shown to considerably enhance assessment reliability levels in comparison to Likert-scale scoring.[63]

The effectiveness of a deliberate practice, proficiency-based training curriculum employing simulation relies on a clear and specific identification, not only of the proper steps that the trainee should perform, but also, of what the trainee is doing wrong and how to prevent or correct their error. Other advantages of creating comprehensive procedure characterizations and explicit operational step and error definitions exist. Detailed metrics provide very clear guidance for the construction of simulation training platforms, specifying exactly what the simulator should be capable of emulating, but more importantly, measuring.[212] Comprehensive procedure characterization is challenging and time consuming the first time it is undertaken and requires robust validation of all of the performance metrics. With experience, however, this methodology is considerably easier to apply to subsequent characterizations of different procedures by the same group.

5.47 Limitations

A limitation of this study relates to the use of cadaveric specimens for the arthroscopic Bankart creation and repair. The specimens lacked some uniformity in the integrity of the capsule and labrum, soft tissue compliance, shoulder mobility / distractibility, and the bulk of the extra-articular tissues. In addition, while specific parameters were utilized for the creation of the Bankart (i.e., 2 – 6 o'clock on the glenoid rim and 6 – 9 mm deep / medial), the lesions could not be made absolutely uniform. Further variability existed in the presence of coexisting pathology, i.e. arthritis, synovial proliferation, rotator cuff partial tears, etc. The "acceptability criteria" for the specimens (listed above) were employed to minimize the impact of this potential problem.

An additional limitation of this study is that there was no confirmation that those serving as Master / Associate Master surgeons and representative of the "experienced" group possessed a specified level of expert skill in performing an arthroscopic Bankart procedure. Nevertheless, the individual surgeons so identified have been recognized by the Arthroscopy Association of North America as valuable educators either from lecture presentations with video exhibiting skilled shoulder arthroscopy techniques or from repeated experience teaching in an arthroscopic lab setting with the ability to demonstrate and teach each of the key components of a Bankart procedure. Thus, 'experienced' rather than 'expert' is a reasonable description of the group. Similarly, other than identifying the year in training, no additional information was obtained to determine the extent of the resident's (novice group) experience with arthroscopic shoulder surgery, i.e., number of arthroscopy / sports medicine rotations previously completed; number of shoulder arthroscopic surgeries as an assistant surgeon, etc. Even with that data, accurate knowledge of the level of skill possessed by an individual resident would not be possible. Thus, while the arthroscopic skill sets of the subjects are representative of their respective groups and experience, those skills are highly likely to be somewhat heterogeneous.

We acknowledge that only a single operative procedure was analyzed for each of the subject surgeons. It is possible that data averaged over several procedures would be somewhat different than that obtained in this study. Cost and time considerations made the performance of a single arthroscopic Bankart repair on a cadaver shoulder most feasible. It should also be noted that the participants in each group had no prior specific knowledge of the metrics to be scored in the review of their procedure and we suspect that the experienced surgeons in particular, might have performed and scored differently (better) for certain non-crucial parts of the procedure, e.g., the diagnostic steps at the beginning, had they been familiar with the metrics to be evaluated.

5.5 Conclusions

The assessment tool comprised of validated arthroscopic Bankart metrics coupled with a cadaver shoulder accurately distinguishes the performance of experienced from novice orthopedic surgeons. A benchmark based on the mean performance of the experienced group includes: completion of a 3 anchor Bankart repair, and enacting no more than 3 total errors and 1 sentinel error.

Appendix 5A – Cadaver Metric Validation: Novice

Video #	33A	33B	33 ave	73A	73B	73 ave	123A	123B	123 ave	53A	53B	53 ave	83A	83B	83 ave	113A	113B	113 ave	13A	13B	13 ave	93A	93B	93 ave	43A	43B	43 ave	63A	63B	63 ave	23A	23B	23 ave	103 A	103 B	103 ave
I - Portals																																				
Steps uncompl.	1	1	1	1	1	1	3	3	3	0	0	0	1	1	1	0	3	1.5	1	1	1	3	1	2	1	0	0.5	0	0	0	0	0	0	0	0	0
Errors made	2	2	2	0	0	0	1	1	1	0	3	1.5	0	0	0	2	1	1.5	1	0	0.5	1	0	0.5	0	1	0.5	0	0	0	0	2	1	0	0	0
II - Instabl Asses.																																				
Steps uncompl.	3	5	4	3	5	4	3	3	3	0	1	0.5	2	3	2.5	1	2	1.5	1	1	1	3	3	3	1	0	0.5	1	0	0.5	2	0	1	1	1	1
Errors made	0	0	0	0	0	0	1	0	0.5	1	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
III - Caps/Gen Prep																																				
Steps uncompl.	1	2	1.5	3	3	3	4	4	4	1	1	1	5	5	5	4	4	4	1	2	1.5	3	2	2.5	3	1	2	5	3	4	3	3	3	1	3	2
Errors made	0	1	0.5	0	0	0	0	0	0	1	1	1	0	1	0.5	0	0	0	0	3	1.5	1	1	1	1	0	0.5	2	1	1.5	0	0	0	0	0	0
IV - 1st Inf Anch Prep																																				
Steps uncompl.	0	0	0	1	1	1	1	0	0.5	0	0	0	1	0	0.5	1	1	1	1	0	0.5	0	0	0	0	0	0	0	1	0.5	1	0	0.5	0	0	0
Errors made	0	0	0	0	0	0	1	2	1.5	0	0	0	1	0	0.5	1	2	1.5	0	0	0	3	2	2.5	1	1	1	0	0	0	0	1	0.5	0	0	0
V - 1st Sut Del / Mgmt																																				
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	1	1	1	0	0	0	2	1	1.5	1	1	1	2	1	1.5	1	1	1	0	0	0	2	3	2.5	0	0	0	4	1	2.5	0	0	0	2	2	2
VI - 1st Knot Tying																																				
Steps uncompl.	0	0	0	0	0	0		Procedure terminated by surgeon		0	0	0	1	0	0.5	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	1	0	0.5				0	1	0.5	2	1	1.5	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5
VII - 2nd Anch Prep																																				
Steps uncompl.	1	1	1	1	0	0.5				1	1	1	2	1	1.5	0	0	0	1	0	0.5	0	1	0.5	0	1	0.5	1	1	1	1	1	1	2	1	1.5
Errors made	0	0	0	0	0	0				1	2	1.5	2	2	2	0	0	0	0	2	1	1	2	1.5	0	0	0	0	0	0	0	0	0	1	2	1.5
VIII - 2nd Sut Del/ Mgmt																																				
Steps uncompl.	0	0	0	0	0	0				0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0				0	0	0	1	3	2	0	1	0.5	0	0	0	3	2	2.5	0	0	0	2	0	1	0	1	0.5	0	0	0
IX - 2nd Knot Tying																																				
Steps uncompl.	0	0	0	0	0	0				0	0	0	3	2	2.5	3	3	3	1	1	1	1	0	0.5	0	0	0	0	1	0.5	0	0	0	0	0	0
Errors made	0	0	0	0	0	0				0	0	0	0	5	2.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
X - 3rd Anch Prep																																				
Steps uncompl.	0	0	0	1	0	0.5				1	1	1	1	1	1	4	4	4	1	0	0.5	4	4	4	0	1	0.5	1	1	1	1	1	1	1	0	0.5
Errors made	0	0	0	0	0	0				0	0	0	1	2	1.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
XI - 3rd Sut Del/ Mgmt																																				
Steps uncompl.	1	0	0.5	0	0	0				0	0	0	0	0	0	2	2	2	0	0	0	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0				1	0	0.5	1	0	0.5	0	0	0	0	0	0	0	0	0	0	1	0.5	2	1	1.5	0	1	0.5	0	0	0
XII - 3rd Knot Tying																																				
Steps uncompl.	0	0	0	0	0	0				0	0	0	0	0	0	3	3	3	0	0	0	3	3	3	0	0	0	1	1	1	0	0	0	0	0	0
Errors made	0	1	0.5	0	0	0				0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	1.5	1	1	1	1	1	1	0	1	0.5
XIII - Eval Repair																																				
Steps uncompl.	0	0	0	0	0	0				0	0	0	0	0	0	2	2	2	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0				0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Portal / Dx Time (B4 Bank)			6			13			26			14			17			36			5			19			9			34			23			7
Bankart Repair Time			50			59			15			45			75			35			31			68			47			39			32			33
Total Time (Dx+Rx)			56			72			41			59			92			71			36			87			56			73			55			40
									DNC																											

(continued)

Video #	33A	33B	33 ave	73A	73B	73 ave	123A	123B	123 ave	53A	53B	53 ave	83A	83B	83 ave	113A	113B	113 ave	13A	13B	13 ave	93A	93B	93 ave	43A	43B	43 ave	63A	63B	63 ave	23A	23B	23 ave	103 A	103 B	103 ave
Rating Pairs																																				
Steps completed (45)	38	37	37.5	35	35	35	*12*	*14*	*13*	41	42	41.5	30	32	31	26	22	24	38	39	38.5	19	25	22	40	42	41	36	37	36.5	37	40	38.5	40	40	40
Errors made (77)	3	5	4	1	0	0.5	*5*	*4*	*4.5*	4	8	6	10	16	13	5	5	5	1	5	3	11	10	10.5	5	3	4	11	4	7.5	2	7	4.5	4	5	4.5
Sentinel errors	1	3	2	0	0	0	*1*	*1*	*1*	1	2	1.5	3	4	3.5	1	1	1	0	1	0.5	2	1	1.5	2	0	1	5	3	4	0	1	0.5	1	1	1
Steps completed																																				
Agreement		39			41						44			40			41			40			37			39			40			42			40	
Disagreement		6			4						1			5			4			5			8			6			5			3			5	
Step IRR		0.87			0.91						0.98			0.89			0.91			0.89			0.82			0.87			0.89			0.93			0.89	
Errors made																																				
Agreement		75			76						68			59			69			71			68			73			70			72			73	
Disagreement		2			1						9			18			8			6			9			4			7			5			4	
Error IRR		0.97			0.99						0.88			0.77			0.9			0.92			0.88			0.95			0.91			0.93			0.95	
Total score (S+E)																																				
Agreement		114			117						112			99			110			113			105			112			110			114			113	
Disagreement		8			5						10			23			12			9			17			10			12			8			9	
Total IRR		0.93			0.96						0.92			0.81			0.9			0.93			0.86			0.92			0.90			0.93			0.93	

NOTE. Data in italics indicate incomplete procedure.
anch, anchor; caps, capsule; DNC, did not complete; Dx, diagnostic; E, errors; eval, evaluate; inf, inferior; instabl, instability; IRR, inter-rater reliability; mgmt, management; Rx, treatment; S, steps; sut del, suture delivery.

Appendix 5B – Cadaver Metric Validation: Experienced

Video #	12A	12B	12 ave	22A	22B	22 ave	112A	112B	112 ave	122A	122B	122 ave	32A	32B	32 ave	42A	42B	42 ave	52A	52B	52 ave	62A	62B	62 ave	82A	82b	82 ave	92A	92B	92 ave
I - Portals																														
Steps uncompl.	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	1	0.5	2	2	2	0	0	0	2	0	1
Errors made	1	1	1	1	2	1.5	1	1	1	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
II - Instabl Asses.																														
Steps uncompl.	1	3	2	0	2	1	1	1	1	0	1	0.5	1	2	1.5	3	3	3	1	1	1	2	0	1	1	1	1	2	0	1
Errors made	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0
III - Caps/Gen Prep																														
Steps uncompl.	2	3	2.5	4	3	3.5	2	1	1.5	0	1	0.5	1	1	1	2	2	2	0	1	0.5	1	1	1	3	0	1.5	4	3	3.5
Errors made	0	0	0	0	1	0.5	2	3	2.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IV - 1st Inf Anch Prep																														
Steps uncompl.	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	0.5
Errors made	0	0	0	0	0	0	0	0	0	2	1	1.5	0	0	0	1	1	1	0	0	0	0	0	0	0	1	0.5	0	1	0.5
V - 1st Sut Del/Mgmt																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0	1	0	0.5	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5
VI - 1st Knot Tying																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
Errors made	0	1	0.5	0	0	0	0	1	0.5	0	0	0	1	2	1.5	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0
VII - 2nd Anch Prep																														
Steps uncompl.	1	1	1	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	1	0	0.5	1	1	1	0	1	0.5	1	1	1
Errors made	0	0	0	2	1	1.5	0	0	0	0	0	0	0	0	0	0	1	0.5	0	2	1	1	1	1	0	0	0	0	0	0
VIII - 2nd Sut Del/Mgmt																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	1	0.5	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.5
IX - 2nd Knot Tying																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
X - 3rd Anch Prep																														
Steps uncompl.	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	0	0.5	1	1	1	1	1	1
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	1	0.5	0	0	0	0	0	0	1	1	1	0	0	0
XI -3rd Sut Del/Mgmt																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	1	1	1	0	0	0	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
XII - 3rd Knot Tying																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	1	1	1	2	2	2
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
XIII - Eval Repair																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Portal/Dx Time (B4 Bank)			8			9			5			6			7			2			18			1			6			3
Bankart Repair Time			36			30			26			28			63			17			24			28			27			23
Total Time (Dx+Rx)			44			39			31			34			70			19			42			29			33			26
Rating Pairs																														

(continued)

Video #	12A	12B	12 ave	22A	22B	22 ave	112A	112B	112 ave	122A	122B	122 ave	32A	32B	32 ave	42A	42B	42 ave	52A	52B	52 ave	62A	62B	62 ave	82A	82b	82 ave	92A	92B	92 ave
Steps completed (45)	37	35	36	39	38	38.5	40	41	40.5	41	40	40.5	36	37	36.5	36	36	36	43	42	42.5	38	41	39.5	38	38	38.5	30	36	33
Errors made (77)	1	4	2.5	4	5	4.5	5	5	5	2	1	1.5	3	9	6	3	6	4.5	0	3	1.5	1	1	1	1	2	1.5	1	2	1.5
Sentinel errors	0	1	0.5	1	1	1	3	2	2.5	1	0	0.5	0	2	1	1	3	2	0	0	0	1	1	1	0	0	0	1	1	1
Steps completed																														
Agreement		40			42			44			42			41			45			42			43			38			39	
Disagreement		5			3			1			3			4			0			3			2			7			6	
Step IRR		0.89			0.93			0.98			0.93			0.91			1			0.93			0.95			0.84			0.87	
Errors made																														
Agreement		73			74			71			76			69			74			74			77			76			74	
Disagreement		4			3			6			1			8			3			3			0			1			3	
Error IRR		0.95			0.96			0.92			0.99			0.9			0.96			0.96			1			0.99			0.96	
Total score (S+E)																														
Agreement		113			116			115			118			110			119			116			120			114			113	
Disagreement		9			6			7			4			12			3			6			2			8			9	
Total IRR		0.93			0.95			0.94			0.97			0.9			0.97			0.95			0.98			0.93			0.93	

anch, anchor; caps, capsule; Dx, diagnostic; E, errors; eval, evaluate; inf, inferior; instabl, instability; IRR, inter-rater reliability; mgmt, management; Rx, treatment; S, steps; sut del, suture delivery.

6. "Objective Assessment of Knot-Tying Proficiency With the Fundamentals of Arthroscopic Surgery Training Program Workstation and Knot Tester" Pedowitz RA, Nicandri GT, Angelo RL, Ryu RK, Gallagher AG. Arthroscopy 2015;31:1872-1879. Appendix 8

(Candidate was key contributor to study design and experimental set-up)

Purpose: To assess a new method for biomechanical assessment of arthroscopic knots, and to establish proficiency benchmarks using the Fundamentals of Arthroscopic Surgery Training (FAST) Program workstation and knot tester.

Methods: The first study group included twenty faculty at an AANA resident arthroscopy course (19.9 $\pm$ 8.25 years in practice). The second group had thirty experienced surgeons attending an AANA fall course (17.1 $\pm$ 19.3 years in practice). The training group included forty-four PGY4/5 orthopedic residents in a randomized, prospective study of proficiency-based training, with three sub-groups: Group A (standard training, n=14), Group B (workstation practice, n=14), Group C (proficiency progression using the knot tester, n=16). Each subject tied five arthroscopic knots backed up by three reversed hitches on alternating posts. Knots were tied under video control around a metal mandrel through a cannula within an opaque dome (the FAST workstation). Each suture loop was stressed statically at 15 pounds for 10 seconds). A calibrated sizer measured loop expansion. Knot failure was defined as $\geq$ 3 mm of loop expansion.

Results: In the faculty group, 24% of knots "failed" under load. Performance was inconsistent: 12 faculty had all knots pass, while 2 had all knots fail. In the second group of practicing surgeons, 21% of the knots failed under load. Overall, 56 of 250 knots (22%) tied by experienced surgeons failed. For the PGY4/5 residents, aggregate knot failure rate was 26% of 220 knots. Group C residents had an 11% knot failure rate (½ the overall faculty rate, p = 0.013).

Conclusion: The FAST workstation and knot tester offer a simple and reproducible educational approach for enhancement of arthroscopic knot-tying skills. Our data suggest that there is significant room for improvement of the quality and consistency of this important arthroscopic skill, even for experienced arthroscopic surgeons.

6.1 Introduction

Knot tying is an essential skill for proficiency in arthroscopic surgery.[221, 222] Arthroscopic knot tying is difficult to teach and to assess objectively. At this time, most trainees are assessed by visual inspection of arthroscopic knots, either by direct view or via an arthroscopic image. Hanypsiak and co-workers[223] recently demonstrated that even experienced practicing surgeons are relatively inconsistent when it comes to arthroscopic knot tying. Technical inconsistency could have a negative impact upon surgical outcomes.

In the laboratory setting, arthroscopic knots (more accurately, the suture loops created after knot tying) are usually tested with expensive material-testing devices that allow sophisticated variation of load magnitude, cyclic versus single pull, loop preload, and the rate of load application.[222, 224-230] However, these devices are not practical for day-to-day education of residents and fellows or for continuing medical education of practicing surgeons. It would be advantageous to have a cost-effective and relatively simple-to-use tester for objective assessment of knot *performance,* as opposed to knot *appearance*.

In the teaching lab, knot-tying skills are generally developed using knot tying boards under direct visualization with the trainee's eyes look directly at the hands, suture and associated surgical instruments. However, in the clinical setting, arthroscopic knots and backup hitches are created outside of the body, delivered through an arthroscopic cannula, and then tensioned within the joint, with visualization provided by a two-dimensional video screen. This combination requires an integrated chain of complex, psychomotor skills that are performed in 3-D space with mostly binocular cues.[231] Such skills are best acquired and rehearsed in a gradual and systematic fashion.[79, 201]

The Fundamentals of Arthroscopic Surgery Training (FAST) Program is a collaborative initiative of the Arthroscopy Association of North America (AANA), the American Academy of Orthopedic Surgeons (AAOS), and the American Board of Orthopedic Surgery (ABOS). The FAST Program offers a basic arthroscopic motor skills curriculum with associated teaching modules to facilitate core training in orthopedic surgery. It is logical to achieve a baseline level of technical proficiency, if possible, prior to operating on patients.[232]

The FAST Program curriculum was developed after task deconstruction of basic arthroscopic skills (available open access at http://www.aana.org/FAST.aspx). The FAST workstation (Sawbones, Vashon Island, WA) was custom-designed for training of these skills. The system allows for initial practice under direct visualization, then advancing to triangulation through simulated portals under direct visual control, and finally moving to skill rehearsal through simulated portals using a video camera, with direct surgeon view eliminated. The purpose of the present study was to assess the FAST knot tester, and to establish benchmarks for knot tying proficiency using this system. Our hypothesis was that the FAST knot tester would facilitate objective, accurate, and immediate mechanical assessment of knot performance.

6.2 Methods

For all groups in the current study, five consecutive knots were created by each subject on the FAST workstation using #2 Fiberwire (Arthrex, Naples, FL) under dry, room temperature conditions via a 7 mm plastic cannula. Each subject created an arthroscopic knot of their choice, backed up by three reversed half hitches on alternating posts. Each suture was labeled, well away from the knot and suture loop, for later identification. The five knots were gently placed within a labeled plastic bag for each subject and set aside for subsequent analysis using the FAST knot tester.

6.21 Faculty Reference Group 1 (n = 20)

The first group (Faculty) was comprised of twenty experienced surgeons teaching at a dedicated AANA resident arthroscopy skills course at the Orthopedic Learning Center (Rosemont, Illinois). This expert group reported clinical practice experience of 19.9 $\pm$ 8.25 years and performed 381 $\pm$ 150 arthroscopies per year.

6.22 Resident Comparison Groups (n = 44)

Orthopedic surgery residents (PGY4 and PGY5) participated in a randomized, prospective study of proficiency-based training at the orthopedic learning center (the AANA "Copernicus Study", described in detail in a separate publication[94] that did not include specific information about knot tying performance, benchmarks, and associated methodology) Residents were divided into three sub-groups. Group A included fourteen residents who were instructed on knot tying skills using standard educational methodology during a regular AANA Resident's Arthroscopy Course. Standard educational methodology included didactic instruction, faculty demonstration, practice with rope, and progression to knot tying using suture around a metal hook and then through an arthroscopic cannula, all under direct visualization. When Group A participants felt ready for testing, each resident used the FAST workstation and USB camera system to create five arthroscopic knots in sequence without interval feedback.

Group B included fourteen residents who received similar didactic instruction as Group A, but they were also allowed to *practice* knot tying skills using the FAST workstation / USB camera system until they were ready to create 5 knots for later testing. Group C had sixteen residents who received the same didactic instruction and practiced with the workstation / USB camera set up. However, Group C residents were allowed to use the FAST knot tester after each knot was tied providing immediate performance feedback *during* the practice phase, until they were ready to create five test knots. Practice time prior to knot testing was not a controlled variable for the resident study groups.

6.23 Surgeon Reference Group (n = 30)

Thirty surgeons volunteered to create 5 knots using the FAST workstation and USB camera set-up at the 2013 AANA Fall Course. For the purposes of setting the benchmark for resident proficiency (see below), we only used the faculty from the Copernicus course as the reference group. We were surprised at the high knot failure rate in the Copernicus faculty, so we pursued an additional cohort of practicing surgeons (whether faculty or non-faculty surgeons) from the AANA fall course. We thought that this would, at a minimum, represent arthroscopic surgeons in practice, and felt that observations in fifty practicing surgeons would enhance overall confidence in the observations. Knots were tested later with no feedback provided to the surgeon during knot creation. All participants were in clinical practice for at least one year (maximum 40 years). The group had an average practice experience of 17.1 $\pm$ 19.3 years (mean $\pm$ SD). Ten surgeons self-reported as course faculty, twelve self-reported as attendees, and eight surgeons did not indicate whether they were faculty or attendee.

6.24 Study Participants

Each subject was verbally informed about the purpose of the study, and all volunteers were assigned a unique identification number. All test knots were labeled using subject identification number (subject name excluded). The study protocol was reviewed by the Western Institutional Review Board (WIRB, Puyallup, WA) and deemed to be exempt.

6.25 FAST Workstation

The FAST workstation is comprised of a base unit, which accommodates various snap-in teaching modules that complement the FAST Program curriculum. The base station can be used for basic skills practice under direct visualization without the need for triangulation. Two snap-in dome units allow for skills rehearsal under either direct visualization with the lucent dome (Figure 6.1) or with video imaging using the opaque dome (Figure 6.2). Both domes have multiple, identically-positioned access holes that mimic portal positions and geometries of knee and shoulder arthroscopy.

Figure 6.1 Fundamentals of Arthroscopic Surgery Training (FAST) base station with knot-tying mandrel and lucent dome for skills rehearsal under direct visualization.

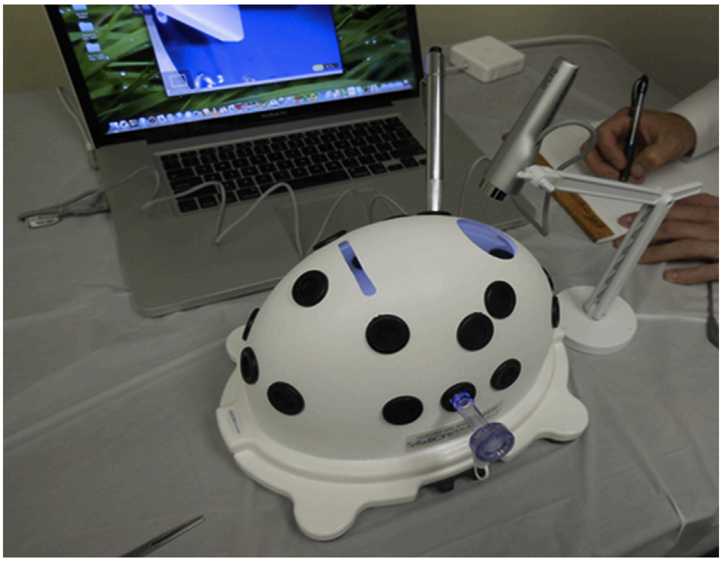


Figure 6.2 Fundamentals of Arthroscopic Surgery Training (FAST) workstation with opaque dome and light-emitting diode penlight. An inexpensive USB camera is directed at the knot-tying mandrel, and the image is displayed on a laptop computer. This arrangement simulates arthroscopic visualization.

The FAST Workstation has a horizontally-positioned smooth stainless steel knot tying mandrel for practice with suture (Figure 6.1). The circumference of the knot tying mandrel matches the first marked position on the conical loop-sizer of the FAST knot tester (Figure 6.3). The loop-sizer is calibrated in one millimeter increments in order to measure up to 5 mm of expansion relative to the knot tying mandrel. Based upon prior literature,[223, 233] 3 mm or more of loop expansion was deemed to be a knot failure. This is considered to be an amount of suture loop elongation that might be associated with biologic healing failure at the tendon-bone interface following rotator cuff repair.[233, 234]

Figure 6.3 Conical loop sizer of Fundamentals of Arthroscopic Surgery Training (FAST) knot tester. The first mark indicates zero loop expansion, and each subsequent mark reflects 1 mm of additional loop expansion compared with the knot-tying mandrel.

6.26 Visualization Protocol

In order to mimic clinical conditions, the FAST workstation was designed for use with an inexpensive USB camera mounted on a stand (Figure 6.2). For the current study, a high resolution Point 2 View camera (IPEVO, Sunnyvale, CA) was mounted on its base station and connected to a laptop computer, which provided the image on the video monitor (Figure 6.2). The camera was directed at the knot tying mandrel through a hole in the opaque workstation dome, which forced subjects to look at the image on the laptop screen during knot tying. The camera was set at 2X screen magnification using the IPEVO software. Illumination within the dome was augmented using a disposable LED penlight (PLED23A, Energizer, St. Louis, MO), although we found that ambient room light was generally sufficient, due to the high sensitivity of the USB camera.

6.27 FAST Knot Tester

The FAST knot tester is comprised of a rigid base with an integrated spring for application of linear tension to a suture loop (Figure 6.4). The tester was designed to apply 15 pounds of tension (60 N) based upon published theoretical modeling of relevant clinical forces, since direct in vivo measurements of post-operative suture tension are not available. Burkhart and co-workers[235] estimated that 60 N would be the maximal force per suture that might be created by muscle contraction after a balanced suture anchor repair of a medium sized rotator cuff tear. Peak loads would potentially be greater with an unbalanced repair or with an abrupt event, such as a postoperative fall. Number 2 high strength sutures are relatively stiff and the material can withstand loads greater than 300 N prior to rupture.[224, 228] Therefore, the 60N load of the FAST knot tester was selected in order to assess *knot performance* as opposed to suture performance.

Figure 6.4 Fundamentals of Arthroscopic Surgery Training (FAST) knot tester. The suture loop is positioned on the 2 tines, and the handle allows for controlled application of 15 lb. of longitudinal load.

After a knot is created on the workstation, the loop is gently slipped off the knot-tying mandrel and transferred to the conical loop-sizer, which gives a baseline measurement of the suture loop. The loop is then transferred to the two tines of the knot tester. One of the tines is solidly attached to the rigid base. Tension is applied via an actuator handle connected via a calibrated spring to the other tine. A force gauge allows the user to apply 15 pounds of axial tension to the suture loop. We chose 10 seconds of steady force application, because pilot studies indicated that significant additional loop expansion did not occur beyond this time point (in fact, most of the loop expansion was observed within a few seconds of force application). After 10 seconds of static load, the actuator handle is released, and the suture loop is removed from the tines and transferred back to the conical loop-sizer. The calibrated markings are used to assess final loop size compared to initial loop size. The loop created with a perfect knot would therefore have zero baseline difference from the knot tying mandrel and zero expansion following load application.

6.28 Reproducibility of Load Application

A digital force scale with a maximum load capacity of 30 pounds (HS-30, CCI Scale Company, Clovis, CA) was used to measure the consistency of load application created by the FAST knot tester. Two sets of 10 measures were acquired by independent observers. One observer applied 15 pounds of load with the actuator handle for ten seconds while looking at the force scale of FAST knot tester. The other observer recorded axial load using the digital force scale after ten seconds. The force scale was connected by a rigid metal link to the suture tine of knot tester. The digital force scale was re-zeroed after each pull of the actuator handle. Roles were reversed for the two observers during the second set of measurements. Using this protocol, mean force application measured by the digital force scale was 15.03 pounds, with mean standard deviation of 0.05 pounds (expressed as the average SD for two independent sets of ten measures).

6.29 Statistical Analysis

Logistic regression analysis was used to compare relative differences between knot tying performance of the three resident trainee groups and faculty knot tying performance. Statistical significance was considered at $p < .05$.

6.3 Results

Performance data from the AANA Copernicus Study participants and from the AANA Fall Course subjects are presented in Table 6.1 and Table 6.2, respectively. This information is stratified according to the number of knots that failed (defined as $\geq$ 3 mm loop expansion) after 15 pounds of static load applied for 10 seconds. Of the twenty Copernicus course faculty, twelve had zero knots fail. Four of the faculty had two knot failures, two faculty had three knot failures, and two of the faculty had all five of their knots fail on the FAST knot tester. Overall, 24% of the faculty knots were considered failures. Only one faculty surgeon tied five consecutive "perfect" knots (zero loop expansion compared to the knot tying mandrel at baseline and zero loop expansion after 15 pounds of load application).

A similarly high rate of knot failure was noted in the second group of experienced surgeons (Table 6.2). Overall, for these practicing surgeons, 21% of the knots were noted to be failures. Five of the 30 participants had three out of 5 knot failures. Taken in aggregate for all knots tied by faculty and practicing surgeons at the two courses, 56 of 250 knots (22%) were deemed to be failures by mechanical testing.

Overall, the knot failure rate was 26% for the 220 knots that were tied by the orthopedic surgery residents (Table 6.1). However, the Group C residents, who were allowed to use the knot tester for feedback *during* the training experience, had an overall 11% knot failure rate, which was exactly one-half the knot rate of the Copernicus Course faculty.

Data for the Copernicus Course faculty were used to create a proficiency benchmark for "passing" resident performance. This "passing" benchmark was applied to the Group C Copernicus Course residents (the proficiency based progression group, see Angelo et al, 2014 for full details[94]). Based upon the Copernicus faculty data, we defined a proficiency benchmark of less than or equal to two knot failures out of five knot attempts as a "passing" grade.

When we apply these same thresholds to the surgeons in practice at the Fall Course (Table 6.2), with the bar set at less than or equal to two knot failures out of five attempts, five of the thirty surgeons (17%) would have missed the "passing" mark.

Relative to the Copernicus Course Faculty, Groups A and B were more likely to have their knots fail, but these differences were not statistically significant (logistic regression analysis, Group A versus Faculty, p = 0.384; Group B versus Faculty, p = 0.07). In contrast, residents in Group C were less than half as likely to have their knots fail in comparison to the faculty reference group (Odds ratio = 2.84) and this difference was found to be statistically significant (p = 0.013).   Logistic regression analysis was also used to compare the relative differences between the three trainee sub-groups, using standard training (Group A) as the reference. There was no statistical difference between Groups A and B (Odds Ratio = 0.725, p = 0.372). In contrast, Group C residents were almost four times as likely to have their knots pass as Group A (Odds ratio = 3.857, p = 0.002).

If we apply a proficiency threshold of no more than two knot failures of five trials (as described above), 6 out of 44 of the orthopedic residents fell below the "passing" bar (14%). Of note, fifteen out of sixteen Group C residents (94%) exceeded the "passing" threshold; this was the best performance for *any* sub-group of the current study.

6.4 Discussion

The FAST workstation and FAST knot tester facilitated direct, objective measurements of arthroscopic knot-tying performance. Overall, 22% of knots tied by practicing surgeons "failed" using this testing protocol. A proficiency-progression training protocol resulted in improved resident knot tying skills (11% knot failure rate) compared to standard training methodology,

The FAST Program provides core education for orthopedic surgery residents, fellows, and practicing surgeons who wish to develop and enhance their arthroscopic motor skills. The FAST Program is intended to create and enhance robust psychomotor skills, right from the outset of training. It is relatively difficult to correct bad surgical habits once they are firmly established.[79, 201] Many surgical skills can be trained outside of the operating room in an efficient and cost-effective manner that maximizes educational quality and eliminates patient morbidity. There has been a significant shift toward structured simulation training, including recent simulation mandates by the American Board of Orthopedic Surgery and the Orthopedic Residency Review Committee of the ACGME.[236-238] The FAST Program was designed to satisfy these educational mandates for arthroscopic surgery with a cost-effective, practical, modular system.

In the traditional approach to training, operative skills were acquired in an apprentice model of training that meant that learning was serendipitous. Resident experience was affected by when they were on duty, which patients and procedures they encountered, and who was supervising and mentoring them. It also relied on learning by repeated practice.[239] The proficiency-progression approach to training, afforded by technologies such as FAST, encourage a 'deliberate practice' approach.[79, 240] This means the trainee receives objective metric-based feedback on THEIR performance proximate to the measured task, thus augmenting the learning experience for them. Seeing knots slip when pulled is a very impactful learning experience, even for very experienced surgeons.

For decades, orthopedic educators have been using knot-tying boards with rope and suture to train arthroscopic knot tying. In most cases, proficiency assessment has been based upon subjective, visual observation of the knot tying process and the visual appearance of the surgical knot.[241-251] However, what is most important, in terms of surgical outcome, is knot *performance* as opposed to knot *appearance*. A "pretty" knot has no clinical value if it doesn't hold under physiologic loads.

Prior to the current study, biomechanical assessment of sutures and knots has generally been restricted to analyses with sophisticated material testing devices (MTS, Instron, etc.). These devices are quite expensive and they are impractical for day-to-day training applications. However, they do have advantages for complex load application paradigms, including cyclic load protocols. Nonetheless, we thought that it would be advantageous to create a very simple and inexpensive knot tester that could be used on the educational front lines. The FAST knot tester was not designed to be a sophisticated bioengineering research tool.

It should be emphasized that the level of load application (15 pounds) for the FAST knot tester was specifically selected for testing of high strength #2 sutures used commonly during many arthroscopic procedures. The knot tester could be adapted with different spring-loads in order to assess performance of other suture materials. Objective proficiency benchmarks could be established for various sutures under specific performance conditions. This strategy is relevant to training and objective assessment of surgical knot-tying across the medical spectrum, since knot tying is a pervasive technical skill requirement for most procedural specialties. The FAST approach to measurement and quality assurance, by achieving performance benchmarks before training progression, fits well with the recommendations of the Institute of Medicine (IOM) report on Graduate Medical Education.[252] The IOM proposed that medical education should move away from training that is process driven (i.e., time in training, number of procedures completed, duration of rotation) to an 'outcome' driven enterprise.[96] This means that trainees would be required to demonstrate a benchmark performance level.[201]

We were quite surprised by the high incidence of knot failures in the current study, for course faculty and surgeons in practice. Subsequent to our data collection, Hanypsiak and co-workers[223] published similar observations in their study of 73 expert orthopedic arthroscopists who tied 365 individual knots with #2 Fiberwire suture. In their study, surgeons created knots under direct visualization, without magnification or video control, and the knots were tested using a sophisticated electromechanical dynamic testing system. The authors observed significant variation between surgeons and between knot configurations. Perhaps even more important, they concluded that "considerable variation and inconsistencies in knot strength exist between arthroscopic knots of the same type tied by the same surgeon."[223] Individual subject performance inconsistency was also noted in the current study for knots created by experts under video control through arthroscopic cannulas. The observations of Hanypsiak et al[223] and the findings of the current study are extremely important, because technical consistency is a hallmark of surgical proficiency and patient safety.

We used the performance data for our expert faculty surgeons to create objective proficiency benchmarks that could be applied to resident and fellow training. Proficiency benchmarks must be reasonable and achievable. It wouldn't make sense to set benchmarks that are unachievable for a high percentage of competent, experienced surgeons. Given the relatively high incidence of knot failure for our experienced surgeons, we defined a proficiency benchmark of no more than two out of five knot failures to achieve a "passing" score for residents. Of course, surgeons should strive for technical perfection, with zero knot "failures", and we observed that level of high performance for some of our expert subjects. However, our data suggests that many arthroscopic surgeons (even experienced and expert surgeons) have substantial opportunities for improvement. Such opportunities are facilitated by direct, objective, and immediate performance feedback.

Based upon the Copernicus faculty data, we defined a proficiency benchmark (a "passing grade") of less than or equal to two knot failures out of five knot attempts. If the threshold had been set to no more than one failure in five attempts, eight of our own Copernicus faculty (40%) would have fallen below the "passing" bar (see Table 6.1). We thought it was important to avoid unrealistic and/or unachievable proficiency benchmarks for the residents, so we selected the more lenient proficiency standard of no more than two knot failures out of five knot attempts. For the surgeons in practice at the Fall Course (Table 6.2), if the threshold was set at less than or equal to one knot failure in five attempts, nine of thirty surgeons in practice (30%) would not have "passed". These data further support use of the more lenient proficiency standard for training purposes.

Overall, the knot failure rate was 26% for the 220 knots that were tied by the orthopedic surgery residents (Table 6.1). Surprisingly, this failure rate was not dramatically different than the overall failure rate for our faculty and surgeons in practice. However, the Group C sub-group (who were allowed to use the knot tester for feedback *during* the training experience) had an overall 11% knot failure rate, which was significantly better than the Copernicus Course faculty ($p = 0.02$). We were impressed by the strong performance of the Group C proficiency-progression sub-group (Table 6.1). These residents could assess their performance based upon direct proximate feedback, make adjustments in knot tying technique, and then see for themselves whether their performance had improved. That approach appears to have resulted in substantial enhancement of this group's performance.

Based upon our observations and the recent findings of Hanypsiak et al,[223] we believe that it would be very challenging to objectively assess knot performance using video review of arthroscopic procedures. Some overt suture failures are easily observed. For example, it is visually obvious when sutures break or entangle, or when there is an overly loose suture loop that doesn't indent soft tissue. However, our findings suggest that some "visually acceptable" knots may fail under relevant mechanical loads, even in the hands of very experienced surgeons.

6.41 Limitations

One of the limitations of the present study was that we did not afford opportunities for self-directed performance feedback to our faculty surgeons or to the practicing surgeons at the AANA Fall Course. This study did not involve a homogeneous population of faculty and practicing surgeons, with prior clinical experience ranging from one to forty years of practice. We did not assess transfer of motor skills to the clinical situation, nor did we examine same-subject test/retest consistency. Maximum practice time prior to knot testing was not a controlled variable. This study was not designed to compare performance differences according to knot type, because we wanted each subject to select their own base knot based upon personal preference and experience. Previous research has looked at biomechanical performance variation as a function of knot type, and it was not our purpose to examine this question. During the study design phase, the authors recognized and discussed the implications of variation of base knot by each study participant. We wanted our subjects, particularly our faculty and experienced surgeons, to pick the base knot that he/she would be most comfortable tying. We did not want to impose a particular knot choice because we were concerned that individual performance could be adversely affected by asking subjects to tie knots that they were unaccustomed to, thereby introducing greater data variability. The current study was not designed to cross-correlate knot performance with knot "appearance". These are important study limitations and represent opportunities for further work.

6.5 Conclusions

The FAST workstation and knot tester offer a simple and reproducible educational approach for enhancement of arthroscopic knot-tying skills. Load-displacement of the suture loop is a direct reflection of mechanical performance of the surgical knot. There is significant room for improvement of the quality and consistency of this important arthroscopic skill, even for experienced arthroscopic surgeons.

Table 6.1

Knot Tying at the AANA Resident's Copernicus Course

Knot Failure defined as $\geq$ 3 mm loop expansion with 10 seconds of 15# load

	Faculty (n = 20)	Group A Residents (n = 14)	Group B Residents (n = 14)	Group C Residents (n = 16)
Years in practice, mean ± SD (range)	19.9 ± 8.3 (4-32)			
Knot performance, n				
0 of 5 failed	12	3	3	11
1 of 5 failed	0	3	5	2
2 of 5 failed	4	7	2	2
3 of 5 failed	2	0	1	1
4 of 5 failed	0	1	1	0
5 of 5 failed	2	0	2	0
No. of knots that failed	24 of 100 (24%)	21 of 70 (30%)	26 of 70 (35%)	9 of 80 (11%)

NOTE. Knot failure was defined as 3 mm of loop expansion or greater with application of 15 lb of load for 15 seconds.

Table 6.2

Knot Tying at the 32nd AANA Fall Course

Knot Failure defined as $\geq$ 3 mm loop expansion with 10 seconds of 15# load

	Faculty (n = 10)	Surgeon Attendees (n = 12)	Faculty or Attendee Not Defined (n = 8)	Total for All Participants (n = 30)
Year in practice, mean ± SD (range)	20.5 ± 7.6 (3-30)	14.6 ± 12.4 (1-40)	19.0 ± 9.3 (9-32)	17.1 ± 19.3 (1-40)
Knot performance, n				
0 of 5 failed	5	3	4	12
1 of 5 failed	3	3	3	9
2 of 5 failed	1	3	0	4
3 of 5 failed	1	3	1	5
4 of 5 failed	0	0	0	0
5 of 5 failed	0	0	0	0
No. of knots that failed	8 of 50 (16%)	18 of 60 (30%)	6 of 40 (15%)	32 of 150 (21%)

NOTE. Knot failure was defined as 3 mm of loop expansion or greater with application of 15 lb of load for 15 seconds.

7.1 "Proficiency-Based Progression Training Curriculum Coupled With a Model Simulator Results in the Acquisition of a Superior Arthroscopic Bankart Skill Set" Angelo RL, Ryu RKN, Pedowitz RA, Beach W, Burns J, Dodds J, Field, Getelman M, Hobgood R, McIntyre L, Gallagher AG. Arthroscopy 2015; 33:1854-1871. Appendix 9

(Candidate is the first and primary author)

Abstract

Purpose: To determine the effectiveness of proficiency based progression (PBP) training using simulation compared to both the same training without proficiency requirements and to a traditional resident course for learning to perform an arthroscopic Bankart repair (ABR).

Methods: In a prospective, randomized, blinded study, 44 PGY 4 or 5 orthopedic residents from 21 approved US orthopedic residency programs were randomly assigned to one of three skills training protocols: Group A: Traditional (control, N = 14)(routine AANA Resident Course); Group B: Simulator (N = 14)(modified curriculum adding a shoulder model simulator); or Group C: Proficiency Based Progression (N=16)(PBP plus the simulator) for learning an ABR. At the completion of training, all subjects performed a 3 suture anchor ABR on a cadaver shoulder, which was videotaped and scored in blinded fashion using previously validated metrics.

Results: The PBP trained Group C made 56% fewer objectively assessed errors than the traditionally trained Group A (p = 0.011) and 41% fewer than Group B (p = 0.049)(both comparisons significant). 68.7% of those in Group C achieved the proficiency benchmark on their final repair compared to 36.7% of Group B and 28.6% of Group A. Group B participants were 1.4 times as likely, Group C 5.5 times, and Group C^{PBP} (who met all intermediate proficiency benchmarks) were 7.5 times as likely to achieve the final proficiency benchmark than Group A.

Conclusions: A PBP training curriculum and protocol, coupled with the use of a shoulder model simulator and previously validated metrics, produces a superior arthroscopic Bankart skill set when compared to traditional and simulator enhanced training methods.

Clinical Relevance: Surgical training combining PBP and a simulator is efficient and effective. Patient safety could be improved if surgical trainees participated in PBP training employing a simulator before treating surgical patients.

7.1 Introduction

Changing work patterns and a reduction in hours available for training[217, 253] have forced the surgical community to consider new methods to augment and enhance training. Surgical simulation-based training, first proposed by Satava in 1993[26] as a potential solution to this problem, has developed in sophistication and adoption amongst medical education and training communities.[204, 254] The first prospective, randomized, double-blind clinical trial of simulation-based training for the operating room demonstrated that surgical residents trained to a "proficiency benchmark" (Table 3.1) on a virtual reality simulator made significantly fewer objectively assessed intra-operative errors (Table 3.1) when compared to the control group.[22] The reader is referred to Table 3.1 for a list of terms used throughout this article. Gallagher et al.[201], and Gallagher and O'Sullivan,[79] have argued that simulation based training is optimal when trainees are given precise feedback on their performance with specific recommendations for improvement, proximate to the performance. They have also suggested that trainees be provided a quantitative performance benchmark to work toward and that this benchmark should be a valid representation of a clinically important performance characteristic or task. Thus, trainees must demonstrate the ability to meet specific performance benchmarks before they are permitted to progress in training (proficiency-based progression training [PBP]") (Table 3.1). The effectiveness of this methodology is well supported.[22, 198, 208]

We sought to study the effectiveness of proficiency-based progression (PBP) training plus simulation for the acquisition of surgical skills. For the patient with unidirectional anterior instability due primarily to a Bankart lesion (capsulolabral detachment from the anteroinferior glenoid) without significant bone loss, a suture anchor repair employing 3 implants is a commonly accepted method utilized to obtain a successful patient outcome.[188, 190, 255-259] In addition, the essential components of the procedure are well outlined regardless of whether the patient is placed in the lateral decubitus or beach chair orientation.[89, 193] Thus, an arthroscopic Bankart repair was selected as the platform for this research.

The investigation into PBP training plus simulation required the development and validation of 3 separate, specific tools to conduct the analysis. The first component to be created was a "**metric tool"** (Table 3.1), which could objectively and accurately characterize an ABR, by clearly defining the essential "steps", "errors", and "sentinel errors" (more serious) (Table 3.1) for a standard reference repair. The metric tool created was demonstrated to have "face" and "content validity"[260] (Table 3.1) using a "modified Delphi panel" (Table 3.1) methodology.[65] Second, a "**training tool"** (a shoulder model simulator coupled with the ABR metrics) was shown to have "construct validity" (Table 3.1), demonstrating the ability to distinguish between novice and experienced surgeon performance. A proficiency benchmark for the use of the metrics with the simulator was established.[93] Lastly, an "**assessment tool"** (a cadaver shoulder coupled with the ABR metrics) was evaluated and also shown to have "construct validity" (Table 3.1).[66]

The purpose of this study was to determine the effectiveness of proficiency based progression training using simulation in comparison to both the same curriculum without the proficiency requirements as well as to a traditional AANA Resident Course for learning to perform an arthroscopic Bankart repair. We hypothesized that a training protocol coupling proficiency based progression training with a shoulder model simulator would be superior to an identical curriculum using a simulator, but without the need to demonstrate proficiency as well as to a traditional curriculum with no simulator or proficiency requirements.

7.2 Methods

7.21 Participants/Subjects

44 PGY (Post Graduate Year) 4 or 5 residents from 21 ACGME (Accreditation Council for Graduate Medical Education) approved orthopedic residency training programs from across the US participated. All subjects were assigned a unique identifying number which gave no indication of their post-graduate year, residency program or study group. The Western Institutional Review Board (WIRB) opined (#1-776362-1) that as an educational curriculum study, it was “Exempt” from the need for full IRB approval (criteria: 45 CFR 46.101(b)(1)). The study protocol was registered with the NIH (ClinicalTrials.gov: #NCT01921621) prior to initiating the investigation.

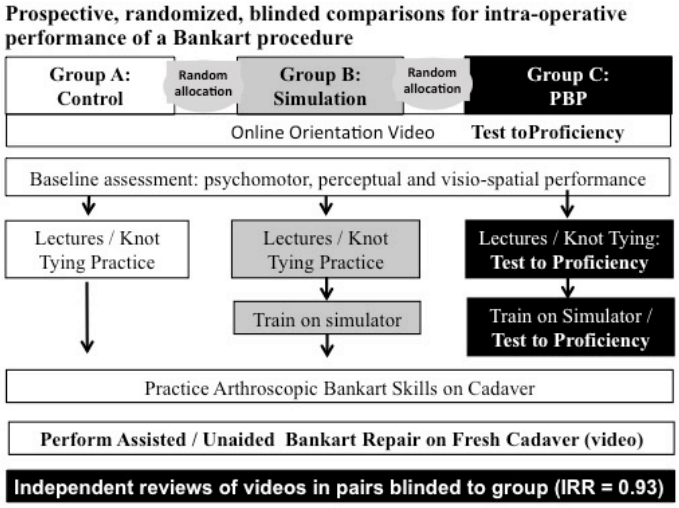


Figure 7.1 Study pathways for 3 separate training protocols, Groups A (traditional training - control), Group B (simulation enhanced), and Group C (proficiency-based progression training plus simulation); All groups underwent baseline assessments followed by lectures on knot tying, but only group C was required to pass a cognitive exam on the metrics and to test to proficiency on knot tying; Groups B and C trained to perform an arthroscopic Bankart repair (ABR) on a model simulator, but only Group C had to test to proficiency; All 3 groups trained / practiced an ABR on a cadaver; All participants from each group performed an unaided ABR on a fresh cadaver shoulder which was videotaped and scored in blinded fashion.

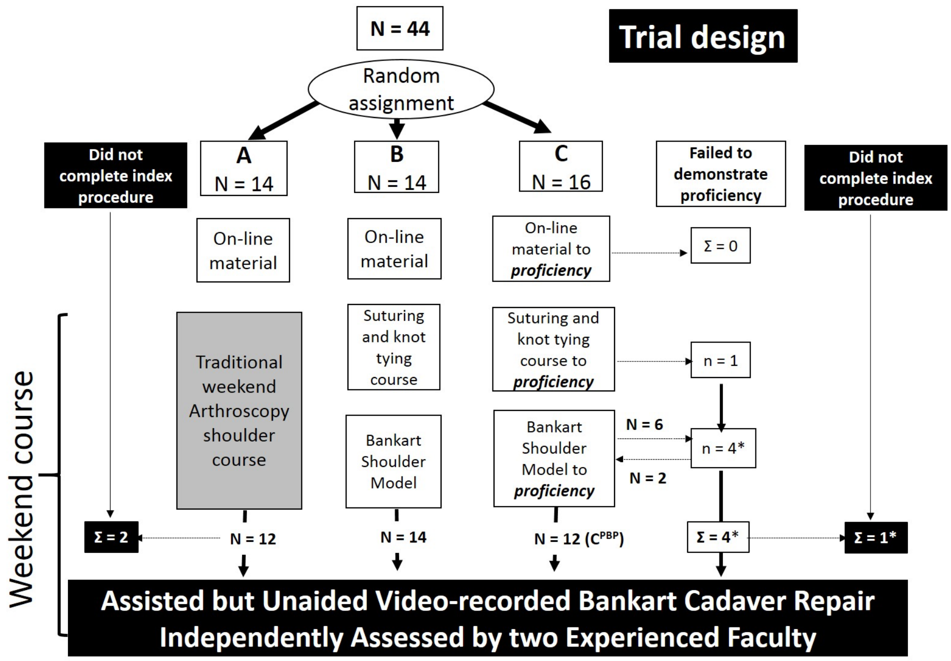


Figure 7.2 Participant Flowchart: Shows participant flowchart detailing that 2 of 14 participants in Group A were unable to complete the procedure; one Group C subject could not complete knot tying to proficiency, and 4 Group C resident were unable to demonstrate proficiency on the shoulder model simulator. The Group C subject who failed both the knot tying and model proficiency also failed to complete the final Bankart repair. The remaining 12 subjects from Group C, designated Group C^{PBP} who met all intermediate proficiency benchmarks would be the only ones allowed to progress per the PBP *protocol*.

7.22 Bankart procedure performance metrics

The surgical residents were evaluated on their skill in performing an arthroscopic Bankart repair on a cadaver specimen. Previously validated "performance metrics" (Table 3.1) formed the basis of that evaluation and include 45 key steps with related steps grouped into one of 13 "phases" (Table 7.1) of the procedure[65] (Table 3.2). 77 potential errors to be avoided were specified (Table 3.3). 20 of those errors were designated as more serious, or "sentinel" errors either because, 1) their enactment had the potential to seriously compromise the success of the procedure, or 2) the error had the potential to create significant iatrogenic damage to the shoulder. The metrics were clearly defined with beginning and endpoints for each step, as well as precisely, what did and did not constitute each potential error. All subject surgeons and faculty were provided a link on the AANA website to 2 full-length orientation videos, one each in the lateral decubitus and beach chair orientations. Access was available four weeks prior to the course in which they were participating. Each video demonstrated all of the steps, in addition to either demonstrating or specifically identifying each of the potential errors (including sentinel errors) to be avoided in performing an arthroscopic Bankart repair safely.

Figure 7.3 AANA F.A.S.T. (Fundamentals of Arthroscopic Surgery Training) workstation with multiple potential portal sites in an opaque dome covering a mandrel around which knots are tied via an arthroscopic cannula (inset picture on the laptop screen). A USB camera (IPevo – Sunnyvale) directed through a window, projects an image of the inside of the dome onto a laptop computer screen.

7.23 Baseline Assessments

To ensure homogeneity among the 44 subjects, all residents completed previously validated assessments of their visuo-spatial,[22, 261] perceptual,[22, 262], and psychomotor[75, 263, 264] abilities (Table 7.1). Visuo-spatial ability is one component of cognitive function that is related to the capacity to process and interpret visual information about where objects are in space. In this assessment, a pencil is used to create the shortest and most appropriate route between two specific points on a block grid street map. A possibility of 20 correct routes between various points existed for each test. The number of correct routes created in a 2-minute time period was scored. Each registrant completed 2 tests. Perceptual ability refers to the capacity to identify, organize, and interpret sensory information about visual depth of field. It is assessed with a computer generated and scored task requiring the subject to orient the axis of a spinning cone perpendicular to a designated face of a cube. Each of 30 trials places the cube in a different 3 dimensional orientation. Psychomotor ability refers to the capacity for coordinated activity involving the arms, hands, fingers, (and potentially movement of the feet). Performance was assessed using a lighted endoscopic box trainer with a fixed overhead view projected onto a laptop screen. A 4 X 8 inch piece of paper had a series of 1 inch long parallel lines drawn perpendicular to and along the long boarder of the sheet. Each of 30 parallel lines was separated by 10 mm. Instruments were passed through openings in the front of the box trainer. An endoscopic grasper was controlled by one hand and used to hold the paper within the box. Endoscopic scissors were controlled by the other hand and used to make cuts in the paper between the designated lines. The number of accurate paper cuts (between, but not touching parallel lines) able to be made in 60 seconds was tabulated. Two trials were run, one with the scissors in the dominant, and the other with them in the non-dominant hand.

7.24 Study Groups

During the weekend courses, all groups were provided similar background shoulder instability lectures which focused on indications, contra-indications, and case-based examples. References to surgical technique were avoided in the lecture presentations. Each of the three groups had separate, dedicated, experienced Master and Associate Master AANA faculty members who worked closely with that cohort of residents. The duration of training was similar for each of the three groups. The 3 training curriculums are outlined in Figure 7.1. All training was conducted at the Orthopedic Learning Center (OLC) in Rosemont, IL.

Group A (Traditional - Control)

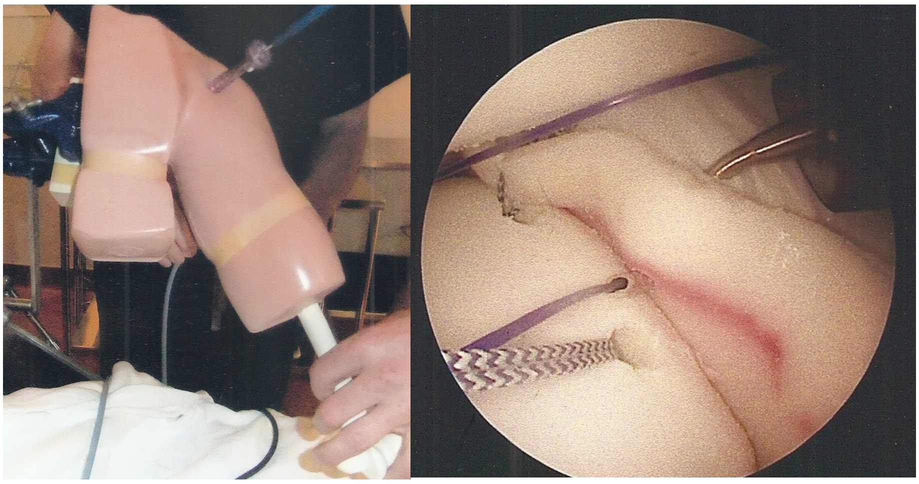
Group A was derived from a cohort of PGY 4 and 5 residents who had independently registered for a 3 day AANA Resident course at the Orthopedic Learning Center (OLC – Chicago). The curriculum involved equal time spent on the practice of knee and shoulder procedures, which included an arthroscopic Bankart repair. The registrants of the course were given the opportunity to be involved in the study, but were not required to do so. A cohort of 14, of a total of 48 residents registered, elected to participate in the research project and complete an arthroscopic Bankart repair at the end of their course of training.

The 14 subjects in Group A served as the control group and followed an agenda typical for an AANA Residents' Course (Figure 7.2). The curriculum included lectures on various topics including shoulder instability. Knot tying skill was practiced under the direction of an experienced faculty member. Both sliding and non-sliding knots as well as ½ hitches to secure the primary knot were practiced. Knot tying boards provided the opportunity to tie knots around hooks using large cord and / or #2 Fiberwire suture (Arthrex, Naples). In addition, practice tying suture knots and delivering them down an 8.5 mm arthroscopic cannula was afforded. Finally, knots could be created and delivered through a cannula with the loop around a mandrel (smooth bar) using the AANA FAST Workstation (Fundamentals of Arthroscopic Surgery Training) (Figure 7.3). The opaque dome eliminated direct surgeon view of the knot being tied and required the use of triangulation skills. A small USB camera was directed through a window in the dome with the image viewable on a laptop computer screen. The field of view contained the knot being tied. The subject surgeon had the opportunity to spend as much time practicing knots as desired and until they believed they were proficient. A series of 5 knot trials were then tied and labeled in sequence for later testing.

Study in the lab using a cadaver specimen then followed under the direct supervision of an experienced faculty member with knowledge of the Bankart metrics. Fresh frozen specimens with a complete shoulder girdle from the scapula to the mid-humerus with associated soft tissues were used. After appropriate thawing, the scapula was mounted with a clamp in the surgeon's orientation of preference (lateral decubitus vs. beach chair). Bony landmarks were identified and marked with a surgical pen, portals established, and a diagnostic arthroscopy performed. An arthroscopic Bankart repair was generally the first intra-articular procedure studied. After a Bankart lesion was created, practice was conducted in the steps necessary to mobilize the capsulolabral tissue and complete a 3 anchor repair. Standard instruments for an arthroscopic Bankart repair were made available. A 45^0 cannulated suture hook was the primary tool used to deliver sutures through the capsulolabral tissue. Single-loaded push-in anchors and a simple loop suture pattern were employed. The resident was able to continue with guided study on the Bankart repair as long as desired and until they believed they were proficient. Subsequently, additional arthroscopic shoulder procedures could be electively studied as well.

Group B (Simulator)

Multiple, randomly selected residency program coordinators were notified of the opportunity for their residents to participate in this proficiency based progression training study. The first 32 PGY 4 and 5 residents to register became study participants. Training for Groups B and C was conducted concurrently at the OLC, but on a different weekend from Group A. Of the 32 pre-registered subjects, 16 were randomly assigned to one of two training protocols (Group B or C) based on a computer generated random allocation[265] (Figure 7.2). Of the 16 residents randomized to Group B, 2 who were pre-registered failed to show up for the week-end course, thus, N = 14 for Group B. The residents in Group B engaged in knot tying study and practice similar to that for Group A until they believed they were proficient. A series of 5 knot trials were then tied and labeled in sequence for later testing. Group B participants were afforded the additional opportunity to train and practice an arthroscopic Bankart repair using a dry shoulder simulator (Figures 7.4 A & B) secured in the orientation of surgeon preference.[93] A standard equipment tower with a 30^0 arthroscope was provided along with all instruments necessary to complete an arthroscopic Bankart repair. The model was composed of a dense plastic endoskeleton palpable through simulated skin and soft tissues. Posterior, anterosuperior, and mid-anterior portals were created. A glenoid, humeral head, biceps, capsule and labrum, in addition to Bankart and Hill-Sachs lesions, were present and provided the opportunity to complete all of the steps demonstrated on the orientation videos for an arthroscopic Bankart repair. Work continued on the simulator as long as desired and until the resident believed they were proficient with the steps and sequences of the capsulolabral repair. Further study and the guided practice of the steps for an arthroscopic Bankart repair on a cadaver specimen then followed and continued as long as the participant desired and until they believed they were proficient.



A **B.**

Figure 7.4A: Exterior view of the anterior aspect of a left shoulder model simulator

supported in the beach chair orientation; hook probe is delivered through a midanterior portal. **Figure 7.4B**: Arthroscopic view from the anterosuperior portal of a left shoulder in the lateral decubitus orientation; the second anchor is in place with its sutures coursing toward the posterior portal; a suture hook device passes through the midanterior portal pierces the labrum and delivers a monofilament suture (to be used later for shuttling).

Group C (Proficiency-Based Progression)

All residents randomized to Group C attended the course (N = 16) and were exposed to a protocol identical to that of group B with the additional requirement of demonstrating proficiency at various stages of the training (Figure 7.2). Each of the individual proficiency benchmarks for the procedural components was established based on the mean performance of separate groups of experienced surgeons on the specific exercise.[22, 66, 79, 93, 198, 201, 208] After arriving at the Orthopedic Learning Center, if the registrant had not yet taken and passed the cognitive test online covering the validated metric steps and errors demonstrated on the orientation videos, they were required to do so on site. A minimum score of 84% was required to pass. Those who initially failed continued to study the material and were provided additional faculty instruction.

Knot tying study progressed in a manner similar to Groups A and B. Once the resident believed they had mastered the knot tying skills, they had the opportunity for the integrity of their knots to be tested using the FAST Workstation knot tensiometer[35] if desired. To pass, the loop / knot construct had to elongate < 3 mm when subjected to a static load of 15# for 15 seconds. The benchmark was set at a minimum of 3 / 5 knots meeting this standard. Once the formal testing process began, the subject tied 5 knot trials, which were labeled in sequence. All 5 were then tested using the tensiometer. Those subjects, who failed, continued to practice until confidence was gained, and then the testing sequence was repeated with a new series of 5 knots being tied. This process continued until the resident achieved proficiency or was unable to do so and failed to demonstrate progressive improvement in the knot tying skill set.[35]

Work and practice then began with the same shoulder model simulator used by Group B. The model was oriented according to physician preference. Landmarks were identified and posterior, midanterior, and anterosuperior portals were established. After a diagnostic examination was performed, the steps for a 3 anchor Bankart repair were practiced. The faculty instructors provided proximate feedback and recommended corrections based on the previously defined step and error metrics demonstrated on the orientation videos. Practice and faculty feedback continued through a complete procedure until the subject and their faculty instructor both believed they had adequately prepared for testing.

A new simulator model was then oriented in the resident's position of preference. Equipment representatives from multiple different vendors served as surgical assistants and were randomly assigned to participating surgeons. The assistants were instructed to act only at the specific direction of the operating surgeon. Prompting and coaching (of technique) were prohibited (the procedures were proctored by staff from the Orthopedic Learning Center). The resident surgeon then proceeded to complete a diagnostic evaluation, and perform a 3 anchor arthroscopic Bankart repair attempting to mimic the key steps identified on the orientation videos. Using the Bankart metric score sheet, one of six designated faculty members, intimately familiar with the Bankart metrics, scored the subject in real time during the arthroscopic repair on the simulator model. The benchmark for a passing score in the shoulder model was established from a prior study[65] and included a 3 anchor repair with no more than 4 total errors, and no more than 1 sentinel error. If a resident failed to meet the benchmark, the faculty who scored the model test, the assigned training instructor, and the resident all conferred together to identify the specific deficiencies exhibited and the appropriate corrections. The subject then worked toward acquisition of the requisite skills with instructor guidance. When confident, they were given 1 additional opportunity to repeat the scored procedure on a new model. In a normal proficiency based progression protocol, residents who failed to meet each of the intermediate proficiency benchmarks would not be allowed to progress in training and would require additional practice until the deficiencies were corrected (and would not have been allowed to progress on to working with the cadaver). However, given the artificial finite time constraints of the study weekend, all Group C participants, regardless of persistent deficiencies, were allowed to proceed on to practice with a cadaver specimen and guided instruction similar to Groups A and B.

7.25 Final Videotaped Bankart Repair Assessment

At the completion of their respective courses, the subjects from each group performed an assisted, unaided arthroscopic diagnostic survey and a 3 anchor Bankart repair on a fresh cadaver shoulder. The cadaver specimens were considered acceptable if: 1) arthroscopic visibility of the target tissues was obtainable; 2) the specimen (flexibility) permitted adequate access to the target tissues; and 3) the integrity of the capsulolabral tissues was sufficient to permit mobilization, suture delivery and knot-tying. All necessary instrumentation and implants were made available. Residents participating in the course served as assistants for each other. They were instructed to act only at the request of the operating surgeon and were prevented from coaching or prompting. The OLC staff proctored the procedures for compliance.

The procedure was videotaped in its entirety beginning with the initial view from the posterior portal. The resident surgeon mapped the bony landmarks and then created their preferred portals. All or a portion of the diagnostic exam was completed. The arthroscope was withdrawn and a red card videotaped for 5 seconds to signal that the subject surgeon was no longer operating. One of 4 designated faculty members then reintroduced the arthroscope and, using a sharp elevator from either the anterosuperior portal, midanterior portal, or both, created a standard Bankart lesion, 6 – 9 mm deep (medial from the bony rim) and from 2 – 6 o'clock along the glenoid. Once the Bankart lesion was created, care was taken to avoid additional mobilization of the capsulolabral tissue. The arthroscope was then withdrawn and a green card videotaped for 5 seconds signaling that the subject surgeon was operating for the balance of the procedure. The arthroscope was reintroduced into the glenohumeral joint by the subject surgeon, and any remaining elements of the diagnostic survey completed. A 3 anchor Bankart repair was then performed attempting to mimic the steps demonstrated in the orientation video and practiced in the simulation model.

All subject surgeons used identical implants: single-loaded (2.8 mm) Gryphon push-in anchors loaded with #2 Orthocord (DePuy Mitek, Raynham). A 45^0 cannulated suture hook was used to deliver a shuttling device with retrograde passage of the anchor sutures through the capsulolabral tissues. Prior to beginning work on their final scored Bankart repair, instructions were given to all residents regarding the protocol for anchor pullout from cadaver bone. If an anchor failed prior to completion of the index sliding knot, the surgeon was permitted to remove the anchor and suture, and replace it with a metal 5.5 mm screw-in anchor (Smith and Nephew, Andover). The procedure then continued with no penalty. The time required for the reintroduction of the screw-in anchor and re-passage of the anchor suture through the capsulolabral tissue was subtracted from the total procedure time. If the anchor failed subsequent to the initial sliding knot being completed (i.e., efforts to back up the primary knot with half hitches), the surgeon was instructed to abandon the first anchor position and proceed on to the second anchor position. No time limit was imposed on the performance of the Bankart repair and each participant was able to continue to work as long as they believed they were making progress. At the point in the procedure that the subject surgeon did not believe they could make further progress in the Bankart repair, they could electively choose to terminate the procedure.

7.26 Video Reviewer Training

Once the construction of the metrics for an arthroscopic Bankart repair was completed and face and content validity verified[65], a final version of a score sheet was formatted. 10 AANA Master / Associate Master faculty surgeons formed the panel of reviewers designated to score the videos. This group included the 3 arthroscopic surgeons (RLA, RKNR, RAP) who, in conjunction with a consultant experimental psychologist

(AGG), developed the arthroscopic Bankart metric "definitions" (Table 3.1). The ten reviewers were randomly assigned to form five fixed pairs, which remained constant throughout the scoring of all videos. Reviewer training was initiated with an 8-hour in-person meeting during which time each metric was studied in detail. Multiple video examples of live patient cases were shown to illustrate each particular metric. Videos of the patients in both the lateral decubitus and beach chair orientations were represented. Discussion helped to clarify how each step and error was to be scored, including the nuances and conventions to be used. Several weeks later, full-length practice videos # 1 and 2 (one each in the lateral decubitus and beach chair orientation) were sent to and independently scored by each of the 10 reviewers, and the scores tabulated. In two subsequent 2-hour group phone conferences, the differences and discrepancies amongst all reviewers were compared and discussed seeking conformity in scoring. In addition, each designated pair of reviewers conducted one to three additional phone conferences to analyze the specific instances in which the two of them scored particular events differently. Subsequently, all reviewers scored practice videos #3 and #4 and the results were tabulated (each patient orientation again represented). The scores for each of the 5 designated pairs of reviewers were compared for the second set of practice videos. In only 1 of 10 comparisons (2 videos X 5 reviewer pairs) did the "inter-rater reliability" (Table 3.1) calculation (see below) fall below an acceptable level[79] of 0.8 at 0.76.

7.27 Video Scoring

The AANA research coordinator randomly assigned the 44 full-length study videos, each with only the designated unique identifying number attached, to a single pair of reviewers. Other than the research coordinator and the study consultant, all video reviewers remained blinded to the source of the video being reviewed. Each video was independently reviewed and scored by the two members of an assigned pair of reviewers. All scores were tabulated for each of 13 phases of the procedure. Each step and error metric was scored as either a "yes" or "no", designating whether the specific event was or was not observed to occur by the reviewer. In addition to scoring steps and errors, each event characterized as "damage to non-target tissue" (DNTT) (Table 3.1), was scored (i.e. gouging the articular cartilage; or tearing of the capsule). There was no limit to the number of individual instances DNTT could be scored, with each occurrence tallied as a single error event. The score sheet also contained a box for specific reviewer comments for each metric.

7.28 Score Tabulation

For each of the 13 separate phases of the procedure, the # of 'uncompleted steps' and 'errors made' were tabulated and the scores for the 2 reviewers averaged. Further, for each subject, the step and error data was pooled for the three repetitive components of the procedure: 1) anchor preparation, 2) suture passage / management, 3) knot tying. These data were used to determine which of the procedural phases demonstrated the greatest differences in performance among the groups (one factor -ANOVA analysis; IBM SPSS statistical software program). Further, for the entire procedure, the total number of steps 'completed', errors made, and sentinel errors enacted were also averaged for the pair of reviewers. The subject's operative time was obtained by subtracting the faculty time to create the Bankart lesion from the total absolute recording time for the procedure.

The two raw score sheets from the designated pair of reviewers were compared for each of the individual steps (N = 45) and the number of 'agreements' tabulated (either both reviewers documented that a step was performed, or both scored the step as not being completed). In addition, the number of 'disagreements' in scoring steps was tabulated (one of the reviewers indicated that the step had and other scored that the step had not been completed). The inter-rater reliability for the steps was calculated according to the following formula:

$$\frac{\textbf{Agreements}}{\textbf{Agreements + Disagreements}}$$

In a similar manner, there was either agreement or disagreement in the 2 scores for each of the potential errors (N = 77). The IRR for error scoring was calculated in the same manner as that for the steps. Finally, the IRR for scoring the entire procedure was calculated using both the step and error agreements / disagreements for the complete procedure (N = 122). Acceptable IRR is = or > 0.80.[79]

7.29 Statistical Methods

The analysis was conducted as a series of multiple regressions. The exogenous variables (covariates) were the three intervention conditions, i.e., proficiency based progression plus simulator (Condition C), simulator (Condition B), and traditional training (Condition A). Group C was used as the reference condition within the analysis. As a check on the veracity and stability of the results, all of the analyses were also conducted using Poisson regression. The substantive interpretation remained unchanged regardless of the model used. All of the reported results are based on the analyses from the multiple regressions. Further, a logistic regression analysis was performed to estimate the probability of those trainees from the different training curriculums being able to attain the proficiency benchmark for the final arthroscopic Bankart repair.

A secondary analysis was conducted to evaluate the subset of Group C subjects who successfully met all of the intermediate benchmarks throughout training, and was designated Group C^{PBP}. Group C^{PBP} was evaluated for the same performance metrics as the other groups, which included steps, errors, sentinel errors, and time as well as the probability of attaining the benchmark on the final repair. All of the participants in Group C followed the proficiency-based progression '*curriculum'*. The proficiency-based progression *'protocol'*, in distinction however, only permits those individuals who meet each proficiency benchmark to progress in training (C^{PBP}) (Figure 7.2).

7.3 Results

The mean and standard deviation scores on the baseline assessments of perceptual, visuo-spatial and psychomotor performance are shown in Table 7.1. Although Group A performed somewhat better on the psychomotor test than Groups B and C, these differences were not statistically significant.

7.31 Intermediate Proficiency Training Benchmarks for Group C

All of the 16 participants in Group C were able to obtain a passing score on the cognitive exam, although several required additional instruction after failing to achieve a passing score of 84% on their initial test. One subject from this group was unable to demonstrate proficiency in knot tying despite repeated training and practice. Six Group C participants failed their first attempt to meet the benchmark for a Bankart repair on the simulator model. After additional guided training and practice, 2 of the 6 were able to show proficiency on their second attempt with the shoulder model. One of the 4, who were unable to demonstrate proficiency on the model, was also the one who failed to show proficiency at knot tying. Thus, 12 of 16 Group C subjects met all of the intermediate benchmarks during training. Based on the PBP protocol, these 12 (Group C^{PBP}) would have been the only participants from Group C allowed to progress to working on the cadaver.

7.32 Final Cadaver Bankart Assessment

Two cadavers, one each from Group B and C failed to meet the acceptability criteria and were replaced with better specimens. The video recording was restarted with the onset of work on the replacement specimen. Of 44 videos scored, the mean inter-rater reliability for the total number of steps performed and errors made was 0.93 (0.84 – 0.99; standard deviation = 0.04).

7.33 Incomplete Final Procedures

Of all 44 subjects, only three failed to complete their final Bankart repair on a cadaver shoulder. Two individuals from Group A were only able to finish the first anchor with an average of 16.25 steps completed, 7 errors and 0.5 sentinel errors enacted. They worked for an average of 99 minutes. In Group C, one subject had the first anchor pull out during efforts to deliver and secure the primary suture knot and the individual elected not to replace that anchor (although they could have done so according to the anchor pullout protocol). They completed all of the 2nd and 3rd anchor components, thus performing only a 2 anchor final repair, which was deemed to be incomplete. During this procedure, an average of 37 steps were completed, along with 4 errors, and 0 sentinel errors. Operative time was 92 minutes. This subject was the one who had previously failed to demonstrate proficiency on both knot tying and on the shoulder model repair components of the training curriculum and would not normally have been allowed to progress to training with the cadaver. It was not possible to estimate or accurately extrapolate the # errors, sentinel errors, or time for the 3 incomplete procedures. Thus, for the comparative analysis of the three groups for steps, errors, sentinel errors, and time, Group A = 12, Group B = 14, and Group C = 15 subjects.

7.34 Steps Completed

Figure 7.5 shows the mean and 95% confidence intervals (CI) of procedure steps completed by Groups A, B, and C. Groups A and B completed a similar number of procedure steps, but Group C on average completed four more steps. The differences between the groups' performances using the regression model with Group C as the reference group is summarized in Table 7.2. The results showed that Group C completed on average 42.2 procedure steps. Those in Group A completed 3.8 fewer steps, while those in Group B completed 4.7 fewer steps. Both of these differences were found to be statistically significant (C vs. A $p < 0.001$ and C vs. B $p < 0.001$).

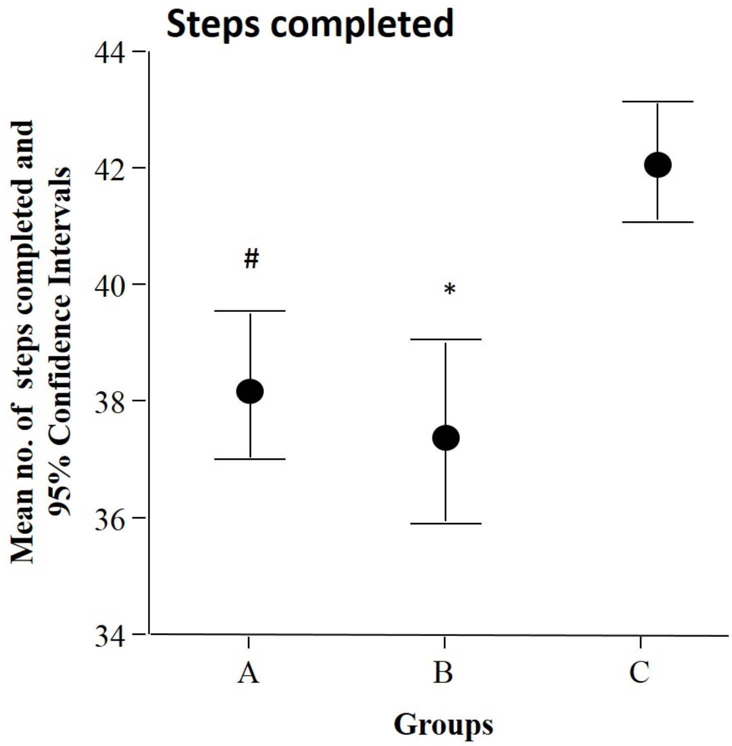


Figure 7.5 Graph depicting the mean number of Bankart procedure steps (and 95% confidence intervals) completed for each study group (A: traditional [control]; B: simulator; C: simulator plus PBP curriculum). Also indicated are the probability value observed for the

differences between Group C performance and Group A (# p < 0.001); and Group B (* p < 0.001).

7.35 Procedure errors

The average number of errors and 95% CI for each group are shown in Figure 7.6. The average number of errors amongst those in Group C was 2.6 (Table 7.3). Those in Group A made on average 3.3 more errors, while those in Group B, 2.4 more errors than Group C. Both of these differences were found to be statistically significant (C vs. A, p = 0.011 and C vs. B, p = 0.049). Overall, those in Group C demonstrated a 56% reduction in the mean number of errors over Group A, and a 41% reduction over the number of errors made by those in Group B. The participants in Group C were also more consistent, with the range in number of errors much smaller when compared to Groups A and B.

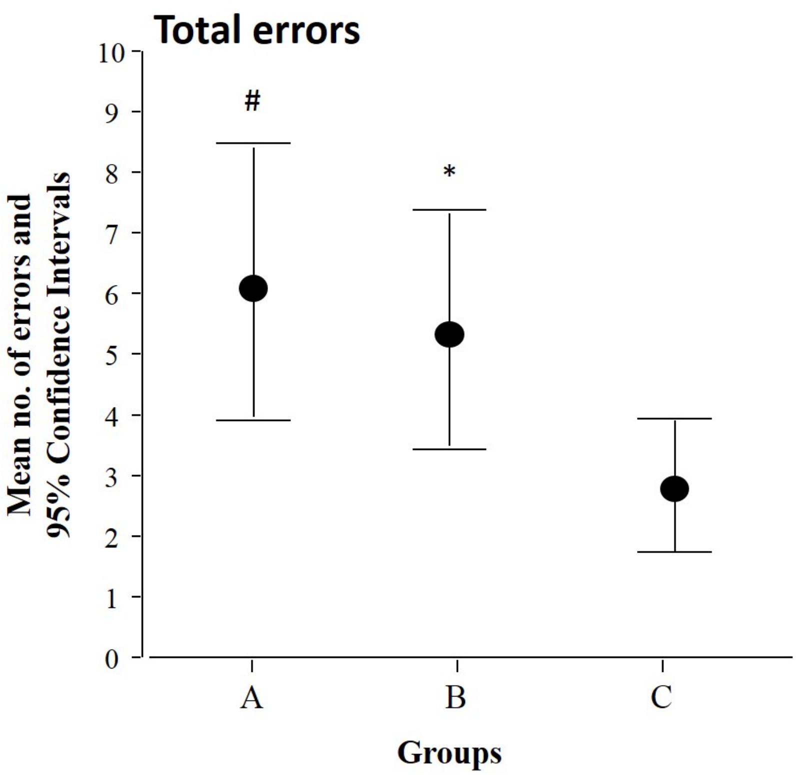


Figure 7.6 Graph depicting the mean number of procedure errors enacted for each study group and the probability value observed for the

differences between Group C performance and Group A (# p = 0.01); and Group B (* p = 0.049).

7.36 Sentinel errors

The mean and CI for the sentinel errors are shown in Figure 7.7. On average those in Group C made 0.53 sentinel errors, while those in Group A made 1.175 more and those in Group B, 0.43 more sentinel errors (Table 7.4). The difference between Group A and Group C for sentinel errors was found to be statistically significant (p = 0.017) but not the difference between Group C and Group B. Overall, those in Group C made 69% fewer sentinel errors than those in Group A and 44% fewer than Group B.

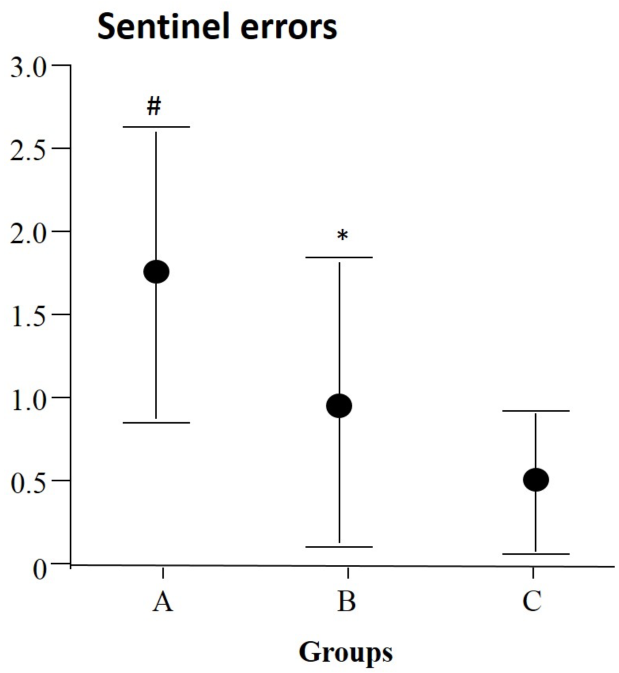


Figure 7.7 Graph depicting the mean number of sentinel errors enacted for each study group and the probability value observed for the differences between 'Group C' performance and Group A (# p = 0.023); and Group B (* p = 0.351).

7.37 Bankart performance time

The mean and 95% CI for time taken by the groups to perform the index procedure is shown in Figure 7.8. Groups A, B, and C took a similar amount of time to complete the procedure with no significant differences observed (Table 7.5).

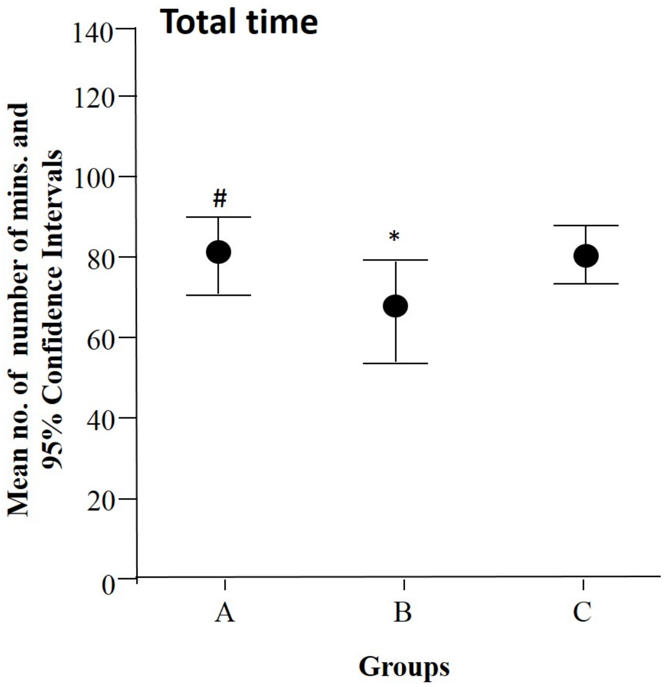


Figure 7.8 Graph depicting the mean procedure duration in minutes for each study group.

7.38 Analysis of Group C^{PBP}

A secondary analysis was conducted to determine the performance of the 12 Group C^{PBP} subjects. The differences between the mean scores of Group C^{PBP} vs. Group C were calculated. The difference in steps completed was marginal (C^{PBP} = 42.46 vs. C = 42.2). The error analysis showed that 12% fewer errors were made (C^{PBP} = 2.29 vs. C = 2.6), 6% fewer sentinel errors were made (C^{PBP} = 0.5 vs. C = 0.53), and 4% less time was required to complete the procedure (C^{PBP} = 77.17 minutes vs. C = 80.44 mins.).

7.39 Final Bankart Proficiency Benchmark

The benchmark (set as the mean performance of an experienced group of surgeons) had previously been established for a cadaver shoulder[66] as 'no more than 3 total errors (one less error than the simulator model benchmark) and no more than 1 sentinel error. In addition, a 3 anchor Bankart repair must have been completed. Overall, 28.6% (4/14) of Group A, 36.7% (5/14) of Group B, 68.7% (11/16) of Group C, and 75% (9/12) of C^{PBP} were able to achieve the final proficiency benchmark.

Logistic regression analysis for the relative differences between the control conditions, Group A (traditionally trained group), Group B (simulator), Group C (simulator + BPB *curriculum*), and Group C^{PBP} (simulator + PBP *protocol*) was performed and used to determine the odds ratios for the comparisons. Relative to Group A, Group B subjects were 1.4 times as likely ($p = 0.121$), Group C were 5.5 times ($p = 0.033$), and Group C^{PBP} were 7.5 times ($p = 0.024$) as likely to achieve the final quantitatively defined proficiency benchmark. Only the comparisons of proficiency between Group A and C, as well as between Group A and C^{PBP} were statistically significant (Figure 7.9). Trainees from Group C^{PBP} had a 36.4% greater probability of achieving the final benchmark than those in the entire C Group.

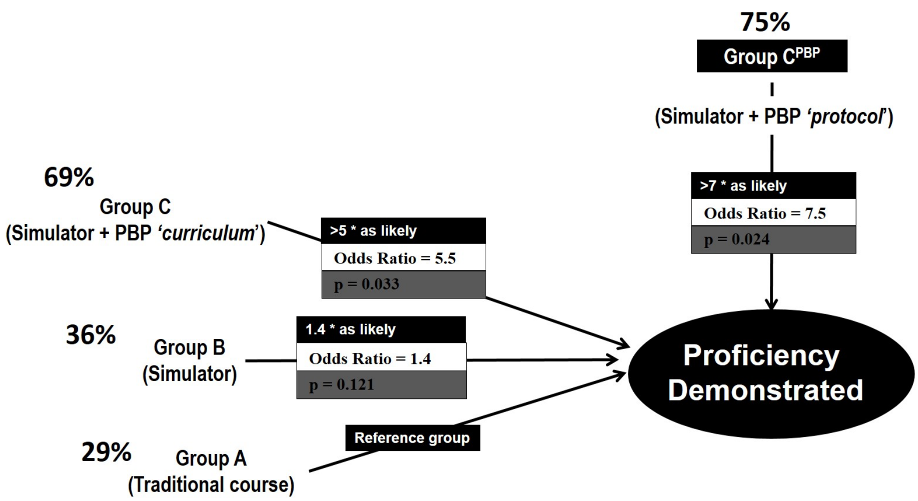


Figure 7.9 Diagram outlining the odds ratios and statistical significance of the differences of Group A from Groups B, C, and C^{PBP} for the final Bankart proficiency demonstration.

7.4 Discussion

7.41 Proficiency Based Progression Paradigm

Two primary conclusions can be drawn from the data in this study. First, the performance of the entire C Group on the final Bankart evaluation shows that the proficiency based progression <u>curriculum</u> employing simulation is superior to both the traditional curriculum (Group A) and to one identical to Group C (including the use of the simulator), but without the requirement to demonstrate proficiency (Group B). Secondly, the performance of the C^{PBP} Group reveals the superiority of the PBP <u>protocol</u> itself, in which only those trainees who meet each sequential proficiency benchmark during training are permitted to progress in the curriculum. The most important and revealing comparison of the three training protocols, therefore, compares Groups A, B, and C^{PBP}. Those in Group C^{PBP} performed more of the operative steps but did not take significantly longer to do so. They also made significantly fewer objectively assessed intra-operative errors, and were over 7 times more likely to achieve the final benchmark than those in Group A who followed a traditional training pathway.

The performance of Group B was only marginally better than Group A, and suggests that it is not simply access to working with the simulator that is important. Rather, it is the metric dependent, proficiency based progression curriculum coupled with the simulator that optimizes the effectiveness of the training. The findings of this investigation strongly support the 'outcomes' (objective assessments) rather than 'process' (time spent / exposure gained) based approach to graduate medical education recommended by the Institute of Medicine.[96]

Due to the artificial constraint of a limited time period for training (a week-end course), all members of Group C were allowed to progress and practice with the cadaver and participate in the final Bankart assessment whether they met all of the intermediate benchmarks or not. During the training process, 25% of the subjects in Group C failed to demonstrate proficiency during the knot-tying phase or on the shoulder simulation model (1 subject failed both). These individual performances are, thus, not representative of the PBP protocol and diminish the performance of Group C as a whole. Had these four subjects had additional time and further opportunities to demonstrate proficiency, it is likely that they may have been able to do so.

Overall, confidence in the observed effects of the training methods was high. There were no statistically significant differences between the groups on pre-course, baseline visuo-spatial, perceptual, or psychomotor assessments. Furthermore, the blinded and objectively assessed videotaped performance of the subjects' final Bankart repairs was scored with a consistently high inter-rater reliability (greater than 90% agreement between the raters for all assessments, with none falling below 80%).

7.42 Performance Outliers

For the 2 Group A subjects who effectively completed only a 1 anchor Bankart repair, it is probable that they would have enacted substantially more errors during a complete 3 anchor procedure. It was not possible to accurately estimate the total number of errors that would have been enacted for a full repair. Consequently, all of their data (including steps completed and errors enacted) had to be excluded from the statistical analysis. The necessary exclusion of these 2 subjects' data results in an over-estimation of the performance of Group A as a whole. The one Group C subject, whose repair was considered incomplete, abandoned the effort on the first anchor repair, as they were unable to complete the knot tying steps. They did, however, complete the second and third anchor components of the repair, and along with performing 37 steps, made only 4 errors and no sentinel errors. Thus, their overall performance was not substantially different from Group C as a whole.

7.43 PBP Superiority

An important finding from this study is that the training process must be more than an educational experience. Simple knowledge of the metrics (steps and errors) and the opportunity to practice with expert feedback (Group A – traditional AANA resident course) resulted in an inferior demonstration of arthroscopic Bankart skills. Further, the addition of the opportunity to work with the simulator (Group B) resulted in a modest improvement in performance over the control, Group A. The employment of the metric dependent, proficiency based progression curriculum coupled with the simulator (Group C) resulted in the acquisition of a statistically superior ABR skill set. Multiple potential reasons exist for the superiority of this protocol:

1) the requirement to obtain a passing score on the cognitive exam at the outset ensures that the trainee is very familiar with the steps to be completed and errors to be avoided for the reference repair;

2) proximate feedback linked to established and validated metrics facilitates the prompt, specific, and effective correction of errors;
3) deliberate practice[96] in attempting to mimic the specific skills demonstrated in the orientation video assures uniformity in acquiring the essential skills needed to perform the reference procedure;
4) the medium fidelity simulator provides the opportunity for practice and repetition of the important skills necessary for effective performance of the Bankart repair
5) the validated performance benchmark for the simulator serves as an intermediate assessment tool and helps identify individual trainee deficiencies requiring correction
6) the trainee's knowledge of the requirement to demonstrate proficiency by meeting each of the intermediate benchmarks to be able to progress in training, help assist the trainee in acquiring the necessary skills
7) trainee performance must be demonstrated to a quality-assured performance level based on validated metrics.[79, 201, 212, 266]

7.44 Metric-Based Curriculum

There are a number of unique aspects to this study and they primarily relate to the development and use of the procedural 'metrics'. This is the first simulation-based study we know of in which the metric characterization and validation of a complete procedure has been carried out. This effort sought to investigate the merits of the emerging 'paradigm shift' in surgical skills training from the apprenticeship model to a proficiency-based progression format and consequently became known as the AANA Copernicus Initiative (Nicolaus Copernicus is credited with the 'paradigm shift' from the earth to the sun being considered the center of the universe). In the approach reported on here, the simulation was simply one of the vehicles for the delivery of a metric-based training curriculum. Much of the effort focused on the development and validation of metric-based performance characteristics that appropriately captured a reference approach to the performance of an arthroscopic Bankart repair.[65] Face and content validity for the Bankart metrics were verified using a modified Delphi Panel meeting with Master and Associate Master AANA shoulder faculty.[65] The construct validity of the metrics coupled with the shoulder simulation model[93] and separately with a cadaver shoulder[66] was confirmed. On the basis of these results, specific performance benchmarks were established separately for the shoulder model, and the cadaver specimens.

The approach to the assessment of performance used in this study employs precise metric definitions of performance and simply requires the scorer to determine whether a specific event did or did not occur. This binary approach to the measurement of performance has been shown to facilitate the reliable scoring of metric-based performance units across a variety of functions during skills training[196-198, 261] with different experience levels.[199, 200] In contrast, "Likert scale assessments" (Table 3.1) result in a less focused approach to minimizing errors due to the fact that the deviations from optimal performance are less clearly defined.[64] A 'Likert-type scale' is a method of ascribing a quantitative value to qualitative data to make it amenable to statistical analysis and was originally designed to assess a range of respondent attitudes.[62] Due to the inherent subjectivity in this method of attempting to rate objective performance, it can be difficult to obtain acceptable levels of inter-rater reliability [>80%] in the scoring of events.[210] It has been shown that Likert-scale scoring may be less reliable than are metric based assessments[63] and simply gives the trainee feedback information on the global aspects of their performance.

7.45 Simulation Platforms

The detailed metrics enabled a simulation platform that already existed (an anatomically accurate shoulder model) to be utilized for training and assessment. This medium fidelity platform is relatively inexpensive, readily available, and serves as an accurate representation of the human shoulder joint. It is not, however, able to capture any performance data or provide feedback to the trainee. A significant investment in time and effort for instructional faculty and video reviewers was required to obtain detailed data and formulate accurate performance assessments. The approach used for simulation-based training in this study holds considerable promise in the short term, as the vast majority of surgical procedures (particularly for traditional open surgery) have no virtual reality platform. Relative to higher fidelity computer based simulators, physical simulation models are much easier to develop in an expedient manner. This capability affords the surgical community the opportunity to develop PBP simulation-based training programs in a reasonable amount of time for traditional and new surgical procedures. The crucial element in the effectiveness will be the creation of appropriate and accurate metric based characterizations and the "operational definitions" (Table 3.1) for those metrics.

Studies have begun to emerge assessing the value of simulation on surgical skills education and training. We were unable to find any studies in the literature with which to compare our investigation. Frank[267] performed a systematic review of the published literature (19 studies) on modern arthroscopic simulator training models. The analysis suggested that practice on arthroscopic simulators improves performance on the simulator, but evidence that skills obtained during simulator training are transferred to the operating room is lacking. Cannon[268] studied the impact on transfer of training using the ArthroSim virtual reality arthroscopic knee simulator, which has previously been shown to have construct validity.[43] Those PGY 3 orthopedic residents who trained on the simulator (for an average of 11 hours) demonstrated greater proficiency on a live diagnostic knee arthroscopy than the control group trained in a traditional fashion.

They performed significantly better on the procedural checklist and probing skills, but not on visualization skills.

7.46 Implications

For well over a century, the apprenticeship model has been the predominant method used to assist surgical trainees in skill acquisition and preparation for the practice of surgery. A paradigm of repeated observation in addition to graded, enhanced responsibility and independence during operations of increasing technical complexity has been employed. Although reasonably effective, this approach is inefficient and produces considerable variability in the skill sets obtained by trainees with equivalent time exposure and experience.[254] Alternatively, proficiency-based progression training using simulation enables the trainee to focus on the acquisition of specific procedural skills, measure their progress, and correct deficiencies. It is through the process of deliberate practice,[22] that they learn not only what to do, but perhaps more importantly, what not to do. The trainee is thus able to enact errors and learn to correct them in an inconsequential manner, and without risk to patients. In addition, this structured approach promotes the acquisition of a more homogeneous skill-set at the completion of training.[266]

One of the concerns of the medical community prior to this study was the generalizability of simulation-based training. Even simulation enthusiasts harbor the concern that simulation-based training effectiveness may be, in part, a function of the effort that enthusiasts put into the training initiatives and the reported science. The results of this study demonstrated that simulation-based training is very effective, even when applied across a large number of residents from training programs throughout the United States using faculty equally dispersed. One of the reasons for this success was a deliberate choice by the investigators to choose a reference approach to a particular procedure. This method of standardization means that at the outset of their learning a particular procedure, trainees do not have myriad of approaches and techniques to master. They can develop their own surgical style once they have acquired safe operative skills for the reference approach.

7.47 Limitations:

Other than their year in training, no additional information was available for the participants. The extent of an individual resident's arthroscopic surgery exposure and experience, particularly shoulder arthroscopy, was likely variable. The diverse group of residency programs represented (21) and the fact that residents from an individual program were randomized to different training protocols should have made them homogeneous and minimized selection bias. Although the residents of Group A performed the best overall on the baseline visuo-spatial, perceptual, and psychomotor assessments, the between-group differences were not statistically significant.

The number of participants enrolled was based in part on previous studies,[22] and secondarily on the logistical challenges of having more than 12 simultaneous recording stations at the OLC during the final Bankart repair. The 14 Group A residents were volunteers from a normal AANA resident course and their Bankart recording was performed over 2 flights. Training for the 14 Group B and 16 Group C participants was conducted on a separate study weekend. The final Bankart repair for these two groups required 3 flights of surgical procedures, some of which lasted 2 hours. With turnover, 8 hours were required to record the 3 flights (30 total recordings).

It is acknowledged that conducting an analysis of proficiency based progression training during the finite time constraint of a weekend course imposes an artificial constraint. PBP training dictates that the trainee continues to study and practice as long as it takes to master the requisite skills. It is probable that the majority of those in Group C who did not reach all of the intermediate benchmarks during the weekend course would likely have been able to do so with additional training and practice. Further, it was not the intent of this investigation to study the efficiency of individual residents' skill acquisition, nor their efficiency in performing the index procedure. The efficiency of the PBP training protocol, however, is clearly implied. In essentially the same time frame as the other groups, a substantially greater % of residents trained using the PBP protocol coupled with simulation achieved the final benchmark compared to those participating in the other training methods.

The final proficiency assessment was performed using cadaveric shoulders. The specimen size, tissue compliance and extent of pre-existing glenohumeral pathology likely introduced some inherent variability in the cadaver shoulders. The impact of those differences was minimized by employing the previously described "acceptability" criteria – those specimens deemed unsuitable for an arthroscopic Bankart repair were replaced and a new video recording was initiated.

Finally, all participants and faculty for each of the 3 study groups were provided a link to 2 full-length videos (lateral decubitus and beach chair) demonstrating each of the 45 steps and showing or identifying all of the errors and sentinel errors. While the link was provided to all participants and faculty 1 month prior to the course in which they were involved, we have no record of how often they viewed / studied the videos.

7.5 Conclusions

A PBP training curriculum and protocol, coupled with the use of a shoulder model simulator and previously validated metrics, produces a superior arthroscopic Bankart skill set when compared to traditional and simulator enhanced training methods.

Table 7.1: Demographic and baseline perceptual, visuo-spatial, and psychomotor assessments for the group participants

Demographic Variable	Group A	Group B	Group C	*P* Value
Gender, n				
Male	13	13	14	
Female	1	1	2	
PGY of training, n				
PGY 4	6	8	7	
PGY 5	8	6	9	
Right hand dominant	82%	93%	94%	
Age, yr, mean (SD)	33 (4)	31 (2)	32 (3)	.54
Perceptual assessment (PicSOr), mean (SD)	0.92 (0.07)	0.93 (0.07)	0.93 (0.03)	.89
Visuospatial assessment, mean (SD)	25 (10)	26 (7)	24 (7)	.76
Psychomotor assessments, mean (SD)				
Dominant hand				
Correct incisions	12 (3)	11 (2)	10 (4)	.45
Incorrect incisions	0.3 (0.5)	0.9 (0.1)	0.1 (0.3)	.15
Nondominant hand				
Correct incisions	10 (4)	9 (5)	9 (5)	.66
Incorrect incisions	0.1 (0.5)	0.1 (0.3)	0.2 (0.6)	.6

PGY, postgraduate year; PicSOr, pictorial surface orientation.

Table 7.2: Steps completed regression analysis.

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	SE	β	*t*	*P* Value
Reference group: group C (constant)	42.167	0.618		68.225	< .001
Group A	−3.833	0.927	−0.559	−4.135	< .001
Group B	−4.702	0.890	−0.715	−5.286	< .001

B, beta; β, standardized beta; SE, standard error; t, test statistic.

Table 7.3: Errors made regression analysis.

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	SE	β	*t*	*P* Value
Reference group: group C (constant)	2.600	0.819		3.173	.003
Group A	3.275	1.229	0.444	2.665	.011
Group B	2.400	1.179	0.339	2.035	.049

B, beta; β, standardized beta; SE, standard error; t, test statistic.

Table 7.4: sentinel errors made regression analysis.

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	SE	β	*t*	*P* Value
Reference group: group C (constant)	0.533	0.313		1.706	.096
Group A	1.175	0.469	0.425	2.506	.017
Group B	0.431	0.450	0.162	0.958	.344

B, beta; β, standardized beta; SE, standard error; t, test statistic.

Table 7.5: Time taken regression analysis.

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	SE	β	*t*	*P* Value
Reference group: group C (constant)	80.438	4.455		18.055	< .001
Group A	0.634	6.522	0.016	0.097	.923
Group B	−13.009	6.522	−0.331	−1.995	.053

B, beta; β, standardized beta; SE, standard error; t, test statistic.

8. Thesis Discussion

Before a physician is able to practice the art of surgery, he or she must master the related science – relevant anatomy, pathology, and requisite technical skills needed to perform an accurate and effective operative procedure. A surgery performed technically well is more likely to lead to a better patient outcome than one done poorly. In order to be reliably implemented, skills must be identified, learned, and polished. Historically, the apprenticeship model served to prepare surgeons to practice their craft. Repeated observation in addition to a gradual increase in responsibility and independence during operations of greater technical complexity has been employed as one of the primary methods of preparation. While reasonably effective regarding exposure to various skills, little assurance could be offered that the necessary techniques had been mastered prior to operating on patients. Greater exposure to some principles and techniques and substantially less to others has led to inconsistent preparation and quality of surgical skills using the apprenticeship model for surgical training. Inefficient use of training resources has resulted as well. A relatively recent paradigm shift in surgical skills education, proficiency based progression (PBP) training, is evolving as both a more effective and efficient strategy than the apprenticeship model for preparing surgeons to offer optimal care to their patients.[208, 212]

The leadership of the Arthroscopy Association of North America (AANA), who's primary mission is arthroscopic education, reviewed their instructional programs, particularly those that related to a focus on technical skills, and posed the question, "Is there a better way to train surgical skills than our current methods?" The attempt to answer that query led to the creation of the AANA Copernicus Initiative, an investigation conducted to determine the relative effectiveness of proficiency-based progression training as an alternative to the apprenticeship model.[269] The change in focus from the apprenticeship model to one of PBP training is considered to be a paradigm shift with respect the approach to surgical training and was thus, labeled, the Copernicus Initiative (Nicolai Copernicus is the individual primarily credited with the paradigm shift from the earth to the sun being the center of the universe).

8.1 AANA Copernicus Investigation

For the patient with unidirectional anterior instability due primarily to a Bankart lesion (capsulolabral detachment from the anteroinferior glenoid) without significant bone loss, a suture anchor repair employing 3 implants is a commonly accepted method utilized to obtain a successful patient outcome.[87, 89, 143, 209, 258] An arthroscopic Bankart repair (ABR) for recurrent shoulder instability was chosen as the index procedure for the related studies. The investigation into PBP training coupled with simulation required the development and validation of 3 separate, specific tools to conduct the analysis. The first component to be created was a metric tool, which could objectively and accurately characterize the performance of an ABR by clearly defining the essential steps, errors, and sentinel errors (more serious) for a standard reference repair.[65] It was critical to subject the drafted metrics to rigorous validation rather than have them be predominantly based on the opinions of a small group of metric authors. The metric tool created was demonstrated to have face and content validity[200] using a modified Delphi panel methodology. Once validated, the metrics were copyright protected. Secondly, a training tool (a shoulder model simulator coupled with the ABR metrics) was shown to have construct validity, demonstrating the ability to distinguish between novice and experienced surgeon performance.[93] A proficiency benchmark for the use of the metrics with the simulator was also established. Lastly, an "assessment tool" (a cadaver shoulder coupled with the ABR metrics) was evaluated and also shown to have "construct validity".[66] A separate and specific benchmark was also established for the use of the metrics with a cadaver shoulder, which is the closest surrogate for live surgery. A final preliminary investigation was needed to verify construct validity for the arthroscopic knot tying and testing protocol and to establish a related proficiency benchmark.[35]

In the initial study format, an intra-corporal knot tying simulator and a dry bench top shoulder model simulator served as vehicles to assist in comparing two groups of residents, one with the requirement to demonstrate skill proficiency along the training pathway, the other without the need to show the ability to perform intermediate steps. As the study design process matured, we elected to add a 3rd group of residents who underwent the current AANA method of resident training that included lectures, knot tying practice, and training using a cadaver without the use of the shoulder model simulator. This last group served as a baseline reference representing the apprenticeship training methodology with which to compare both the impact of adding simulation as well as simulation coupled with proficiency-based progression training. The reference group assisted us more directly in answering our original research question as to whether there was a better way to train surgical skills than our existing methods.

In a prospective, randomized, blinded study, 44 PGY 4 or 5 residents from 21 approved US orthopedic residency programs were randomly assigned to one of three skills training protocols: Group A: Traditional (control, N = 14)(routine AANA Resident Course representing the apprenticeship model); Group B: Simulator Enhanced (N = 14)(modified curriculum adding a medium fidelity shoulder model simulator 'experience', but without the need to demonstrate skill or proficiency); or Group C: Proficiency Based Progression (N=16)(PBP curriculum in which a passing score on an exam demonstrating knowledge of the Bankart metrics was required as well as a demonstration of the ability to meet the proficiency benchmarks for arthroscopic knot tying and the performance of an arthroscopic Bankart repair on the model simulator)**.**[94] Baseline visuo-spatial, perceptual, and psychomotor skill tests confirmed that the study group subjects were homogeneous. At the completion of training, all subjects independently performed a 3 suture anchor ABR on a cadaver shoulder, which was videotaped in its entirety. The training and assessment of all groups was performed at the Orthopedic Learning Center in Rosemont, Illinois.

8.11 Copernicus Study Results

While Group C completed a significantly greater number of steps than both Groups A and B, it was not the most profound difference. The enactment of errors was by far the greatest discriminator. The PBP trained Group C made 56% fewer objectively assessed errors than the traditionally trained Group A ($p = 0.011$) and 41% fewer than Group B ($p = 0.049$)(both comparisons significant) [94].

For the purposes of determining the impact of the PBP 'curriculum' in comparison to the 'protocol', Group C was further subdivided. The performance of the entire C Group (whether or not they achieved all of the intermediate benchmarks) represented the PBP 'curriculum' as they were all subjected to the same training methodology with proximate feedback, deliberate practice, and focused correction of deficiencies. The Group C residents who successfully met each of the intermediate benchmarks (cognitive knowledge of the metrics, knot tying and an ABR on the model simulator) were designated Group C^{PBP} and represented the PBP 'protocol' in which only those trainees who are able to demonstrate each of the intermediate proficiency benchmarks are able to progress in training. In a two-day course, 75% of Group C^{PBP} and 68.7% of the entire C Group achieved the proficiency benchmark on their final Bankart repair compared to 36.7% of Group B and 28.6% of Group A. A logistic regression analysis was performed to estimate the probability of those trainees from the different groups being able to attain the proficiency benchmark for the final ABR. Using Group A as a reference, Group B participants were 1.4 times as likely, and Group C residents (entire cohort) were 5.5 times as likely to achieve the final proficiency benchmark. The results of the C Group in general indicate that the PBP curriculum is superior to both the traditional apprenticeship model as well as one that added a simulator experience. The PBP curriculum was not simply a “weeding out” of less innately skilled residents as the entire C Group, some of whom did not meet all of the proficiency benchmarks, dramatically outperformed those in reference Group A. Subgroup Group C^{PBP} members who met all of the intermediate proficiencies were 7.5 times more likely to meet the benchmark than Group A. In a true PBP model, those in this last group would have been the only ones allowed to progress in training and their performance is representative of the superiority of the PBP protocol. The results of this investigation were unequivocal in demonstrating that a PBP training curriculum using previously validated metrics coupled with the use of a medium fidelity shoulder model simulator facilitated the acquisition of a far superior arthroscopic Bankart skill set when compared to traditional and simulator enhanced training curriculums.

8.2 Lessons Learned

8.21 Communication

The small core group of individuals (3 – 5) who compose the procedural metrics must be intimately familiar with the index procedure to be characterized, not only the appropriate steps to execute, but also the most important errors to avoid. In addition, experience with teaching the techniques involved offers insight into aspects of the procedure that frequently prove most challenging to trainees. Members should have good linguistic skills to facilitate the crafting of metrics with clear, unambiguous definitions**.** The goal of drafting accurate and meaningful metrics must be kept at the forefront and the value of each member’s input respected. There is a risk that if an individual member becomes entrenched in advocating for their perspective, the drafting process can be halted or an important metric deleted altogether for lack of consensus. Each developer must be able to discern when to advocate strongly for their perspective and when to acquiesce and accept that a metric definition might not reflect how they would prefer to perform the task, but that the specific metric is not wrong, inappropriate, or inadvisable as drafted. A strong group leader is essential in facilitating the deliberations. A broad and balanced perspective should be sought, but the leader should recommend and advocate for compromise when necessary.

It is critical that the core group of experienced surgeons responsible for drafting the step and error metrics for a specific procedure be able to communicate effectively and efficiently. Considerations, deliberations, and dialogue are most valuable if conducted in real time, particularly in relation to assessing surgical performance on video recordings. A potential challenge exists in that having the core group of metric developers physically present for each of the necessary meetings can be associated with substantial cost, impart significant scheduling challenges, and require onerous absences from practice responsibilities. After multiple unsuccessful attempts were made to identify a method of media transmission with adequate bandwidth to facilitate concurrent video review by metric developers in different geographic locations (various sites across the U.S. and Ireland), an effective and inexpensive method was identified. For the online sessions, the use of Skype (www.skype.com) group videoconferencing enabled the investigators to simultaneously review arthroscopic videos in real-time with acceptable resolution. An inexpensive USB camera (I-PEVO $69) was connected to the call initiator's laptop and designated as the 'source input' in the Skype settings. A separate desktop computer with a high-resolution monitor was used the play the subject videos. When the USB camera was directed toward the image playing on the desktop computer, all members of the group call were able to simultaneously view the arthroscopic video images.

COMMUNICATION LESSONS:

1) It is essential that the metric developers be intimately familiar with the procedure and possess a collegial and respectful working dynamic to avoid undermining the process.
2) While initially in-person meetings are critical, a practical and cost-effective means of communication during performance review meetings is essential. Skype teleconferencing proved effective, efficient, and inexpensive.

8.22 Index Procedure Selection

In developing a PBP training curriculum for a particular surgical specialty, it is impractical to create a set of validated metrics for all commonly performed procedures. A number of factors should be considered in selecting which potential procedures should be chosen for metric characterization. In the progressive acquisition of skills, trainees should master the essential components of a standard reference procedure before taking on more complex pathology. With time and experience, they will develop their own preferences even for standard index operations. The procedure selected should be one that is relatively frequently performed and also be comprised of various skills that are commonly employed across the breadth of a particular surgical discipline (and will thus represent a broader scope of proficiency once acquired). Finally, the procedure should have a set of recognized and preferred techniques rather than one with numerous different and accepted ways of performing the procedure. The metric development exercise becomes substantially more complex and difficult when the metrics must be written to recognize multiple different methods of accomplishing the same task. In that scenario, bias is often introduced in the attempt to uniformly score the metrics for different methods. However, in the instance where there are two commonly accepted surgical approaches for the same technique (as is the case with an arthroscopic Bankart repair), the step and error metrics must be drafted to avoid scoring bias or penalizing surgeons that use one approach rather than another (i.e. – both the lateral decubitus and beach chair orientations have wide acceptance among those performing arthroscopic shoulder procedures).

INDEX PROCEDURE SELECTION LESSONS:

1) The procedure selected for metric characterization should be straightforward and employ commonly used techniques.
2) Those surgeries with multiple different and recognized methods of accomplishing the same objective should be avoided. In that scenario, drafting metrics to avoid scoring bias is substantially more difficult.

8.23 Task Deconstruction / Metric Development

The intent of procedure characterization is to capture all of the essential elements of a specific surgery and be able to employ them to accurately discriminate between the performances of different operators.[65] The step and error metrics must be tightly defined (rather than simply described) and should be unambiguous to the extent that a video rater is able to simply verify that a particular event "was", or "was not" observed to occur. We found that some drafted metrics were ambiguous or could not be reliably and unambiguously defined even with editing, and therefore needed to be dropped. Quantitative dimensions proved impossible to use in defining certain metrics as no means of measurement or calibrated tool was available on the videos (i.e. the specific location in mm from the glenoid rim for anchor holes). Instruments of known size served as the best quantitative reference when they were present in the field-of-view, i.e. a 3.5 mm drill guide. We found that specifying anatomic regions rather than distances proved to be more practical (i.e. – 'guide is located in the inferior region of the anteroinferior quadrant of the glenoid'). Several metrics that were well defined could not be reliably and consistently identified on the practice videos and needed to be deleted. In contrast to steps, error metrics proved to be a much more powerful discriminator of performance and should be particularly robust in their definition. Every effort should be made to draft and revise potential error metrics until they are acceptable rather than delete them. Not all potential intra-operative errors, however, are represented in the metrics because they; a) occurred too infrequently, b) could not be unambiguously defined or c) could not reliably be identified from operative performance on videos. Compared to videos of experienced surgeons, novice videos contained many more mistakes and were of much greater value in helping to identify and draft error metrics.

In general, time as a performance metric by itself is a relatively poor surrogate for performance quality. A procedure could be performed relatively quickly either due to the knowledge and skill of the operator, or because steps and sequences were skipped. In that scenario, however, if total time for the procedure was coupled with precise operational metrics, those two performances could be easily distinguished. Total procedure time was therefore retained as a metric. 'Lack of progress' (LOP) was considered as an error metric, but identifying the specific beginning and ending points on a video as to when meaningful progress was not being made or resumed proved unreliable. LOP was discarded as a potential metric.

Once the metrics were refined and stress tested by the core group using multiple full-length videos, a modified Delphi panel of experienced surgeons was convened. Their task was to further evaluate how representative and valid the units of behavior were in characterizing the arthroscopic Bankart procedure (i.e. to obtain face and content validity). We found that it was imperative that a thorough orientation for the Delphi panel members to clarify the purpose for convening, and the rules for deliberation take place. It proved critical that during the cyclic process of Delphi Panel voting, the participants understood that a ‘yes’ indicated that the metric definition presented was accurate and acceptable as written, but not necessarily that it was in the manner in which the particular panelist might have chosen to complete the step (i.e., the step was not incorrect, inadvisable, or unsafe). Frequently, the panel provided useful discussion, illumination and clarification for the intra-operative steps and errors. Metrics were either accepted as written, revised until satisfactory, or deleted. Consensus should be reached for each of the metrics characterizing the procedure (we elected to use unanimity for consensus).

TASK DECONSTRUCTION / METRIC DEVELOPMENT LESSONS

1) Performance metric definitions must be unambiguous and reliably identifiable on surgical videos.

2) Numeric dimensions should be avoided and instead, reference should be made to tools of known dimension or to identifiable anatomic regions.

3) Error metrics are the most valuable in discriminating between levels of performance - videos of novice operators are the most helpful for identifying specific errors.

4) Operative time is most useful when coupled with precise operational metrics.

5) An orientation session for the modified Delphi panel should take place in person, and is essential to facilitate orderly, efficient, and meaningful deliberations.

8.24 Video Reviewer Training

It proved to be impractical for the 3 arthroscopic shoulder surgeons (who, along with Dr. Gallagher drafted the metrics) to score all of the videotapes from the 2 construct validity studies and the randomized trial (>80 videos). Extensive training of 7 additional reviewers was conducted. Training was initiated with an individual review and study of the metrics. An 8-hour in-person group review of each of the step and error metrics was conducted, which included video examples of those most difficult to score. Discussion helped immensely to clarify how each step and error was to be scored, including the nuances and conventions to be used. In two subsequent 2-hour group phone conferences, the differences and discrepancies amongst all reviewers for the practice videos were compared and discussed, seeking conformity in scoring. In addition, each designated pair of reviewers conducted one to three additional phone conferences to analyze the specific instances in which the two of them scored particular events differently. This training methodology proved invaluable in clarifying and ensuring consistency in the scoring of videos. Following training, an acceptable inter-rater reliability (>0.08) was achieved for each designated rater pair on practice videos.

An attempt was made to conduct an abbreviated training of 5 additional pairs of raters several months before the scheduled randomized trial but proved unsuccessful. Although the principles of PBP training and the metrics were all reviewed by conference call and multiple practice videos were scored, acceptable accuracy and reliability could not be consistently attained. Without a face-to-face meeting of the new raters at the start of training, a more rudimentary understanding of the proficiency-based progression training principles may have resulted. Further, the lack of a thorough detailed in-person group review of the individual metrics may have sacrificed valuable educational dialogue amongst members of the group. Helpful insights into scoring some of the more difficult metrics may have been lost. Finally, the failure of each of the members of a rating pair to compare their scoring differences after independently reviewing practice videos likely also contributed to the inability to successfully demonstrate uniformity in scoring. Had this second group of raters had sufficient time and the same thorough training program that the first group was afforded, it is likely that they would also have been able to achieve acceptable IRR.

VIDEO REVIEWER TRAINING LESSONS

1) The orientation sessions involving a review of the principles of PBP training as well as the initial review of the specific individual metrics should occur in an in-person setting.

2) Communicating that information by conference call risks a lack of comprehension, committed engagement, and focus on the objectives.

3) The methodical and comprehensive preparation for those scoring the arthroscopic videos cannot be compromised or abbreviated.

4) Acceptable IRR (> 80%) amongst rater pairs on practice videos must be verified or further training conducted.

8.25 'Experienced' Group / Proficiency Benchmark

The level of skill exhibited by the experienced group members must be consistent with the designation as a 'reference' group, whose performance can be used to assist in the establishment of an accurate and acceptable proficiency benchmark. Frequent subject authorship, podium presentations, and academic position do not serve as reliable gauges of surgical skill. Members of the experienced group who perform substantially below their peers could potentially serve to lower the reference standard. Conceivably, those trainees who met a diminished standard might be inadequately prepared to perform the specific procedure on patients. Indeed, for the experienced groups in our investigation, one of 12 faculty member's scores were removed from the calculation of the benchmark for performance on the simulator model, and one of 10 faculty subject's scores were deleted in establishing the shoulder cadaver benchmark. Both outliers had performed greater that 2 SD below the mean of their peers.

Historically, a "proficiency benchmark" has been defined as the mean performance of a group of 'experienced' operators.[22, 198, 208] Deliberation was undertaken to consider whether this standard was too high for a group of PGY 4 and 5 orthopedic residents. To meet the benchmark, they would have to perform better than 50% of those experienced surgeons currently out in practice, operating on patients. It was acknowledged that if an acceptable number of residents were unable to meet the benchmark and demonstrate proficiency, it would not be possible to make meaningful comparisons between the different training curriculums in the randomized trial. We considered using a less demanding proficiency reference, but could not agree on a different standard that would be recognized, have precedent, and be readily acceptable to the readership. Our concerns proved to be unfounded. For the cadaver Bankart repair, 75% of the PBP trained resident study group was able to achieve the benchmark in a single training weekend (compared to only 50% of the experienced group by the proficiency definition). The construct validity study for arthroscopic knot tying revealed similar results.[35] 24% of faculty knots failed compared to 11% of the PBP trained residents.

REFERENCE EXPERIENCED GROUP / PROFICIENCY BENCHMARK LESSONS

1) It is imperative that a policy be established at the outset regarding performance outliers in the experienced group so as to maintain the appropriate character of the reference standard.

2) The scores of those surgeons in the experienced group performing > 2 SD below the mean should be deleted from the calculation to establish a proficiency benchmark.

3) The mean performance of a group of experienced surgeons constitutes a reasonable, practical, and attainable proficiency benchmark.

8.26 Simulation

Simulators themselves may appear to provide to a shortcut to effective surgical training, and perhaps even by themselves to be the answer to training tomorrow's surgeons in ever more complex surgical techniques – they are not. The number of high fidelity simulators collecting dust, testify to that fact. Too often, a piece of simulation equipment with impressive capabilities is obtained and only later is the question asked, "How will we use it for training?" Undoubtedly, simulators will play an increasing role in the training of surgeons and proceduralists, but they must be appropriately harnessed to be a cost effective contributor to skills education. A PBP curriculum must be developed first. Only then can the identification and application of appropriate simulations to support and deliver specific components of the training curriculum take place. Uncoupled from a carefully designed curriculum, a simulator may only be providing an 'experience' that, even with repetition, may not effectively train the necessary skills. In our randomized trial evaluating different training protocols, the use of the simulator without the proficiency based progression curriculum (Group B) only resulted in a modest improvement over traditional training (1.4 times greater chance of achieving the previously established benchmark than those in the traditional apprenticeship pathway).

The fidelity of the simulator should correspond to the task being trained. For example, in the Copernicus Investigation, we observed that learning to tie arthroscopic knots down a cannula under an opaque dome while viewing the image on a laptop monitor provided all of the realism that would be required in vivo, including tactile feel. It would make little sense to attempt to teach a student to tie arthroscopic knots using a virtual reality 3D simulator, even with excellent haptics, when simpler, lower fidelity, and more cost-effective tools are available. For an arthroscopic Bankart repair, a medium fidelity shoulder model simulator was readily available and required only minor modifications. It provided the opportunity to use all of the same tools and equipment normally employed for the procedure and to perform each of the 45 specified tasks in exactly the same manner as in a live surgical case.

Lower fidelity simulators, however, have drawbacks – they normally only permit a single use, possess the ability to represent a relatively small variety of pathologic conditions, and are limited in their ability to represent fine anatomic detail. By far, however, the greatest challenge to the use of lower fidelity simulations is the requirement for a large investment of individual faculty time to provide instruction and proximate feedback throughout the training process. The requirement for manual scoring of operative performance is also very time consuming.

SIMULATION LESSONS

1) Simulators are most useful when they serve as the vehicle to deliver a strong, metric based curriculum, which must be developed before specific simulators are chosen to train the necessary skills.

2) Low and medium fidelity simulators are readily available and can be very effective for training many of the key steps involved in completing various surgical procedures.

3) A drawback of lower fidelity simulators is the requirement for a significant investment of faculty time for training and performance assessment.

8.27 Investment Related to the Development of a PBP Curriculum

Task deconstruction, metric development, validation, and the formulation of a proficiency based progression curriculum are laborious processes. The entire complex of Copernicus Investigations took place over 2 ½ years.[35, 65, 66, 93, 94] The hard costs for development of the PBP arthroscopic Bankart repair curriculum and associated validation investigations were approximately $475,000 (Table 4). This deceptively small figure is due in part to the substantial volunteer contributions by members of the Arthroscopy Association of North America (AANA) who invested in excess of 2,100 hours in the Copernicus Initiative. Had the volunteers been paid consulting fees for their efforts, the costs would have added approximately $850,000. The ability to effectively communicate by employing Skype teleconferencing dramatically decreased expenses related to investigator and consultant travel, housing, and conference facilities that would have been necessary for in-person meetings. Expenses were further reduced by conducting components of the investigation as an add-on feature to numerous, routinely scheduled AANA resident educational courses.

Table 4

AANA Copernicus Initiative Expenses

Meetings / Travel	$166,600
Orthopedic Learning Center Rental	$23,690
Videos (recording, duplication, distribution)	$25,670
Simulators (shoulder models / cadavers)	$192,450
Administrative (office, lab staffing)	$12,100
Consulting (study design, metric development, implementation, statistical analysis)	$53,100
Total Hard Costs	**$473,610**
Physician Volunteer Contribution (study design, implementation, metric development, rater training, video scoring, data compilation, etc. [2100 hours X $400/hour])	
'Sweat Equity'	**$840,000**
Total	**$1,313,610**

Without question, similar efforts to create objective metrics and performance standards for additional procedures will be streamlined and more efficient. The creation of PBP training modules for fundamental arthroscopic skills have been developed (AANA Fundamentals of Arthroscopic Surgery Training – F.A.S.T.) and will prepare the trainee for more advanced procedure specific skill requirements. Performance metrics are currently being created for three additional arthroscopic procedures: anterior cruciate ligament reconstruction for the knee, rotator cuff repair in the shoulder, and labral repair for the hip.

INVESTMENT LESSONS

1) The hard costs for the development of a PBP training curriculum are manageable and can be reduced by employing low cost, efficient strategies for communication (which will limit the high cost of in-person meetings).

2) When less expensive lower fidelity simulations are utilized, the greatest 'cost' is related to faculty and reviewer consultation time.

3) Curriculum validation studies may be most efficiently conducted during other regularly scheduled instructional courses.

8.3 Considerations and Deliberations

8.31 Arthroscopic Bankart Repair

For arthroscopic shoulder surgery, both the lateral decubitus and the beach chair positions are in common use. The initial set of metrics drafted for the arthroscopic Bankart procedure was created with the perspective of the patient in the lateral decubitus orientation, which was the preference of the members of the core group creating the metrics. In reviewing videos of a broad group of experienced surgeons, it became apparent that the beach chair orientation was often employed. It proved necessary redraft some of the metrics (particularly those working on the anterior glenoid) to permit unbiased scoring, regardless of the position of the patient.[65] Similarly, it is commonplace for some experienced surgeons to view the anterior glenoid with the arthroscope in the posterior portal, and some from the anterosuperior portal. Several metrics that were related to the specific approach of the drill guide to the anteroinferior glenoid needed to be dropped, as they could not be consistently scored with the view afforded from the posterior portal. In summary, it is beneficial to select a surgery with a routine and predominant technique, but the metrics must be drafted to avoid scoring bias if more than one commonly accepted patient position or arthroscopic viewpoint are routinely employed by experienced surgeons.

8.32 Training vs. Discrimination

A dilemma was encountered in drafting the metrics with a two-fold purpose. On the one hand, we intended for the metrics to be thorough and present to the trainee all of the important steps that should be performed to execute an arthroscopic Bankart procedure well. On the other hand, we needed to demonstrate construct validity for the metrics with the ability to validly discriminate between novice and experienced surgeon performance. To meet the first intent, a thorough list of diagnostic steps requiring a "pause to view" or probing needed to be included. To meet the second purpose, it was necessary for the experienced surgeons to complete each of the diagnostic steps in the manner defined by the particular metric or they would fail to gain credit for performing that specific step. For example, the experienced arthroscopist may perform step #8, "view or probe the superior labral attachment onto the glenoid" in a relatively rapid manner (although accurate), without probing and without specifically pausing to clearly indicate that they were evaluating the superior labrum. In that instance, the video rater may not give the experienced surgeon credit for completing that particular diagnostic step. Depending on the number of these instances, the performance of the experienced surgeon could be degraded compared to the novice performer, which would impact the ability to discriminate between novice and experienced surgeons. In addition, a proportionately large number of diagnostic step metrics can result in an excessive and undesirable weighting of the diagnostic compared to the therapeutic procedural steps. For the Bankart investigation, the experienced group tended to miss more diagnostic steps than the novice group. As a result, if we had eliminated the diagnostic metric steps altogether, the discrimination between the novice and experienced operators would have been even greater than we reported.[66, 93] Thus, there must be a balance between the thoroughness of the diagnostic steps for the novice trainee and the avoidance of penalizing the experienced group participants who may be brief but accurate in their diagnostic assessments. In future PBP studies, it may prove useful and valuable to retain the diagnostic metrics for the purposes of training, but include only the treatment steps and errors for the purposes of determining construct validation.

8.33 Error Metrics

Error metrics proved to be the most challenging to craft, but are often the most important for discriminating between different levels of performance.[30, 79] Relative to the steps, much more time was required to write and rewrite the error metrics to ensure that they were robust --- both able to capture the most common mistakes that novices enact, as well as be able to be unambiguously scored. Although many potential errors could be conceived to possibly occur, those creating the metrics should resist the tendency to include errors that were not seen to actually occur in the videos of novice performance used to assist in drafting the metrics. Doing so would spuriously inflate the number of agreements between raters (both scorers would indicate that particular errors did not occur) and render the inter-rater reliability ratio less meaningful. In addition, listing potential metric errors that never occurred would do little to aid the trainee in improving their operative performance. From the large pool of potential Bankart error metrics drafted, we elected to delete those that were never exhibited on any of the fourteen complete videos used to create the Bankart metrics.

Videos showing the poorest operative performance were generally the most valuable as they exhibited the greatest number of mistakes and errors in technique. A review of surgical video recordings grounds the metric developers in actual operative performance and the enactment of specific errors. Full-length surgical videos eliminate the need for them to rely primarily on their recall of errors that might potentially occur based on past experience or prior instruction of those learning the procedure. Finally, observing errors as they are actually enacted facilitates the creation of unambiguous and accurate operational definitions for the related error metrics.

8.34 Sentinel Errors

It was impractical if not impossible to accurately and appropriately weight each of the error metrics on any reliable scale. However, it was acknowledged that some errors are more egregious and potentially accompanied by more severe consequences. In previous research, the term 'critical' has been used to designate those more serious errors.[22, 79, 212] Depending on the severity and consequences of the particular surgery and error (i.e. a vascular laceration during a laparoscopic procedure), the designation as critical may be very appropriate. In those instances, enactment of a critical error may justify automatic failure for the trainee being assessed. It would be extremely rare, however, for a complication related to an arthroscopic surgical procedure to result in a life-threatening complication or one of dire consequences. In our investigation using an arthroscopic Bankart repair, the term critical did not match the typical circumstances. In addition, it was felt that the label critical could also potentially have unwarranted medico-legal implications.

We considered the term 'sentinel' ('to watch out for or be observant of') to be a more appropriate designation for more serious errors during an arthroscopic shoulder procedure. A sentinel error was defined as one that either: a) had the potential to jeopardize the outcome of the entire procedure, or b) led to iatrogenic damage to the involved joint. To the extent that a particular metric based evaluation may be used for high stakes assessment, even the designation of some errors as sentinel may create significant difficulties. While the first definition appears appropriate at face value, it is in fact only the opinion of the Delphi panel that the particular error could jeopardize the outcome – no supportive data exists to objectively confirm that assertion. As the definition of sentinel relates to iatrogenic damage to the joint, only those insults that are directly observed on the video can be scored. Thus, a number of significant tissue insults could occur, not be observed, and therefore not scored. As a result, scoring only those sentinel errors that were observed might not accurately represent the overall technique quality in avoiding damage to the joint tissues. We did, however, find it helpful to designate particular errors as sentinel for instructional purposes.[94] For example, the awareness of potential laceration of the labrum (a sentinel error) led trainees to exercise caution in using the liberator/elevator to mobilize the capsulolabral tissues from the glenoid.

Sentinel errors occur relatively infrequently. If they were used to establish a component of the benchmark, they could have an unwarranted impact on whether or not a surgeon demonstrated proficiency. It is our belief that if the number of sentinel errors created is used as a component of the benchmark for high stakes assessments, rigorous study should be conducted and objective evidence presented to validate the significance of that special designation. Whether sentinel or critical errors are used as a component of the proficiency benchmark or not depends heavily on the nature of the surgical procedure and the probable consequences of the particular error.

8.35 Novice and Experienced Cohorts

In an attempt to verify construct validity of the Bankart metrics, it was necessary to identify representative experienced and inexperienced groups to determine whether the metrics are able to discriminate between different levels of operative performance. We sought to identify a group of experts to serve as the reference cohort. However, no established, objective definition exists as to what constitutes 'expert' performance for an arthroscopic Bankart repair, and thus, the term 'experienced' was chosen instead. It was believed that those routinely serving as faculty for shoulder arthroscopy training courses would accurately represent the skilled group. We were able to obtain an adequate number of full-length videos from the group of experienced performers.

The novice group proved more challenging to identify. Our initial intent was to have those who had been in practice for 1 – 3 years participate as the novices. Despite the incentives of free textbooks and reduced course registration fees, we failed to obtain more than a few videos from those early in practice. Our impression was that less experienced surgeons were reluctant to submit videos that exhibited mistakes and suboptimal performance. We next considered sports medicine and arthroscopy fellows, but could not obtain 'pure' full-length videos from them as there was often 'attending take-over' during a case. When those transitions occurred was impossible to identify. Understandably, the responsible attending surgeon would not permit errors and suboptimal performance to go unchecked during surgical procedures on their patients.

Since videos of live surgery representing the novice group could not be obtained, we ultimately elected to use cadaver shoulders for the arthroscopic Bankart repair investigation. Cadaver shoulders were the most feasible platform and their use constituted the closest approximation to live surgery. In part, for practical reasons as well as efficiency in data collection, we chose to obtain videos during routine AANA resident arthroscopy courses using cadaver specimens.[65] The faculty participated as the experienced group and the 4rd and 5th year resident registrants served as the novice group for the determination of construct validity. As there was no clinical consequence to procedural errors committed on the cadaver shoulders, true, unbiased novice performance was obtained and represented on the full-length videos captured.

8.36 Knot Tying Assessment

An attempt was made to write metrics for the evaluation of knot tying. It proved extremely difficult and unreliable to determine the quality of the primary knot on video assessment and whether or not half hitches were delivered on alternating posts with reversed throws. Even when the components of the knot were observed with sufficient clarity, only a visual appearance of the final knot was possible, which provided little information about the integrity or function of the knot construct. For the assessment of knot integrity, the use of an Instron or similar mechanical testing device was too expensive and impractical. It was determined that an accurate, objective method to evaluate knots was needed. Although it was not possible to evaluate each of the knots thrown during the performance of a Bankart repair, we believed it was nevertheless essential that endoscopic knot tying skill be assessed and verified. We initially struggled to identify an economical and acceptable testing methodology for quantitatively determining how much a knot slipped and loop enlarged. After various trials, a testing method was formatted and validated. A spring-loaded lever attached to a base plate was devised to deliver a static stress of 15#s for 15 seconds. When applied to the suture loop after an endoscopic knot was tied, this simple device was found to be accurate and reliable. The resulting loop elongation could be accurately measured by passing a previously stressed suture loop along a graduated cone (a simple nail punch was as the first prototype). The greater the loop in length (circumference), the further it would pass up the reference cone. The knots tested were all created around a cylinder of uniform diameter. In previous investigations, loop elongation of 3 mm or more was considered to be a failure of the knot construct.[221, 234, 270] An etching on the cone was placed to identify a loop 3 mm larger than the circumference of the reference cylinder. Thus, a previously stressed loop that slid up the cone past the etching constituted a failure of the knot construct. Despite the measuring cone and spring-loaded tensiometer being created from inexpensive materials that were readily available, they proved to be accurate, reliable, and of great utility.

8.37 Proficiency Based Progression Randomized Trial Subjects

The intent of the randomized trial was to determine the relative effect of three different training curriculums on the acquisition of the skills necessary to perform a technically well-done arthroscopic Bankart repair. It was felt to be optimal for the study participants to all be relatively untrained and homogeneous. Postgraduate fellows in sports medicine or arthroscopy were initially considered to serve in the study groups for the randomized trial. This group was rejected as some fellows, toward the end of their fellowship training, have become quite proficient at performing the index Bankart procedure. Other fellows may have had little opportunity to participate in shoulder instability cases. In that scenario, the study population would lack homogeneity. The disparate level of experience could potentially confound the observed effects of the different training regimens and result in the investigation failing to show a difference in the 3 curriculums, when in fact one existed. Ultimately, we chose to enlist 4th and 5th year orthopedic residents to serve as the relatively homogeneous and untrained novice study population.

8.38 Study Duration

A true proficiency based progression curriculum would enable the trainee to practice repeatedly over time until proficiency was demonstrated before progressing to the next skill level. A pure PBP training program was not feasible for our study design as repeated training and testing efforts could not practically be conducted for individual residents from different geographic locations over an extended period of time. Uniformity with respect to faculty, cadaver labs, and testing equipment necessitated that the investigation be conducted over a 2-day course at the Orthopedic Learning Center in Rosemont. Every effort was made to adhere to the principles of PBP training and allow for repeated practice and performance evaluations to occur. For most of the assessments, no time limit was imposed nor was the number of trainee attempts to demonstrate proficiency restricted. However, after careful consideration, we elected to limit the residents in the study population to two attempts to reach the proficiency benchmark for the arthroscopic instability repair using the dry model simulator. For those who failed the initial attempt to demonstrate proficiency on the model, a review session with their faculty instructor and the rater who scored their unsuccessful attempt was conducted. Once the trainee's deficits were thoroughly reviewed and the corrections practiced, a new model simulator was provided and a second attempt afforded for the trainee to demonstrate proficiency. The review, practice, and repeat testing for an individual trainee often required up to 4 or 5 hours. It was not possible to repeat this process more than once (for a total of 2 attempts). We believe that it is probable that given additional time, study, and directed practice, most if not all of the residents across the different curriculums would have been able to achieve the all of the intermediate proficiency benchmarks and the final Bankart evaluation benchmark. Despite the limitations in study duration, an accurate assessment of the impact of a PBP training curriculum was obtained.

9. Conclusions

The AANA Copernicus Initiative was a rigorous investigation into the premise query, "Is there a better way to train surgical skills than our current methods?" The related research enabled an accurate assessment of the impact of proficiency-based progression (PBP) training for the surgical skills required to perform an arthroscopic Bankart shoulder stabilization procedure well. A PBP curriculum coupled with simulation training was dramatically more effective than simulation-based training as well as current AANA training methods. All of the three training courses were staffed with faculty of the same level of experience. Simulation-based training groups had the same resources expended during the actual course. Despite this matching of time and resources for the 3 curriculums, the PBP - simulation training group who met all of the intermediate proficiency benchmarks completed significantly more procedure steps, made significantly fewer errors and sentinel errors, and were more than seven times as likely to be able to demonstrate the final proficiency benchmark as the standard resident training group representing the apprenticeship model. So why was PBP simulation training more effective at achieving these performance levels?

PBP is a method of training that is systematic, quantitative and evidence-based. Precisely what should be trained for a given procedure and how the curriculum should be constructed is derived by practitioners who are very experienced at what they do and are intimately familiar with both the performance of the index procedure as well as instructing trainees in the techniques involved. In deconstructing the arthroscopic Bankart procedure, the metric authors systematically identified and operationally defined characteristics of surgical performance that captured the essence of optimal performance for the selected repair. This comprehensive characterization was then stress tested, and rigorously examined and verified for face, content, and construct validity when coupled with either a medium fidelity shoulder model simulator or a cadaver shoulder. These processes ensure that the metrics derived from the characterization are not only able to discriminate between novice and experienced surgeon performance, but also appropriately depict and capture the essence of optimal performance for a reference approach to the given surgical procedure. The validation process helped to identify the metrics that best distinguished between experienced and novice surgeons, providing clear guidance as to what aspects of the curriculum need particular attention during training. Furthermore, the intermediate assessments provided the individual trainee with the identification of their specific individual deficiencies that required correction.

Knowledge does not equal skill, but provides a foundation for skill development. The cognitive exam ensured that the trainees had a firm grasp on the steps to be completed and the errors to be avoided. The orientation video exhibited all of the steps of the arthroscopic Bankart procedure and either demonstrated or identified all of the errors to be avoided. During the course of instruction, the faculty 'taught to the metrics' providing reinforcement of the metrics for the individual trainee and uniformity in the objectives for the entire group of PBP trained residents. The strategy of proximate feedback linked to deliberate practice facilitated the prompt, specific, and effective rectification of errors and facilitated an efficient and effective approach to skill acquisition. Intermediate performance reviews afforded an accurate assessment of the trainee's progress and helped identify specific skills requiring further refinement. Specific, unambiguous, and clinically relevant performance benchmarks were not based on best guess or aspiration but rather on the actual mean performance of experienced practitioners. Thus, the benchmarks are not an abstract goal but a standard grounded in clinical reality and difficult to dismiss by the trainee as unrealistic targets.

The development of the components of the metric based training curriculum must precede selection of the appropriate simulation to train for specific tasks and skills. Simulators of varying fidelity are often needed from box trainers for knot tying and bench top anatomic models for routine techniques to VR simulators for more complex skills. The medium fidelity simulator used in the Copernicus investigations provided the opportunity for repeated practice of all of the important skills necessary for the effective performance of an arthroscopic shoulder instability repair. Expensive, high fidelity virtual reality simulators are not essential for all aspects of surgical skills training.

All of the training components combined to constitute a curriculum that is substantially more than ‘an interesting educational experience’. The acquisition and demonstration of the important component skills of a procedure to an equitable, transparent, and relevant standard helps to ensure that the surgeon is optimally prepared to provide their patient a successful surgical experience. This training methodology was shown to be very effective even when applied across a large number of residents from training programs throughout the U.S.

The evidence supporting a PBP approach to training surgical skills is compelling. The challenge for the surgical community is not so much to believe that PBP is a superior training paradigm, but rather to embark on a strategy to formulate and implement those concepts into our current and future preparation of surgeons in a timely manner. The costs incurred to implement the principles of PBP training in our educational programing will be substantial, but are likely to pale in comparison to the price tag for managing surgical complications as well as the inestimable cost to patients of suboptimal outcomes. Knowing that substantially better training for surgical skills is possible, it is incumbent upon the surgical community to employ that preparation for the benefit of our patients.

References

1. Ney ZA, Looper CA: **Development and Implementation of an Extravehicular Activity Skills Program for Astronauts**. In.: SAE Technical Paper; 2005.
2. Rodriguez-Paz J, Kennedy M, Salas E, Wu A, Sexton J, Hunt E, Pronovost P: **Beyond "see one, do one, teach one": toward a different training paradigm**. *Quality and Safety in Health Care* 2009, **18**(1):63-68.
3. Shea JA, Healey MJ, Berlin JA, Clarke JR, Malet PF, Staroscik RN, Schwartz JS, Williams SV: **Mortality and complications associated with laparoscopic cholecystectomy. A meta-analysis**. *Annals of surgery* 1996, **224**(5):609.
4. See WA, Cooper CS, Fisher RJ: **Predictors of laparoscopic complications after formal training in laparoscopic surgery**. *Jama* 1993, **270**(22):2689-2692.
5. Beyea SC, Kilbridge P: **Setting a research agenda on patient safety in surgical settings**. In: *Seminars in laparoscopic surgery: 2003*: Sage Publications Sage CA: Thousand Oaks, CA; 2003: 79-83.
6. Kohn L, Corrigan J, Donaldson M: **To err is human: building a safer health system. 1999 Washington**. *DC: Institute of Medicine*.
7. Bridges M, Diamond DL: **The financial impact of teaching surgical residents in the operating room**. *The American journal of surgery* 1999, **177**(1):28-32.
8. Farnworth LR, Lemay DE, Wooldridge T, Mabrey JD, Blaschak M, DeCoster TA, Wascher DC, Schenck R: **A comparison of operative times in arthroscopic ACL reconstruction between orthopaedic faculty and residents: the financial impact of orthopaedic surgical training in the operating room**. *Iowa Orthopaedic Journal* 2001, **21**:31-35.

9. Egan C, Elliott R, Fleming P: **European Working Time Directive and the use of simulators and models in Irish orthopaedics**. *Irish journal of medical science* 2012, **181**(1):143-146.
10. Zuckerman JD, Kubiak EN, Immerman I, DiCesare P: **The early effects of code 405 work rules on attitudes of orthopaedic residents and attending surgeons**. *J Bone Joint Surg Am* 2005, **87**(4):903-908.
11. Breen K, Hogan A, Mealy K: **The detrimental impact of the implementation of the European working time directive (EWTD) on surgical senior house officer (SHO) operative experience**. *Irish journal of medical science* 2013, **182**(3):383-387.
12. Skipworth R, Terrace J, Fulton L, Anderson D: **Basic surgical training in the era of the European Working Time Directive: what are the problems and solutions?** *Scottish medical journal* 2008, **53**(4):18-21.
13. Stefanidis D, Heniford BT: **The formula for a successful laparoscopic skills curriculum**. *Archives of Surgery* 2009, **144**(1):77-82.
14. Jenison EL, Gil KM, Lendvay TS, Guy MS: **Robotic surgical skills: acquisition, maintenance, and degradation**. *Jsls* 2012, **16**(2):218-228.
15. Dunn JC, Belmont PJ, Lanzi J, Martin K, Bader J, Owens B, Waterman BR: **Arthroscopic Shoulder Surgical Simulation Training Curriculum: Transfer Reliability and Maintenance of Skill Over Time**. *J Surg Educ* 2015, **72**(6):1118-1123.
16. Club TSS, Moore MJ, Bennett CL: **The learning curve for laparoscopic cholecystectomy**. *The American Journal of Surgery* 1995, **170**(1):55-59.
17. Shah J, Darzi A: **Virtual reality flexible cystoscopy: a validation study**. *BJU international* 2002, **90**(9):828-832.

18. Cass O: **Training to competence in gastrointestinal endoscopy: a plea for continuous measuring of objective end points**. *Endoscopy* 1999, **31**(09):751-754.
19. Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL: **Surgeon volume and operative mortality in the United States**. *New England Journal of Medicine* 2003, **349**(22):2117-2127.
20. Stavrakis AI, Ituarte PH, Ko CY, Yeh MW: **Surgeon volume as a predictor of outcomes in inpatient and outpatient endocrine surgery**. *Surgery* 2007, **142**(6):887-899.
21. Katz JN, Barrett J, Mahomed NN, Baron JA, Wright RJ, Losina E: **Association between hospital and surgeon procedure volume and the outcomes of total knee replacement**. *J bone joint surg am* 2004, **86**(9):1909-1916.
22. Seymour NE, Gallagher AG, Roman SA, O'brien MK, Bansal VK, Andersen DK, Satava RM: **Virtual reality training improves operating room performance: results of a randomized, double-blinded study**. *Annals of surgery* 2002, **236**(4):458-464.
23. Gallagher A, Satava R: **Objective assessment of experience, junior and novice laparoscopic performance with virtual reality: learning curves and reliability measures**. *Surg Endosc* 2002, **16**:1746-1752.
24. Gallagher A, Lederman A, McGlade K, Satava R, Smith C: **Discriminative validity of the Minimally Invasive Surgical Trainer in Virtual Reality (MIST-VR) using criteria levels based on expert performance**. *Surgical Endoscopy And Other Interventional Techniques* 2004, **18**(4):660-665.
25. Palter VN, Orzech N, Reznick RK, Grantcharov TP: **Validation of a structured training and assessment curriculum for technical skill acquisition in minimally invasive surgery: a randomized controlled trial.** *Annals of surgery* 2013, **257**(2):224-230.

26. Satava RM: **Virtual reality surgical simulator**. *Surgical endoscopy* 1993, **7**(3):203-205.
27. Stefanidis D, Scerbo MW, Montero PN, Acker CE, Smith WD: **Simulator training to automaticity leads to improved skill transfer compared with traditional proficiency-based training: a randomized controlled trial**. *Annals of surgery* 2012, **255**(1):30-37.
28. Price J, Naik V, Boodhwani M, Brandys T, Hendry P, Lam B-K: **A randomized evaluation of simulation training on performance of vascular anastomosis on a high-fidelity in vivo model: the role of deliberate practice**. *The Journal of thoracic and cardiovascular surgery* 2011, **142**(3):496-503.
29. Korndorffer JR, Dunne JB, Sierra R, Stefanidis D, Touchard CL, Scott DJ: **Simulator training for laparoscopic suturing using performance goals translates to the operating room**. *Journal of the American College of Surgeons* 2005, **201**(1):23-29.
30. Gallagher AG, Seymour NE, Jordan-Black J-A, Bunting BP, McGlade K, Satava RM: **Prospective, randomized assessment of transfer of training (ToT) and transfer effectiveness ratio (TER) of virtual reality simulation training for laparoscopic skill acquisition**. *Annals of surgery* 2013, **257**(6):1025-1031.
31. Butler A, Olson T, Koehler R, Nicandri G: **Do the skills acquired by novice surgeons using anatomic dry models transfer effectively to the task of diagnostic knee arthroscopy performed on cadaveric specimens?** *J Bone Joint Surg Am* 2013, **95**(3):e15(11-18).
32. Dwyer T, Slade Shantz J, Chahal J, Wasserstein D, Schachar R, Kulasegaram KM, Theodoropoulos J, Greben R, Ogilvie-Harris D: **Simulation of Anterior Cruciate Ligament Reconstruction in a Dry Model.** *Am J Sports Med* 2015, **43**(12):2997-3004.

33. Middleton RM, Alvand A, Garfjeld Roberts P, Hargrove C, Kirby G, Rees JL: **Simulation-Based Training Platforms for Arthroscopy: A Randomized Comparison of Virtual Reality Learning to Benchtop Learning**. *Arthroscopy* 2017.
34. Leong J, Leff D, Das A, Aggarwal R, Reilly P, Atkinson H, Emery R, Darzi A: **Validation of orthopaedic bench models for trauma surgery**. *Bone & Joint Journal* 2008, **90**(7):958-965.
35. Pedowitz RA, Nicandri GT, Angelo RL, Ryu RK, Gallagher AG: **Objective Assessment of Knot-Tying Proficiency With the Fundamentals of Arthroscopic Surgery Training Program Workstation and Knot Tester**. *Arthroscopy* 2015, **31**(10):1872-1879.
36. Holland JP, Waugh L, Horgan A, Paleri V, Deehan DJ: **Cadaveric hands-on training for surgical specialties: is this back to the future for surgical skills development?** *Journal of surgical education* 2011, **68**(2):110-116.
37. Camp CL, Krych AJ, Stuart MJ, Regnier TD, Mills KM, Turner NS: **Improving Resident Performance in Knee Arthroscopy: A Prospective Value Assessment of Simulators and Cadaveric Skills Laboratories**. *J Bone Joint Surg Am* 2016, **98**(3):220-225.
38. Chang J, Banaszek DC, Gambrel J, Bardana D: **Global Rating Scales and Motion Analysis Are Valid Proficiency Metrics in Virtual and Benchtop Knee Arthroscopy Simulators**. *Clin Orthop Relat Res* 2016, **474**(4):956-964.
39. Muller W, Bockholt U: **The virtual reality arthroscopy training simulator**. *Studies in health technology and informatics* 1998, **50**:13-19.
40. Cannon WD, Eckhoff DG, Garrett WE, Jr., Hunter RE, Sweeney HJ: **Report of a group developing a virtual reality simulator for arthroscopic surgery of the knee joint**. *Clin Orthop Relat Res* 2006, **442**:21-29.

41. Mabrey JD, Gillogly SD, Kasser JR, Sweeney HJ, Zarins B, Mevis H, Garrett WE, Jr., Poss R, Cannon WD: **Virtual reality simulation of arthroscopy of the knee**. *Arthroscopy* 2002, **18**(6):E28.
42. Gomoll AH, Pappas G, Forsythe B, Warner JJ: **Individual skill progression on a virtual reality simulator for shoulder arthroscopy: a 3-year follow-up study**. *Am J Sports Med* 2008, **36**(6):1139-1142.
43. Cannon WD, Nicandri GT, Reinig K, Mevis H, Wittstein J: **Evaluation of skill level between trainees and community orthopaedic surgeons using a virtual reality arthroscopic knee simulator**. *J Bone Joint Surg Am* 2014, **96**(7):e57.
44. Rose K, Pedowitz R: **Fundamental arthroscopic skill differentiation with virtual reality simulation**. *Arthroscopy* 2015, **31**(2):299-305.
45. Cannon WD, Garrett WE, Jr., Hunter RE, Sweeney HJ, Eckhoff DG, Nicandri GT, Hutchinson MR, Johnson DD, Bisson LJ, Bedi A *et al*: **Improving residency training in arthroscopic knee surgery with use of a virtual-reality simulator. A randomized blinded study**. *J Bone Joint Surg Am* 2014, **96**(21):1798-1806.
46. Akhtar K, Sugand K, Sperrin M, Cobb J, Standfield N, Gupte C: **Training safer orthopedic surgeons. Construct validation of a virtual-reality simulator for hip fracture surgery**. *Acta orthopaedica* 2015, **86**(5):616-621.
47. Garfjeld Roberts P, Guyver P, Baldwin M, Akhtar K, Alvand A, Price AJ, Rees JL: **Validation of the updated ArthroS simulator: face and construct validity of a passive haptic virtual reality simulator with novel performance metrics**. *Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA* 2016.
48. Aim F, Lonjon G, Hannouche D, Nizard R: **Effectiveness of Virtual Reality Training in Orthopaedic Surgery**. *Arthroscopy* 2016, **32**(1):224-232.

49. McCarthy AD, Moody L, Waterworth AR, Bickerstaff DR: **Passive haptics in a knee arthroscopy simulator: is it valid for core skills training?** *Clinical orthopaedics and related research* 2006, **442**:13-20.
50. Hall MP, Kaplan KM, Gorczynski CT, Zuckerman JD, Rosen JE: **Assessment of arthroscopic training in US orthopaedic surgery residency programs: a resident self-assessment**. *Bulletin of the NYU hospital for joint diseases* 2010, **68**(1):5-5.
51. Coughlin RP, Pauyo T, Sutton JC, 3rd, Coughlin LP, Bergeron SG: **A Validated Orthopaedic Surgical Simulation Model for Training and Evaluation of Basic Arthroscopic Skills**. *J Bone Joint Surg Am* 2015, **97**(17):1465-1471.
52. Martin KD, Belmont PJ, Schoenfeld AJ, Todd M, Cameron KL, Owens BD: **Arthroscopic basic task performance in shoulder simulator model correlates with similar task performance in cadavers**. *J Bone Joint Surg Am* 2011, **93**(21):e1271-1275.
53. Colaco HB, Hughes K, Pearse E, Arnander M, Tennent D: **Construct Validity, Assessment of the Learning Curve, and Experience of Using a Low-Cost Arthroscopic Surgical Simulator**. *J Surg Educ* 2017, **74**(1):47-54.
54. Rebolledo BJ, Hammann-Scala J, Leali A, Ranawat AS: **Arthroscopy skills development with a surgical simulator: a comparative study in orthopaedic surgery residents**. *Am J Sports Med* 2015, **43**(6):1526-1529.
55. Peres LR, Junior WM, Coelho G, Lyra M: **A new simulator model for knee arthroscopy procedures**. *Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA* 2016.
56. Martin KD, Cameron K, Belmont PJ, Schoenfeld A, Owens BD: **Shoulder arthroscopy simulator performance correlates with resident and shoulder arthroscopy experience**. *J Bone Joint Surg Am* 2012, **94**(21):e160.

57. Henn RF, 3rd, Shah N, Warner JJ, Gomoll AH: **Shoulder arthroscopy simulator training improves shoulder arthroscopy performance in a cadaveric model**. *Arthroscopy* 2013, **29**(6):982-985.
58. Pollard TC, Khan T, Price AJ, Gill HS, Glyn-Jones S, Rees JL: **Simulated hip arthroscopy skills: learning curves with the lateral and supine patient positions: a randomized trial**. *J Bone Joint Surg Am* 2012, **94**(10):e68.
59. Tuijthof GJ, van Sterkenburg MN, Sierevelt IN, van Oldenrijk J, Van Dijk CN, Kerkhoffs GM: **First validation of the PASSPORT training environment for arthroscopic skills**. *Knee surgery, sports traumatology, arthroscopy : official journal of the ESSKA* 2010, **18**(2):218-224.
60. Waterman BR, Martin KD, Cameron KL, Owens BD, Belmont PJ, Jr.: **Simulation Training Improves Surgical Proficiency and Safety During Diagnostic Shoulder Arthroscopy Performed by Residents**. *Orthopedics* 2016, **39**(3):e479-485.
61. Middleton RM, Baldwin MJ, Akhtar K, Alvand A, Rees JL: **Which Global Rating Scale? A Comparison of the ASSET, BAKSSS, and IGARS for the Assessment of Simulated Arthroscopic Skills**. *J Bone Joint Surg Am* 2016, **98**(1):75-81.
62. Likert R: **A technique for the measurement of attitudes**. *Archives of psychology* 1932.
63. Gallagher AG, O'Sullivan GC, Leonard G, Bunting BP, McGlade KJ: **Objective structured assessment of technical skills and checklist scales reliability compared for high stakes assessments**. *ANZ journal of surgery* 2014, **84**(7-8):568-573.
64. Gallagher A, Ritter E, Satava R: **Fundamental principles of validation, and reliability: rigorous science for the assessment of surgical education and training**. *Surgical Endoscopy and other Interventional Techniques* 2003, **17**(10):1525-1529.

65. Angelo RL, Ryu RK, Pedowitz RA, Gallagher AG: **Metric Development for an Arthroscopic Bankart Procedure: Assessment of Face and Content Validity**. *Arthroscopy* 2015, **31**(8):1430-1440.
66. Angelo RL, Ryu RK, Pedowitz RA, Gallagher AG: **The Bankart Performance Metrics Combined With a Cadaveric Shoulder Create a Precise and Accurate Assessment Tool for Measuring Surgeon Skill**. *Arthroscopy* 2015, **31**(9):1655-1670.
67. Gomoll AH, O'Toole RV, Czarnecki J, Warner JJ: **Surgical experience correlates with performance on a virtual reality simulator for shoulder arthroscopy.** *The American journal of sports medicine* 2007, **35**(6):883-888.
68. Pedowitz RA, Esch J, Snyder S: **Evaluation of a virtual reality simulator for arthroscopy skills development**. *Arthroscopy* 2002, **18**(6):E29.
69. Srivastava S, Youngblood PL, Rawn C, Hariri S, Heinrichs W, Ladd AL: **Initial evaluation of a shoulder arthroscopy simulator: establishing construct validity**. *Journal of shoulder and elbow surgery* 2004, **13**(2):196-205.
70. Larsen CR, Soerensen JL, Grantcharov TP, Dalsgaard T, Schouenborg L, Ottosen C, Schroeder TV, Ottesen BS: **Effect of virtual reality training on laparoscopic surgery: randomised controlled trial**. *Bmj* 2009, **338**:b1802.
71. Sturm LP, Windsor JA, Cosman PH, Cregan P, Hewett PJ, Maddern GJ: **A systematic review of skills transfer after surgical simulation training.** *Annals of surgery* 2008, **248**(2):166-179.
72. Ikonen T, Antikainen T, Silvennoinen M, Isojärvi J, Mäkinen E, Scheinin T: **Virtual reality simulator training of laparoscopic cholecystectomies—a systematic review**. *Scandinavian Journal of Surgery* 2012, **101**(1):5-12.
73. Feldman LS, Sherman V, Fried GM: **Using simulators to assess laparoscopic competence: ready for widespread use?** In.: Mosby; 2004.

74. Howells NR, Gill HS, Carr AJ, Price AJ, Rees JL: **Transferring simulated arthroscopic skills to the operating theatre: a randomised blinded study**. *The Journal of bone and joint surgery British volume* 2008, **90**(4):494-499.
75. Gallagher A, McClure N, McGuigan J, Ritchie K, Sheehy N: **An ergonomic analysis of the fulcrum effect in the acquisition of endoscopic skills**. *Endoscopy* 1998, **30**(07):617-620.
76. Gallagher AG, Smith CD: **Human-factors lessons learned from the minimally invasive surgery revolution**. In: *Seminars in Laparoscopic Surgery: 2003*: Sage Publications Sage CA: Thousand Oaks, CA; 2003: 127-139.
77. Broadbent DE: **Failures of attention in selective listening**: MRC.; 1952.
78. Baddeley A: **Working memory: looking back and looking forward**. *Nature reviews neuroscience* 2003, **4**(10):829-839.
79. Gallagher AG, O'Sullivan GC: **Fundamentals of surgical simulation : principles and practices**. London: Springer; 2012.
80. Gallagher AG, Satava RM, O'Sullivan GC: **Attentional capacity: an essential aspect of surgeon performance**. *Annals of surgery* 2015, **261**(3):e60-e61.
81. Bankart AS: **The pathology and treatment of recurrent dislocation of the shoulder‐joint**. *British Journal of Surgery* 1938, **26**(101):23-29.
82. Taylor DC, Arciero RA: **Pathologic Changes Associated with Shoulder Dislocations Arthroscopic and Physical Examination Findings in First-Time, Traumatic Anterior Dislocations**. *The American journal of sports medicine* 1997, **25**(3):306-311.
83. Rowe C, Patel D, Southmayd W: **The Bankart procedure: a long-term end-result study**. *J Bone Joint Surg Am* 1978, **60**(1):1-16.

84. Thomas SC, Matsen FA: **An approach to the repair of avulsion of the glenohumeral ligaments in the management of traumatic anterior glenohumeral instability**. *J Bone Joint Surg Am* 1989, **71**(4):506-513.
85. Richmond JC, Donaldson WR, Fu F, Harner CD: **Modification of the Bankart reconstruction with a suture anchor report of a new technique**. *The American journal of sports medicine* 1991, **19**(4):343-346.
86. Karlsson J, Järvholm U, Swärd L, Lansinger O: **Repair of Bankart lesions with a suture anchor in recurrent dislocation of the shoulder**. *Scandinavian journal of medicine & science in sports* 1995, **5**(3):170-174.
87. Wolf EM, Wilk RM, Richmond JC: **Arthroscopic Bankart repair using suture anchors**. *Operative techniques in orthopaedics* 1991, **1**(2):184-191.
88. Bacilla P, Field L, Savoie Fr: **Arthroscopic Bankart repair in a high demand patient population**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 1997, **13**(1):51-60.
89. Angelo RL: **Arthroscopic bankart repair for unidirectional shoulder instability**. *Instr Course Lect* 2009, **58**:305-313.
90. Abrams JS, Bradley JP, Angelo RL, Burks R: **Arthroscopic management of shoulder instabilities: anterior, posterior, and multidirectional**. *Instructional course lectures* 2009, **59**:141-155.
91. Angelo RL: **Arthroscopic Setup: Approaches and Tips for Success**. *Operative Arthroscopy* 2012:1.
92. Angelo RL: **The overhead athlete: how to examine, test, and treat shoulder injuries. Intra-articular pathology**. *Arthroscopy* 2003, **19 Suppl 1**:47-50.
93. Angelo RL, Pedowitz RA, Ryu RK, Gallagher AG: **The Bankart Performance Metrics Combined With a Shoulder Model Simulator Create a Precise and Accurate Training Tool for Measuring Surgeon Skill**. *Arthroscopy* 2015, **31**(9):1639-1654.

94. Angelo RL, Ryu RK, Pedowitz RA, Beach W, Burns J, Dodds J, Field L, Getelman M, Hobgood R, McIntyre L *et al*: **A Proficiency-Based Progression Training Curriculum Coupled With a Model Simulator Results in the Acquisition of a Superior Arthroscopic Bankart Skill Set**. *Arthroscopy* 2015, **31**(10):1854-1871.
95. Aggarwal R, Crochet P, Dias A, Misra A, Ziprin P, Darzi A: **Development of a virtual reality training curriculum for laparoscopic cholecystectomy**. *British Journal of Surgery* 2009, **96**(9):1086-1093.
96. Asch DA, Weinstein DF: **Innovation in medical education**. *New England Journal of Medicine* 2014, **371**(9):794-795.
97. Satava RM: **Modeling and simulation in healthcare future directions**. *10th Annual MOVES Research and Education Summit, Monterey, California* 2010.
98. Levitz CL, Dugas J, Andrews JR: **The use of arthroscopic thermal capsulorrhaphy to treat internal impingement in baseball players**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2001, **17**(6):573-577.
99. Gerber A, Warner JJ: **Thermal capsulorrhaphy to treat shoulder instability**. *Clinical orthopaedics and related research* 2002, **400**:105-116.
100. Abrams J: **Thermal capsulorrhaphy for instability of the shoulder: concerns and applications of the heat probe**. *Instructional course lectures* 2000, **50**:29-36.
101. Savoie FH, Field LD, Atchinson S: **Anterior superior instability with rotator cuff tearing: SLAC lesion**. *Orthopedic Clinics* 2001, **32**(3):457-461.
102. Snyder SJ, Karzel RP, Del Pizzo W, Ferkel RD, Friedman MJ: **SLAP lesions of the shoulder**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 1990, **6**(4):274-279.

103. Kuhn J, Huston L, Blasier R, Soslowsky L: **Ligamentous restraints and muscle effects limiting external rotation of the glenohumeral joint in the neutral and abducted positions.** *J Shoulder Elbow Surg* 1999, **8**:664.
104. Fleisig GS, Andrews JR, Dillman CJ, Escamilla RF: **Kinetics of baseball pitching with implications about injury mechanisms**. *The American journal of sports medicine* 1995, **23**(2):233-239.
105. Andrews JR, Carson JR WG, Mcleod WD: **Glenoid labrum tears related to the long head of the biceps**. *The American journal of sports medicine* 1985, **13**(5):337-341.
106. Mileski RA, Snyder SJ: **Superior labral lesions in the shoulder: pathoanatomy and surgical management**. *Journal of the American Academy of Orthopaedic Surgeons* 1998, **6**(2):121-131.
107. Walch G, Boileau P, Noel E, Donell S: **Impingement of the deep surface of the supraspinatus tendon on the posterosuperior glenoid rim: an arthroscopic study**. *Journal of Shoulder and Elbow Surgery* 1992, **1**(5):238-245.
108. Jobe CM: **Posterior superior glenoid impingement: expanded spectrum**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 1995, **11**(5):530-536.
109. Davidson PA, Elattrache NS, Jobe CM, Jobe FW: **Rotator cuff and posterior-superior glenoid labrum injury associated with increased glenohumeral motion: a new site of impingement**. *Journal of Shoulder and Elbow Surgery* 1995, **4**(5):384-390.
110. Harryman DT, Sidles JA, Clark JM, McQuade KJ, Gibb TD, Matsen FA: **Translation of the humeral head on the glenoid with passive glenohumeral motion.** *J Bone Joint Surg Am* 1990, **72**(9):1334-1343.
111. Kibler WB: **The role of the scapula in athletic shoulder function.** *The American journal of sports medicine* 1998, **26**(2):325-337.

112. Burkhart SS, Morgan CD, Kibler WB: **The disabled throwing shoulder: spectrum of pathology Part I: pathoanatomy and biomechanics**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2003, **19**(4):404-420.
113. Burkhart SS, Morgan CD: **The peel-back mechanism: its role in producing and extending posterior type II SLAP lesions and its effect on SLAP repair rehabilitation**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 1998, **14**(6):637-640.
114. Morgan CD, Burkhart SS, Palmeri M, Gillespie M: **Type II SLAP lesions: three subtypes and their relationships to superior instability and rotator cuff tears**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 1998, **14**(6):553-565.
115. Ticker JB, Beim GM, Warner JJ: **Recognition and treatment of refractory posterior capsular contracture of the shoulder**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2000, **16**(1):27-34.
116. Jobe FW, Giangarra CE, Kvitne RS, Glousman RE: **Anterior capsulolabral reconstruction of the shoulder in athletes in overhand sports**. *The American journal of sports medicine* 1991, **19**(5):428-434.
117. Andrews JR, Angelo RL: **Shoulder arthroscopy for the throwing athlete**. *Techniques in Orthopaedics* 1988, **3**(1):75-82.
118. Nakajima T, Rokuuma N, Hamada K, Tomatsu T, Fukuda H: **Histologic and biomechanical characteristics of the supraspinatus tendon: reference to rotator cuff tearing**. *Journal of Shoulder and Elbow Surgery* 1994, **3**(2):79-87.
119. Payne LZ, Altchek DW, Craig EV, Warren RF: **Arthroscopic treatment of partial rotator cuff tears in young athletes: a preliminary report**. *The American Journal of Sports Medicine* 1997, **25**(3):299-305.

120. Jobe F, Kvitne R, Giangarra C: **Shoulder pain in the overhand or throwing athlete. The relationship of anterior instability and rotator cuff impingement**. *Orthopaedic review* 1989, **18**(9):963-975.

121. Correa MC, Gonçalves LB, Andrade RP, Carvalho LH: **Beach chair position with instrumental distraction for arthroscopic and open shoulder surgeries**. *Journal of Shoulder and Elbow Surgery* 2008, **17**(2):226-230.

122. Pohl A, Cullen DJ: **Cerebral ischemia during shoulder surgery in the upright position: a case series**. *Journal of clinical anesthesia* 2005, **17**(6):463-469.

123. Papadonikolakis A, Wiesler ER, Olympio MA, Poehling GG: **Avoiding catastrophic complications of stroke and death related to shoulder surgery in the sitting position**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2008, **24**(4):481-482.

124. Weber SC, Jain R: **Scalene regional anesthesia for shoulder surgery in a community setting: an assessment of risk**. *J Bone Joint Surg Am* 2002, **84**(5):775-779.

125. Bishop JY, Sprague M, Gelber J, Krol M, Rosenblatt MA, Gladstone J, Flatow EL: **Interscalene regional anesthesia for shoulder surgery**. *J Bone Joint Surg Am* 2005, **87**(5):974-979.

126. Urban MK, Urquhart B: **Evaluation of brachial plexus anesthesia for upper extremity surgery**. *Regional Anesthesia and Pain Medicine* 1994, **19**(3):175-182.

127. Ritchie ED, Tong D, Chung F, Norris AM, Miniaci A, Vairavanathan SD: **Suprascapular nerve block for postoperative pain relief in arthroscopic shoulder surgery: a new modality?** *Anesthesia & Analgesia* 1997, **84**(6):1306-1312.

128. Matsumoto D, Suenaga N, Oizumi N, Hisada Y, Minami A: **A new nerve block procedure for the suprascapular nerve based on a cadaveric study**. *Journal of Shoulder and Elbow Surgery* 2009, **18**(4):607-611.
129. Barber FA: **Suprascapular nerve block for shoulder arthroscopy**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2005, **21**(8):1015. e1011-1015. e1014.
130. Nottage W: **Arthroscopic portals: anatomy at risk**. *The Orthopedic clinics of North America* 1993, **24**(1):19-26.
131. Stanish WD, Peterson DC: **Shoulder arthroscopy and nerve injury: pitfalls and prevention**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 1995, **11**(4):458-466.
132. Lo IK, Lind CC, Burkhart SS: **Glenohumeral arthroscopy portals established using an outside–In technique: neurovascular anatomy at risk.** *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2004, **20**(6):596-602.
133. Meyer M, Graveleau N, Hardy P, Landreau P: **Anatomic risks of shoulder arthroscopy portals: anatomic cadaveric study of 12 portals.** *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2007, **23**(5):529-536.
134. Woolf SK, Guttmann D, Karch MM, Graham RD, Reid JB, Lubowitz JH: **The superior-medial shoulder arthroscopy portal is safe**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2007, **23**(3):247-250.
135. Baker CL, Uribe JW, Whitman C: **Arthroscopic evaluation of acute initial anterior shoulder dislocations**. *The American journal of sports medicine* 1990, **18**(1):25-28.
136. Perthes G: **Über operationen bei habitueller schulterluxation.** *Dtsch Z Chir* 1906, **85**(199):10.1007.
137. Levine WN, Richmond JC, Donaldson WR: **Use of the suture anchor in open Bankart reconstruction: a follow-up report**. *The American journal of sports medicine* 1994, **22**(5):723-726.

138. Caspari RB: **Arthroscopic reconstruction for anterior shoulder instability**. *Techniques in Orthopaedics* 1988, **3**(1):59-66.
139. McIntyre L, Caspari R: **The rationale and technique for arthroscopic reconstruction of anterior shoulder instability using multiple sutures**. *The Orthopedic clinics of North America* 1993, **24**(1):55-58.
140. Burkhart SS, De Beer JF: **Traumatic glenohumeral bone defects and their relationship to failure of arthroscopic Bankart repairs: significance of the inverted-pear glenoid and the humeral engaging Hill-Sachs lesion**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2000, **16**(7):677-694.
141. Sperber A, Hamberg P, Karlsson J, Swärd L, Wredmark T: **Comparison of an arthroscopic and an open procedure for posttraumatic instability of the shoulder: a prospective, randomized multicenter study**. *Journal of shoulder and elbow surgery* 2001, **10**(2):105-108.
142. Warner JJ: **Treatment options for anterior instability: Open versus arthroscopic**. *Operative Techniques in Orthopaedics* 1995, **5**(3):233-237.
143. Cole BJ, Warner JJ: **Arthroscopic versus open Bankart repair for traumatic anterior shoulder instability**. *Clinics in sports medicine* 2000, **19**(1):19-48.
144. Kim S-H, Ha K-I, Kim S-H: **Bankart repair in traumatic anterior shoulder instability: open versus arthroscopic technique**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2002, **18**(7):755-763.
145. Weber S: **A prospective evaluation comparing arthroscopic and open treatment in the management of recurrent anterior glenohumeral dislocation**. *Orthop Trans* 1991, **15**:763.

146. Mologne TS, McBride MT, Lapoint JM: **Assessment of failed arthroscopic anterior labral repairs: findings at open surgery**. *The American journal of sports medicine* 1997, **25**(6):813-817.
147. Kirkley A, Griffin S, Richards C, Miniaci A, Mohtadi N: **Prospective randomized clinical trial comparing the effectiveness of immediate arthroscopic stabilization versus immobilization and rehabilitation in first traumatic anterior dislocations of the shoulder**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 1999, **15**(5):507-514.
148. Bottoni CR, Wilckens JH, DeBerardino TM, D'Alleyrand J-CG, Rooney RC, Harpstrite JK, Arciero RA: **A prospective, randomized evaluation of arthroscopic stabilization versus nonoperative treatment in patients with acute, traumatic, first-time shoulder dislocations.** *The American Journal of Sports Medicine* 2002, **30**(4):576-580.
149. Abrams JS, Savoie III FH, Tauro JC, Bradley JP: **Recent advances in the evaluation and treatment of shoulder instability: anterior, posterior, and multidirectional**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2002, **18**(9):1-13.
150. Wolf EM, Eakin CL: **Arthroscopic capsular plication for posterior shoulder instability**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 1998, **14**(2):153-163.
151. Antoniou J, Duckworth DT, Harryman DT: **Capsulolabral augmentation for the management of posteroinferior instability of the shoulder**. *J Bone Joint Surg Am* 2000, **82**(9):1220-1220.
152. Lu Y, Hayashi K, Edwards RB, Fanton GS, Thabit G, Markel MD: **The Effect of Monopolar Radiofrequency Treatment Pattern on Joint Capsular Healing In Vitro and In Vivo Studies Using an Ovine Model**. *The American journal of sports medicine* 2000, **28**(5):711-719.

153. Marquardt B, Witt K-A, Liem D, Steinbeck J, Pötzl W: **Arthroscopic Bankart repair in traumatic anterior shoulder instability using a suture anchor technique**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2006, **22**(9):931-936.
154. Carreira DS, Mazzocca AD, Oryhon J, Brown FM, Hayden JK, Romeo AA: **A prospective outcome evaluation of arthroscopic Bankart repairs minimum 2-year follow-up**. *The American Journal of Sports Medicine* 2006, **34**(5):771-777.
155. Kim S-H, Ha K-I, Cho Y-B, Ryu B-D, Oh I: **Arthroscopic anterior stabilization of the shoulder**. *J Bone Joint Surg Am* 2003, **85**(8):1511-1518.
156. Kim S-H, Ha K-I, Kim Y-M: **Arthroscopic revision Bankart repair: a prospective outcome study**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2002, **18**(5):469-482.
157. Fabbriciani C, Milano G, Demontis A, Fadda S, Ziranu F, Mulas PD: **Arthroscopic versus open treatment of Bankart lesion of the shoulder: a prospective randomized study**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2004, **20**(5):456-462.
158. Rhee YG, Ha JH, Cho NS: **Anterior shoulder stabilization in collision athletes arthroscopic versus open Bankart repair**. *The American journal of sports medicine* 2006, **34**(6):979-985.
159. Ide J, Maeda S, Takagi K: **Arthroscopic Bankart Repair Using Suture Anchors in Athletes Patient Selection and Postoperative Sports Activity**. *The American journal of sports medicine* 2004, **32**(8):1899-1905.
160. Marquardt B, Hurschler C, Schneppendahl J, Witt K-A, Pötzl W, Steinbeck J: **Quantitative assessment of glenohumeral translation after anterior shoulder dislocation and subsequent arthroscopic Bankart repair**. *The American journal of sports medicine* 2006, **34**(11):1756-1762.

161. Mazzocca AD, Brown FM, Carreira DS, Hayden J, Romeo AA: **Arthroscopic anterior shoulder stabilization of collision and contact athletes**. *The American journal of sports medicine* 2005, **33**(1):52-60.
162. Jones KJ, Wiesel B, Ganley TJ, Wells L: **Functional outcomes of early arthroscopic Bankart repair in adolescents aged 11 to 18 years**. *Journal of Pediatric Orthopaedics* 2007, **27**(2):209-213.
163. Habermeyer P, Gleyze P, Rickert M: **Evolution of lesions of the labrum-ligament complex in posttraumatic anterior shoulder instability: a prospective study**. *Journal of Shoulder and Elbow Surgery* 1999, **8**(1):66-74.
164. Boileau P, Villalba M, Héry J-Y, Balg F, Ahrens P, Neyton L: **Risk factors for recurrence of shoulder instability after arthroscopic Bankart repair**. *J Bone Joint Surg Am* 2006, **88**(8):1755-1763.
165. Sugaya H, Moriishi J, Kanisawa I, Tsuchiya A: **Arthroscopic osseous Bankart repair for chronic recurrent traumatic anterior glenohumeral instability**. *J Bone Joint Surg Am* 2005, **87**(8):1752-1760.
166. Mologne TS, Provencher MT, Menzel KA, Vachon TA, Dewing CB: **Arthroscopic Stabilization in Patients With an Inverted Pear Glenoid Results in Patients With Bone Loss of the Anterior Glenoid**. *The American journal of sports medicine* 2007, **35**(8):1276-1283.
167. Porcellini G, Paladini P, Campi F, Paganelli M: **Long-term outcome of acute versus chronic bony Bankart lesions managed arthroscopically**. *The American journal of sports medicine* 2007, **35**(12):2067-2072.
168. Kropf EJ, Sekiya JK: **Osteoarticular allograft transplantation for large humeral head defects in glenohumeral instability**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2007, **23**(3):322. e321-322. e325.

169. Kazel MD, Sekiya JK, Greene JA, Bruker CT: **Percutaneous correction (humeroplasty) of humeral head defects (Hill-Sachs) associated with anterior shoulder instability: a cadaveric study**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2005, **21**(12):1473-1478.

170. Re P, Gallo RA, Richmond JC: **Transhumeral head plasty for large Hill-Sachs lesions**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2006, **22**(7):798. e791-798. e794.

171. Rhee YG, Ha JH, Park KJ: **Clinical outcome of anterior shoulder instability with capsular midsubstance tear: A comparison of isolated midsubstance tear and midsubstance tear with Bankart lesion**. *Journal of shoulder and elbow surgery* 2006, **15**(5):586-590.

172. Richards DP, Burkhart SS: **Arthroscopic humeral avulsion of the glenohumeral ligaments (HAGL) repair**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2004, **20**:134-141.

173. Thal R, Nofziger M, Bridges M, Kim JJ: **Arthroscopic Bankart repair using Knotless or BioKnotless suture anchors: 2-to 7-year results**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2007, **23**(4):367-375.

174. Cho NS, Lubis AM, Ha JH, Rhee YG: **Clinical results of arthroscopic Bankart repair with knot-tying and knotless suture anchors**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2006, **22**(12):1276-1282.

175. Marquardt B, Garmann S, Schulte T, Witt K-A, Steinbeck J, Pötzl W: **Outcome after failed traumatic anterior shoulder instability repair with and without surgical revision**. *Journal of Shoulder and Elbow Surgery* 2007, **16**(6):742-747.

176. Adams JA: **A closed-loop theory of motor learning**. *Journal of motor behavior* 1971, **3**(2):111-149.

177. Adams JA: **Historical review and appraisal of research on the learning, retention, and transfer of human motor skills**. *Psychol Bull* 1987, **101**(1):41-74.
178. Adams JA: **The changing face of motor learning**. *Human movement science* 1990, **9**(3-5):209-220.
179. Dreyfus HL, Dreyfus SE, Athanasiou T: **Mind over machine**: Free press New York; 1986.
180. Gallagher AG, O'Sullivan GC: **Fundamentals of Surgical Simulation; Principles & Practices** London: Springer Verlag; 2011.
181. Gallagher AG: **Metric-based simulation training to proficiency in medical education:-What it is and how to do it**. *Ulster Med J* 2012, **81**(3):107-113.
182. Andersen DK: **How can educators use simulation applications to teach and assess surgical judgment?** *Academic medicine : journal of the Association of American Medical Colleges* 2012, **87**(7):934-941.
183. Burks RT, Presson AP, Weng H-Y: **An analysis of technical aspects of the arthroscopic Bankart procedure as performed in the United States.** *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2014, **30**(10):1246-1253.
184. Morgan CD, Bodenstab AB: **Arthroscopic Bankart Suture Repair: Technique and Early Results.** *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2010, **26**(6):819-820.
185. Streubel PN, Krych AJ, Simone JP, Dahm DL, Sperling JW, Steinmann SP, O’Driscoll SW, Sanchez-Sotelo J: **Anterior glenohumeral instability: a pathology-based surgical treatment strategy**. *Journal of the American Academy of Orthopaedic Surgeons* 2014, **22**(5):283-294.
186. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Andersen DK, Satava RM: **Analysis of errors in laparoscopic surgical procedures**. *Surg Endosc* 2004, **18**(4):592-595.

187. Cuschieri A, Francis N, Crosby J, Hanna GB: **What do master surgeons think of surgical competence and revalidation?** *The American Journal of Surgery* 2001, **182**(2):110-116.
188. Ee GW, Mohamed S, Tan AH: **Long term results of arthroscopic Bankart repair for traumatic anterior shoulder instability**. *Journal of orthopaedic surgery and research* 2011, **6**(1):28.
189. van der Linde JA, van Kampen DA, Willems JW: **How should we define failure after surgical shoulder stabilization?** *Arthroscopy* 2013, **29**(10):1602.
190. Waterman BR, Burns TC, McCriskin B, Kilcoyne K, Cameron KL, Owens BD: **Outcomes after Bankart repair in a military population: Predictors for surgical revision and long-term disability**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2014, **30**(2):172-177.
191. Shibata H, Gotoh M, Mitsui Y, Kai Y, Nakamura H, Kanazawa T, Okawa T, Higuchi F, Shirahama M, Shiba N: **Risk factors for shoulder re-dislocation after arthroscopic Bankart repair**. *Journal of orthopaedic surgery and research* 2014, **9**(1):53.
192. Angelo RL: **Controversies in arthroscopic shoulder surgery: arthroscopic versus open bankart repair, thermal treatment of capsular tissue, acromioplasties--are they necessary?** *Arthroscopy* 2003, **19 Suppl 1**:224-228.
193. Ryu RK: **Arthroscopic approach to traumatic anterior shoulder instability**. *Arthroscopy* 2003, **19 Suppl 1**:94-101.
194. Cole BJ, L'insalata J, Irrgang J, Warner JJ: **Comparison of arthroscopic and open anterior shoulder stabilization**. *J Bone Joint Surg Am* 2000, **82**(8):1108-1108.
195. McClusky DA, Gallagher AG, Ritter EM, Lederman AB, Van Sickle KR, Baghai M, Smith CD: **Virtual reality training improves junior residents' operating room performance: results of a prospective, randomized, double-blinded study of the complete laparoscopic cholecystectomy**. *Journal of the American College of Surgeons* 2004, **199**(3):73.

196. Van Sickle KR, Smith B, Mcclusky DA, Baghai M, Smith CD, Gallagher AG: **Evaluation of a tensiometer to provide objective feedback in knot-tying performance**. *The American Surgeon* 2005, **71**(12):1018-1023.
197. Van Sickle K, Gallagher A, Smith C: **The effect of escalating feedback on the acquisition of psychomotor skills for laparoscopy**. *Surgical endoscopy* 2007, **21**(2):220-224.
198. Van Sickle KR, Ritter EM, Baghai M, Goldenberg AE, Huang I-P, Gallagher AG, Smith CD: **Prospective, randomized, double-blind trial of curriculum-based training for intracorporeal suturing and knot tying**. *Journal of the American College of Surgeons* 2008, **207**(4):560-568.
199. Neary PC, Boyle E, Delaney CP, Senagore AJ, Keane FB, Gallagher AG: **Construct validation of a novel hybrid virtual-reality simulator for training and assessing laparoscopic colectomy; results from the first course for experienced senior laparoscopic surgeons**. *Surgical endoscopy* 2008, **22**(10):2301-2309.
200. Nicholson WJ, Cates CU, Patel AD, Niazi K, Palmer S, Helmy T, Gallagher AG: **Face and content validation of virtual reality simulation for carotid angiography: results from the first 100 physicians attending the Emory NeuroAnatomy Carotid Training (ENACT) program**. *Simulation in Healthcare* 2006, **1**(3):147-150.
201. Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G, Smith CD, Satava RM: **Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training**. *Annals of surgery* 2005, **241**(2):364-372.
202. Healy GB: **The college should be instrumental in adapting simulators to education**. *Bull Am Coll Surg* 2002, **87**(11):10-11.

203. Pellegrini CA: **Surgical education in the United States: navigating the white waters.** *Annals of surgery* 2006, **244**(3):335-342.
204. Sachdeva AK, Pellegrini CA, Johnson KA: **Support for simulation-based surgical education through American College of Surgeons–accredited education institutes**. *World journal of surgery* 2008, **32**(2):196-207.
205. Selzer DJ, Dunnington GL: **Surgical Skills Simulation: A Shift in the Conversation.** *Ann Surg* 2013 **257**(4):594-595.
206. Zendejas B, Brydges R, Hamstra SJ, DA. C: **State of the Evidence on Simulation-Based Training for Laparoscopic Surgery: A Systematic Review.** *Ann Surg* 2013, **257**(4):586-593.
207. Grantcharov TP, Kristiansen V, Bendix J, Bardram L, Rosenberg J, Funch ‐ Jensen P: **Randomized clinical trial of virtual reality simulation for laparoscopic skills training**. *British Journal of Surgery* 2004, **91**(2):146-150.
208. Ahlberg G, Enochsson L, Gallagher AG, Hedman L, Hogman C, McClusky DA, Ramel S, Smith CD, Arvidsson D: **Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies**. *The American journal of surgery* 2007, **193**(6):797-804.
209. Morgan CD, Bodenstab AB: **Arthroscopic Bankart suture repair: technique and early results**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 1987, **3**(2):111-122.
210. **American Education and Research Association: Standards for Education and Psychological Testing.** 2014.
211. Rossi MJ, Lubowitz JH, Provencher MT, Poehling GG: **Precision versus accuracy: A case for common sense**. In.: WB Saunders; 2012.
212. Gallagher AG: **Metric-based simulation training to proficiency in medical education:-What it is and how to do it**. *Ulster Med J* 2012, **81**(3):107-113.

213. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Andersen DK, Satava RM: **Analysis of errors in laparoscopic surgical procedures**. *Surg Endosc* 2004, **18**(4):592-595.

214. Reinhardt-Rutland A, Gallagher A: **Visual depth perception in minimally invasive surgery**. 1995.

215. Kohn LT, Corrigan JM, Donaldson MS: **To err is human: building a safer health system**. *Washington DC* 2000:196–197.

216. Smith R: **All changed, changed utterly**. *Br Med J* 1998, **316**(7149):1917 - 1918.

217. Barden CB, Specht MC, McCarter MD, Daly JM, Fahey TJ: **Effects of limited work hours on surgical training.** *J Am Coll Surg* 2002, **195**(4):531-538.

218. Temple J: **Time for Training: A Review of the Impact of the European Working Time Directive on the Quality of Training** In.; 2010.

219. Bell Jr RH, Biester TW, Tabuenca A, Rhodes RS, Cofer JB, Britt LD, Lewis Jr FR: **Operative experience of residents in US general surgery programs: a gap between expectation and experience.** *Ann Surg* 2009, **249**(5):719 - 724.

220. Birkmeyer JD, Finks JF, O'reilly A, Oerline M, Carlin AM, Nunn AR, Dimick J, Banerjee M, Birkmeyer NJ: **Surgical skill and complication rates after bariatric surgery**. *New England Journal of Medicine* 2013, **369**(15):1434-1442.

221. Mishra DK, Cannon Jr WD, Lucas DJ, Belzer JP: **Elongation of arthroscopically tied knots**. *The American journal of sports medicine* 1997, **25**(1):113-117.

222. Swan KG, Baldini T, McCarty EC: **Arthroscopic suture material and knot type an updated biomechanical analysis.** *The American journal of sports medicine* 2009, **37**(8):1578-1585.

223. Hanypsiak BT, DeLong JM, Simmons L, Lowe W, Burkhart S: **Knot strength varies widely among expert arthroscopists**. *The American journal of sports medicine* 2014:0363546514535554.

224. Abbi G, Espinoza L, Odell T, Mahar A, Pedowitz R: **Evaluation of 5 knots and 2 suture materials for arthroscopic rotator cuff repair: very strong sutures can still slip**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2006, **22**(1):38-43.

225. Baums M, Sachs C, Kostuj T, Schmidt-Horlohé K, Schultz W, Klinger H-M: **Mechanical testing of different knot types using high-performance suture material**. *Knee Surgery, Sports Traumatology, Arthroscopy* 2015, **23**(5):1351-1358.

226. Ilahi OA, Younas SA, Alexander J, Noble PC: **Cyclic testing of arthroscopic knot security**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2004, **20**(1):62-68.

227. Ilahi OA, Younas SA, Ho DM, Noble PC: **Security of knots tied with ethibond, fiberwire, orthocord, or ultrabraid**. *The American journal of sports medicine* 2008, **36**(12):2407-2414.

228. Mahar AT, Moezzi DM, Serra-Hsu F, Pedowitz RA: **Comparison and performance characteristics of 3 different knots when tied with 2 suture materials used for shoulder arthroscopy**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2006, **22**(6):610-614. e612.

229. Riboh JC, Heckman DS, Glisson RR, Moorman III CT: **Shortcuts in arthroscopic knot tying: Do they affect knot and loop security?** *The American journal of sports medicine* 2012, **40**(7):1572-1577.

230. Shah MR, Strauss EJ, Kaplan K, Jazrawi L, Rosen J: **Initial loop and knot security of arthroscopic knots using high-strength sutures**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2007, **23**(8):884-888.

231. Gallagher AG, Ritter EM, Lederman AB, McClusky DA, Smith CD: **Video-assisted surgery represents more than a loss of three-dimensional vision**. *The American journal of surgery* 2005, **189**(1):76-80.

232. **A framework for technology enhanced learning.** In. Edited by Health; 2011.

233. Chan KC, Burkhart SS, Thiagarajan P, Goh JC: **Optimization of stacked half-hitch knots for arthroscopic surgery**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2001, **17**(7):752-759.

234. Lo IK, Burkhart SS, Chan KC, Athanasiou K: **Arthroscopic knots: determining the optimal balance of loop security and knot security**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2004, **20**(5):489-502.

235. Burkhart SS, Wirth MA, Simonich M, Salem D, Lanctot D, Athanasiou K: **Knot security in simple sliding knots and its relationship to rotator cuff repair: How secure must the knot be?** *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2000, **16**(2):202-207.

236. Karam MD, Pedowitz RA, Natividad H, Murray J, Marsh JL: **Current and future use of surgical skills training laboratories in orthopaedic resident education: a national survey**. *J Bone Joint Surg Am* 2013, **95**(1):e4.

237. Pedowitz RA, Marsh LJ: **Motor Skills Training in Orthopaedic Surgery: A Paradigm Shift Toward a Simulation ‐ based Educational Curriculum**. *Journal of the American Academy of Orthopaedic Surgeons* 2012, **20**(7):407-409.

238. Thomas GW, Johns BD, Marsh JL, Anderson DD: **A review of the role of simulation in developing and assessing orthopaedic surgical skills**. *The Iowa orthopaedic journal* 2014, **34**:181.

239. Cuschieri A: **Whither minimal access surgery: tribulations and expectations**. *The American Journal of Surgery* 1995, **169**(1):9-19.

240. Krampe RT, Ericsson KA: **Maintaining excellence: deliberate practice and elite performance in young and older pianists**. *Journal of experimental psychology: general* 1996, **125**(4):331.

241. Cheung J, Rojas D, Weber B, Kapralos B, Carnahan H, Dubrowski A: **Evaluation of tensiometric assessment as a measure of skill degradation**. *Studies in health technology and informatics* 2011, **173**:97-101.
242. Ching SS, Mok CW, Koh YX, Tan S-M, Tan YK: **Assessment of surgical trainees' quality of knot-tying**. *Journal of surgical education* 2013, **70**(1):48-54.
243. Horeman T, Blikkendaal MD, Feng D, van Dijke A, Jansen F, Dankelman J, van den Dobbelsteen JJ: **Visual force feedback improves knot-tying security**. *Journal of surgical education* 2014, **71**(1):133-141.
244. Huang E, Chern H, O'Sullivan P, Cook B, McDonald E, Palmer B, Liu T, Kim E: **A better way to teach knot tying: a randomized controlled trial comparing the kinesthetic and traditional methods**. *The American Journal of Surgery* 2014, **208**(4):690-694.
245. Hu Y, Tiemann D, Brunt LM: **Video self-assessment of basic suturing and knot tying skills by novice trainees**. *Journal of surgical education* 2013, **70**(2):279-283.
246. Mashaud LB, Arain NA, Hogg DC, Scott DJ: **Development, validation, and implementation of a cost-effective intermediate-level proficiency-based knot-tying and suturing curriculum for surgery residents**. *Journal of surgical education* 2013, **70**(2):193-199.
247. Rodrigues S, Horeman T, Sam P, Dankelman J, van den Dobbelsteen J, Jansen FW: **Influence of visual force feedback on tissue handling in minimally invasive surgery**. *British Journal of Surgery* 2014, **101**(13):1766-1773.
248. Empel PJ, Verdam MG, Huirne JA, Bonjer HJ, Meijerink WJ, Scheele F: **Open knot‐tying skills: Resident skills assessed.** *Journal of Obstetrics and Gynaecology Research* 2013, **39**(5):1030-1036.

249. van Leeuwen N, Trimbos JB: **Strength of sliding knots in multifilament resorbable suture materials**. *Gynecological surgery* 2012, **9**(4):433-437.
250. Wong IH, Denkers MR, Urquhart NA, Farrokhyar F: **Systematic instruction of arthroscopic knot tying with the ArK Trainer: an objective evaluation tool**. *Knee Surgery, Sports Traumatology, Arthroscopy* 2015, **23**(3):912-918.
251. Wong IH, Denkers M, Urquhart N, Farrokhyar F: **Construct validity testing of the Arthroscopic Knot Trainer (ArK)**. *Knee Surgery, Sports Traumatology, Arthroscopy* 2015, **23**(3):906-911.
252. Eden J, Berwick D, Wilensky G: **Graduate medical education that meets the nation′ s health needs. Committee on Governance and Financing of Graduate Medical Education. Board on Health Care Services**. *Institute of Medicine of the National Academies, Washington, DC* 2014.
253. Pickersgill T: **The European working time directive for doctors in training**. *Br Med J* 2001, **323**(7324):1266.
254. Pellegrini C, Sachdeva A, Johnson K: **Accreditation of education institutes by the American College of Surgeons: a new program folowing an old tradition**. *Bulletin of the American College of Surgeons* 2006, **91**(3):8.
255. Petrera M, Dwyer T, Tsuji MR, Theodoropoulos JS: **Outcomes of arthroscopic Bankart repair in collision versus noncollision athletes**. *Orthopedics* 2013, **36**(5):e621-e626.
256. Harris JD, Gupta AK, Mall NA, Abrams GD, McCormick FM, Cole BJ, Bach BR, Romeo AA, Verma NN: **Long-term outcomes after Bankart shoulder stabilization**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2013, **29**(5):920-933.

257. Netto NA, Tamaoki MJS, Lenza M, dos Santos JBG, Matsumoto MH, Faloppa F, Belloti JC: **Treatment of Bankart lesions in traumatic anterior instability of the shoulder: a randomized controlled trial comparing arthroscopy and open techniques**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2012, **28**(7):900-908.
258. Chahal J, Marks PH, MacDonald PB, Shah PS, Theodoropoulos J, Ravi B, Whelan DB: **Anatomic Bankart repair compared with nonoperative treatment and/or arthroscopic lavage for first-time traumatic shoulder dislocation**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2012, **28**(4):565-575.
259. Stein T, Linke RD, Buckup J, Efe T, von Eisenhart-Rothe R, Hoffmann R, Jäger A, Welsch F: **Shoulder sport-specific impairments after arthroscopic Bankart repair: a prospective longitudinal assessment**. *The American journal of sports medicine* 2011, **39**(11):2404-2414.
260. Smith R: **All changed, changed utterly**. *British Medical Journal* 1998, **316**:1917-1917.
261. McClusky D, Ritter E, Lederman A, Gallagher A, Smith C: **Correlation between perceptual, visuo-spatial, and psychomotor aptitude to duration of training required to reach performance goals on the MIST-VR surgical simulator**. *The American Surgeon* 2005, **71**(1):13-21.
262. Gallagher A, Cowie R, Crothers I, Jordan-Black J-A, Satava R: **PicSOr: An objective test of perceptual skill that predicts laparoscopic technical skill in three initial studies of laparoscopopic performance**. *Surgical endoscopy* 2003, **17**(9):1468-1471.
263. Crothers I, Gallagher A, McClure N, James D, McGuigan J: **Experienced laparoscopic surgeons are automated to the "fulcrum effect": an ergonomic demonstration**. *Endoscopy* 1999, **31**(05):365-369.

264. Gallagher AG, Smith CD, Bowers SP, Seymour NE, Pearson A, McNatt S, Hananel D, Satava RM: **Psychomotor skills assessment in practicing surgeons experienced in performing advanced laparoscopic procedures**. *Journal of the American College of Surgeons* 2003, **197**(3):479-488.
265. [http://www.randomization.com]
266. Zendejas B, Brydges R, Hamstra SJ, Cook DA: **State of the evidence on simulation-based training for laparoscopic surgery: a systematic review**. *Annals of surgery* 2013, **257**(4):586-593.
267. Frank RM, Erickson B, Frank JM, Bush-Joseph CA, Bach BR, Cole BJ, Romeo AA, Provencher MT, Verma NN: **Utility of modern arthroscopic simulator training models**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2014, **30**(1):121-133.
268. Cannon WD, Garrett WE, Hunter RE, Sweeney HJ, Eckhoff DG, Nicandri GT, Hutchinson MR, Johnson DD, Bisson LJ, Bedi A: **Improving residency training in arthroscopic knee surgery with use of a virtual-reality simulator**. *J Bone Joint Surg Am* 2014, **96**(21):1798-1806.
269. Angelo RL: **Magellan and Copernicus: Arthroscopy Association of North America Seeking Excellence in Education**. *Arthroscopy* 2015, **31**(8):1428-1429.
270. Barber FA, Herbert MA, Beavis RC: **Cyclic load and failure behavior of arthroscopic knots and high strength sutures**. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2009, **25**(2):192-199.

Appendix 1, (Chapter 2.1)

Instructional Course 104

The Overhead Athlete: How to Examine, Test, and Treat Shoulder Injuries. Intra-articular Pathology

Richard L. Angelo, M.D.

ANTEROINFERIOR INSTABILITY

The overhead athlete typically places demands on his or her shoulder that far exceed activities for the normal population. Intra-articular pathology in the overhead athlete includes microinstability, SLAP tears, internal impingement, biceps tendinopathy, and partial articular surface rotator cuff tears. The fact that these pathologies are interrelated and often coexist creates a challenge in identifying specific etiologies. The pathomechanics involved in the overhand throwing motion and the internal impingement phenomenon in particular are complex. Treatment of these entities in the past has produced mixed results. As the results of valuable research accumulate, more unified models are evolving that begin to better explain the breadth of clinical findings. With a more complete understanding comes the hope of more effective treatment strategies.

MICROINSTABILITY

Progressive acquired capsular and ligamentous attenuation unrelated to specific traumatic events may create progressive dysfunction in the shoulder. Symptomatic microinstability can be anteroinferior, straight anterior, or anterosuperior. Although posterosuperior instability has been described, the pathology is not truly one of instability. Rather than being caused by capsular and ligamentous laxity, it is due to capsular tightness and aberrant posterosuperior glenohumeral translation. This is addressed later with internal impingement.

Repetitive overhead sports activities including throwing, volleyball, tennis, and gymnastics may create anteroinferior glenohumeral instability, which is the most common acquired symptomatic laxity. On examination, the "load and shift" test identifies excessive anteroinferior laxity but is difficult to quantitate. The anterior Jobe relocation test is a helpful sign. It is considered to be positive when the patient's apprehension created by the examiner placing the shoulder in abduction and external rotation is eliminated by a posteriorly directed force on the upper arm. Arthroscopic findings may include a positive "drive-through" sign in which the arthroscope can be passed through a generous glenohumeral interval from posterior to anterior across the shoulder joint. Depression or concavity of the labrum and glenoid chondromalacia are footprints of excessive translation of the humeral head on the glenoid. Occasionally, an incomplete Bankart lesion is present and may also create subtle anteroinferior instability.

Treatment options include an arthroscopic capsular plication that allows the surgeon, with some "guestimation," to roughly quantitate the magnitude of capsuloligamentous shortening that is produced. A rasp or whisker shaver blade is used to lightly excoriate the capsule along a 1.5-cm band adjacent to the labrum. A tuck or fold of capsule is then initiated by inserting a curved, cannulated suture hook into the capsule 1.0 to 1.2 cm lateral to the glenoid rim, passing it immediately deep to the capsule, and exiting approximately 5 mm lateral to the glenoid rim. The fold of capsule is then created by delivering the suture hook beneath the intact labrum to exit centrally. Long-term follow-up is not available for this technique, but early results are encouraging and the risk is relatively low because minimal tissue destruction is created.

Address correspondence to Richard L. Angelo, M.D., 12911 120th Ave NE, Suite H-210, Kirkland, WA 98034, U.S.A. E-mail: rlamdortho@aol.com

0749-8063/03/1910-0107$30.00/0
doi:10.1016/j.arthro.2003.09.046

A second option to stabilize the shoulder is to perform a thermal capsulorrhaphy. A great deal of debate has surrounded this technique in the past several years regarding its safety and efficacy. The limited clinical studies that are available show widely varying success rates. Levitz et al.[1] reported 85% success rate in 122 throwing athletes when thermal capsulorrhaphy was used as an adjunctive tool. Others have reported failure rates from 15% to 60% depending on the primary clinical pattern of instability.[2] The visible tissue response to the heat probe is quite variable and is an unreliable guide to the magnitude of the thermal effect on the tissues. Reports of permanent capsular damage have led to recommendations for "striping" or creating a grid pattern rather than painting the tissues. It is advisable to leave as much healthy untreated tissue as that which is thermally altered. Capsular necrosis, stiffness, and axillary nerve injury are concerns and this modality must be used with caution.[3]

The open capsulolabral reconstruction has been reported to permit return of up to 75% of professional baseball players for at least one full season subsequent to shoulder repair. A transverse incision rather than a vertical detachment of the subscapularis avoids much of the morbidity associated with an open procedure and permits earlier and more aggressive rehabilitation. A "pants over vest" imbrication of the capsule along the glenoid rim is created to reduce the capsular volume and restore stability to the shoulder.

STRAIGHT ANTERIOR INSTABILITY

Straight anterior glenohumeral instability is relatively uncommon and may result from tearing of the midlabrum and detatchment of the middle glenohumeral ligament origin. Associated partial articular surface rotator cuff tears are identified in approximately 2 thirds of patients. In addition to repetitive overhead activities, glenohumeral hyperextension at neutral rotation and 45° abduction may also produce this pathology. Examination findings include a positive load and shift test and a positive anterior Jobe relocation test. A positive Whipple test denotes pain on resisted elevation of the arm in the scapular plane if associated supraspinatus tearing is present. Treatment includes a suture anchor repair of the anterior labrum and associated middle glenohumeral ligament complex along with arthroscopic debridement of the articular surface rotator cuff tear.

ANTEROSUPERIOR INSTABILITY

Anterosuperior instability is also relatively rare. The eponym, SLAC (superior labrum anterior cuff) has been used to describe this lesion.[4] An anterosuperior labral lesion, a superior glenohumeral ligament tear, and a partial articular surface supraspinatus tear have been noted. Occasionally, chondromalacia in the superior glenoid quadrant is detected. Approximately 50% of the patients who have been recognized with this entity have been overhead athletes, and 50% have sustained significant shoulder trauma. The superior labral and glenohumeral ligament damage is either repaired or debrided along with the rotator cuff.

SLAP-BICEPS LESIONS

The superior labrum is typically more meniscoid in appearance than the inferior region. The biceps anchor has a variable attachment to the supraglenoid tubercle with approximately 25% to 50% attaching to the bony tubercle and 50% to 75% attaching predominantly to the posterosuperior labrum. Normal variants include an anterosuperior sublabral foramen and the "Buford complex" or cord-like middle glenohumeral ligament. Snyder[5] was the first to classify superior labral tears. Type I consists of superior labral fraying (20%); type II, biceps-labral detachment (55%); type III, a superior bucket-handle tear (9%); type IV, a bucket handle tear with extension into the biceps tendon (10%) and complex (5%). Microinstability, internal impingement (discussed later), marked external rotation of the abducted arm,[6] and traction on the long head tendon of the biceps during deceleration of the throwing arm are possible mechanisms of injury creating SLAP lesions in the overhead athlete.[7,8]

Many tests have been described to diagnose superior labral tears but lack specificity and sensitivity. The following are more reliable. The posterior Jobe relocation test (for posterior or superior SLAP lesions) is begun by placing the patient's arm in 90° abduction and full external rotation. If posterosuperior pain is relieved by the examiner creating a posteriorly directed force on the upper arm, the test is considered positive. The O'Brien test is performed by placing the patient's arm in 90° flexion, 25° adduction, and full internal rotation. Downward pressure on the arm by the examiner may create anterosuperior pain. The test is considered positive for an anterosuperior labral tear if the pain on resisted flexion is eliminated with the arm in a similar position but with the forearm supinated. Kibler's anterior slide test is begun by asking the patient to their place hands on their hips with the elbows directed posteriorly. With one hand, the examiner supports the scapula. The other hand creates an anterosuperiorly directed force on the patient's elbow. If anterosuperior pain is generated, an anterior or superior labral tear is suspected.

A challenge is often presented to the arthroscopist in deciding which superior labral tears are significantly pathologic and require treatment. A large recess between the superior labrum and glenoid may be a normal occurrence. The findings that suggest significant pathology include hemorrhage and irregularity at the biceps anchor, superior labral arching with biceps traction, biceps "peelback" with abduction and external rotation of the shoulder, and a positive "drive-through" sign seen arthroscopically. Treatment includes debridement for type I tears, suture anchor repair for type II tears, resection versus repair for type III lesions, and repair, debridement, or occasionally tenodesis for type IV tears.[9]

INTERNAL IMPINGEMENT

The constellation of pathology found with internal impingement includes posterosuperior SLAP tears, a partial articular surface tear of the posterior supraspinatus, and posterosuperior glenoid chondromalacia. Walch et al.[10] were among the first to describe this entity. Though contact between the greater tuberosity and posterosuperior glenoid was thought to occur normally in full abduction and external rotation, it was believed that the repetitive frequency and intensity with which it occurred, especially during throwing, led to labral and rotator cuff damage. It was also believed that decreased humeral retroversion may exacerbate the problem. Jobe[11] and Davidson et al.[12] attributed the pathologic findings of internal impingement to acquired anteroinferior microinstability that compromises the obligate posterior rollback of the humeral head during abduction and external rotation.[13] The anterior translation and lack of roll-back of the humeral head were believed to permit increased impact of the greater tuberosity on the posterosuperior glenoid. In addition, hyperangulation of the glenohumeral joint in the transverse plane was thought to increase the frequency and magnitude of the greater tuberosity-rotator cuff contact on the posterosuperior glenoid. Kibler[14] suggested that a loss of scapular synchrony with inefficient scapular elevation and retraction also created hyperangulation of the glenohumeral articulation. Eventually, the increased stress on the anterior capsuloligamentous structures was believed to create an acquired anteroinferior microinstability. Components of several of these models likely coexist in any one particular shoulder patient suffering from internal impingement.

Recently, Burkhart et al.[15] offered a model that unifies a number of these concepts used to explain internal glenohumeral impingement. A key finding thought to initiate the pathologic cascade is a glenohumeral internal rotation deficit (GIRD) due to a contracted posterior capsule. As the arm comes into abduction and external rotation during the throwing motion, the contracted posterior capsule "slings" beneath the humeral head. After the excursion in the posterior capsule reaches it limit, the humeral head then begins to "roll up" the capsule much like a tire on a rope and results in an aberrant posterosuperior shift of the humeral head. This shift creates shear forces that produce posterosuperior labral tearing and glenoid chondromalacia. Accompanying the posterosuperior shift in the axis of humeral head rotation is a pseudolaxity of the anterior capsule. As the humeral head translates posterosuperiorly, it no longer "fills" the anterior capsule and permits a capsular redundancy. Both this anterior pseudolaxity and the increased clearance between the greater tuberosity and the posterosuperior glenoid, caused by the posterior humeral head shift, permit hyperexternal rotation of the shoulder. As the humeral head externally rotates excessively, there are 2 consequences. First, along with excessive torsion of the biceps tendon, the vector of the tendon becomes more posteriorly directed than normal and leads to a "peelback" of the posterior and superior labrum.[16,17] Second, excessive torsion of the rotator cuff may contribute to tearing of the articular surface fibers. Finally, a "break in the ring" of the posterosuperior labrum (circle concept) is thought to add to the anterior capsular pseudolaxity that may be manifested as a positive "drive-through" sign during arthroscopy.

Examination findings of significance include a loss of internal rotation greater than 25° and a positive posterior Jobe relocation test. Excessive anteroinferior glenohumeral translation may be detected but is often difficult to quantitate.

The initial treatment is directed toward activity modification, nonsteroidal anti-inflammatory medication (NSAIDs), focused posterior capsule stretching, and rotator cuff and periscapular strengthening. If posterior capsular stretching is unsuccessful, a limited arthroscopic posterior capsular release may be indicated in a small number of patients.[18] If there appears to be significant anterior inferior microinstability, consideration may need to be given for an arthroscopic capsular plication, thermal capsulorrhaphy, or open anterior capsule labral reconstruction.[19] When a posterior or superior SLAP II lesion is present, it should be repaired and will usually eliminate the drive-through sign seen at diagnostic arthroscopy. Partial articular surface cuff tears are debrided to stable healthy tissue.

PARTIAL ARTICULAR SURFACE ROTATOR CUFF TEARS

The etiology for partial articular surface rotator cuff tears is likely multifactorial but may include repetitive traction on articular surface fibers such as during deceleration of the throwing arm[20] and internal impingement as described above. Biomechanically, the articular surface of the cuff may be more likely to fail under tensile rather than compressive forces.[21] A grade 1 tear describes a defect less than 3 mm (< 25%); grade 2, 3 to 6 mm defect thickness (< 50%); and grade 3, greater than 6 mm defect thickness (> 50%). Treatment includes an arthroscopic debridement to stable healthy rotator cuff tissue. For the few that are grade 3 tears, consideration may need to be given for an arthroscopic or mini-open rotator cuff repair. An arthroscopic subacromial decompression may be considered part of the management for grade 1 and 2 articular surface tears as suggested by Payne et al.[22] Great caution should be exercised if instability is a component of the pathology because a subacromial decompression may worsen symptoms.[23]

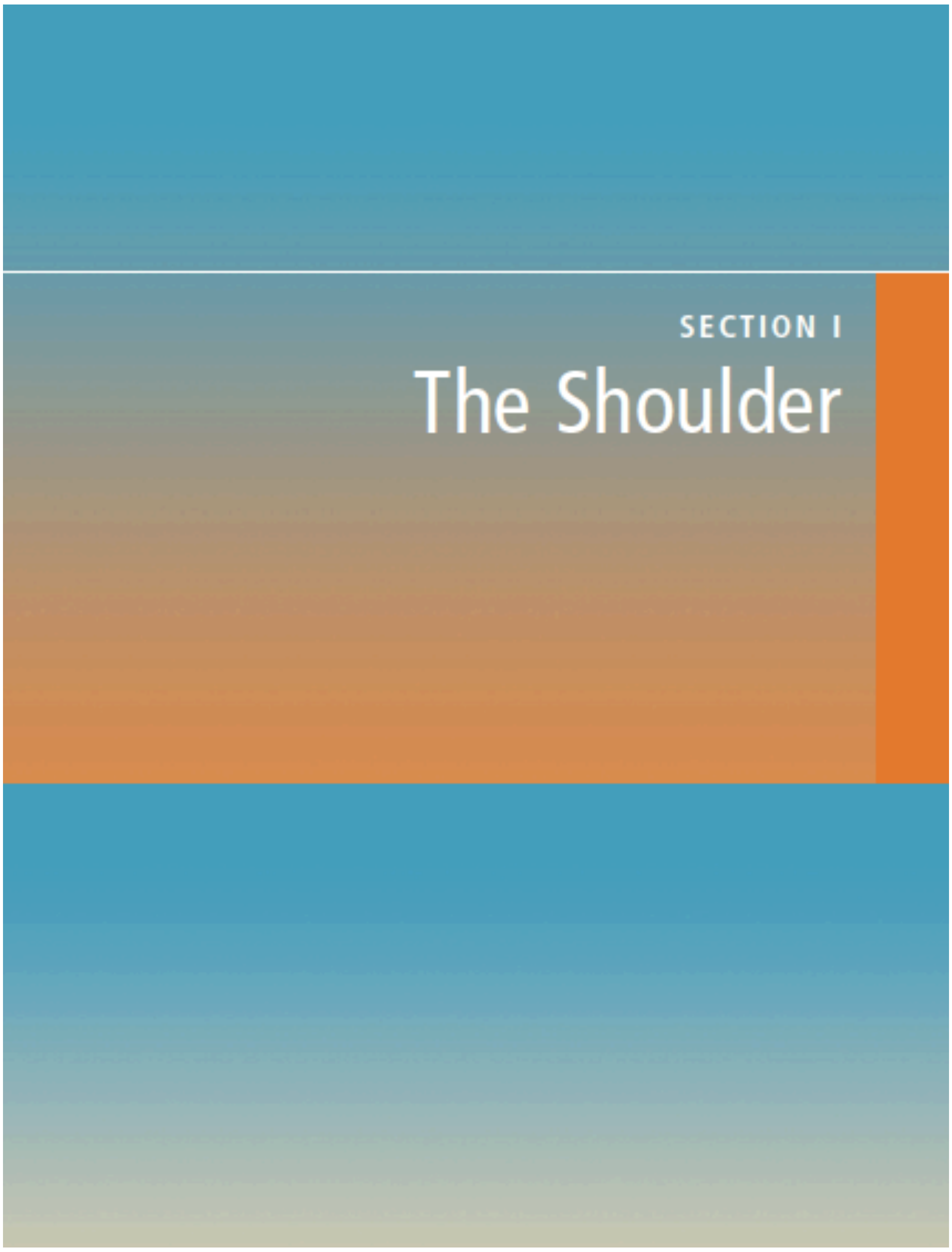
SUMMARY

The pathology in the overhead athlete's shoulder is often complex, with substantial overlap between microinstability, labral pathology, internal impingement, and partial articular surface rotator cuff tears. An accurate diagnosis demands the careful integration of the history, physical examination findings, imaging studies, examination under anesthesia, and findings at diagnostic arthroscopy. The treatment options described have relatively little intermediate or long-term follow-up and remain controversial.

REFERENCES

1. Levitz CL, Dugas J, Andrews JR. The use of arthroscopic thermal capsulorrhaphy to treat internal impingement in baseball players. *Arthroscopy* 2001;17:573-577.
2. Gerber A, Warner JP. Thermal capsulorrhaphy to treat shoulder instability. *Clin Orthop* 2002;400:105-116.
3. Abrams JS. Thermal capsulorrhaphy for instability of the shoulder: Concerns and applications of the heat probe. *Instr Course Lect* 2001;50:29-36.
4. Savioe FH III, Field LD, Atchinson S. Anterior superior instability with rotator cuff tearing: SLAC lesion. *Orthop Clin North Am* 2001;32:457-461.
5. Snyder SJ, Karzel RP, Del Pizzo W, et al. SLAP lesions of the shoulder. *Arthroscopy* 1990;6:274-279.
6. Kuhn JE, Lindholm SR, Huston LJ, et al. Ligamentous restraints and muscle effects limiting external rotation of the glenohumeral joint in the neutral and abducted positions. *J Shoulder Elbow Surg* 1999;8:664.
7. Fleisig GS, Andrews JR, Dillman CD, et al. Kinetics of baseball pitching with implications about injury mechanisms. *Am J Sports Med* 1995;23:233-239.
8. Andrews JR, Carson WG, McLeod WD. Glenoid labrum tears related to the long head of the biceps. *Am J Sports Med* 1985;13:337-341.
9. Mileski RA, Snyder SJ. Superior labral lesions in the shoulder: Pathoanatomy and surgical management. *J Am Acad Orthop Surg* 1998;6:121-131.
10. Walch G, Boileau J, Noel E, et al. Impingement of the deep surface of the supraspinatus tendon on the posterior superior glenoid rim: An arthroscopic study. *J Shoulder Elbow Surg* 1992;1:238-243.
11. Jobe CM. Superior glenoid impingement, expanded spectrum. *Arthroscopy* 1995;11:530-537.
12. Davidson PA, El Attrache NW, Jobe CM, Jobe FW. Rotator cuff and posterior-superior glenoid labrum injury associated with increased glenohumeral motion: A new site of impingement. *J Shoulder Elbow Surg* 1995;4:384-390.
13. Harryman DT II, Sidles JA, Clark JM, et al. Translation of the humeral head on the glenoid with passive glenohumeral motion. *J Bone Joint Surg Am* 1990;72:1334-1343.
14. Kibler WB. The role of the scapula in athletic shoulder function. *Am J Sports Med* 1998;26:325-337.
15. Burkhart SS, Morgan CD, Kibler WB. The disabled throwing shoulder: Spectrum of pathology Part I: Pathoanatomy and biomechanics. *Arthoscopy* 2003;19:404-420.
16. Burkhart SS, Morgan CG. The peel-back mechanism: Its role in producing and extending posterior Type II SLAP lesions and its effect on SLAP repair rehabilitation. *Arthroscopy* 1998;14:637-640.
17. Morgan CD, Burkhart SS, Palmeri M, Gillespie M. Type II SLAP lesions: Three subtypes and their relationships to superior instability and rotator cuff tears. *Arthroscopy* 1998;14:553-565.
18. Ticker JB, Beim GM, Warner JP. Recognition and treatment of refractory posterior capsular contracture of the shoulder. *Arthroscopy* 2000;16:27-34.
19. Jobe FW, Giangarra CE, Kvitne RS, et al. Anterior capsulolabral reconstruction of the shoulder in athletes in overhead sports. *Am J Sports Med* 1991;19:428-434.
20. Andrews JR, Angelo RL. Shoulder arthroscopy for the throwing athlete. *Tech Orthop* 1988;3:75-81.
21. Nakajima T, Rokuuma N, Hamada K, et al. Histologic and biomechanical characteristics of the supraspinatus tendon: Reference to rotator cuff tearing. *J Shoulder Elbow Surg* 1994;3:79-87.
22. Payne LZ, Altchek DW, Craig EV, Warren RF. Arthroscopic treatment of partial rotator cuff tears in young athletes. *Am J Sports Med* 1997;25:299-305.
23. Jobe FW, Kvitne RS. Shoulder pain in the overhand or throwing athlete. *Orthop Rev* 1989;18:963-975.

Appendix 2, (Chapter 2.2)



Arthroscopic Setup: Approaches and Tips for Success

Richard L. Angelo

PATIENT POSITIONING

The safety and ease with which an arthroscopic shoulder procedure is accomplished frequently relates to how the patient is positioned and how accurate and utilitarian are the portals that have been established. Although minor variations exist, most surgeons employ either the lateral decubitus or the beach chair positions, and each has its proponents. The choice is largely influenced by the familiarity gained while the surgeon was learning shoulder arthroscopy, the ease and anticipated likelihood of converting to a mini-open procedure, and the availability of surgical assistants and supportive devices for arm positioning. Equipment is readily available to facilitate the use of either position.

Lateral Decubitus Orientation

The supine position is used during the induction of general anesthesia. The patient is then repositioned in the lateral decubitus orientation on a vacuum bag (Fig. 1.1A, B). A gel pad can be layered on top of the bean bag, particularly if there is the anticipation that the procedure may be prolonged. A soft axillary roll is placed beneath the upper thorax to minimize direct pressure on the axilla, and the head is supported in a neutral orientation. The patient is allowed to roll back approximately 15 degrees orienting the glenoid parallel with the floor. The vacuum bag is then evacuated to maintain support. All bony prominences must be appropriately padded, in particular the fibular head to protect the peroneal nerve. The operating table is then rotated to position the anesthesiologist and necessary equipment in an area near the middle of the operating table near the patient's abdomen. The surgeon is thus provided with unrestricted access to the involved shoulder. Monitors are located for easy viewing.

If the primary procedures are to be performed in the subacromial region, that is, rotator cuff repair, the primary monitor is positioned superior and anterior to the patient's head. A secondary monitor for use by the surgical assistant may be located in front of and above the patient's abdomen. When the work to be completed is primarily in the glenohumeral joint, that is, a Bankart or SLAP repair, the monitor is set across from the surgeon near the patient's abdomen as the general viewing direction for glenohumeral procedures is anterior. The arm is supported in 30 to 40 degrees of abduction and 15 degrees of forward flexion using 10 lb (4.5 kg) to suspend rather than place significant traction on the arm. This shoulder position is varied during the case depending on the access necessary to specific locations. Numerous sterile sleeves and gauntlet devices are commercially available to support the arm.

Arthroscopic Bankart repairs may be facilitated by directing 10 lb (4.5 kg) of accessory traction laterally (perpendicular to the humerus) to distract the shoulder and improve access to the anterior aspect of the glenohumeral joint. Alternatively, a similar manual maneuver can be accomplished by an assistant. A routine sterile prep and drape are then performed. The lateral decubitus method eliminates the need for an assistant or mechanical device to support the arm. Internal and external rotation of the suspended arm affords acceptable access to the entire rotator cuff. If range of motion is to be assessed, that is, at the completion of a Bankart repair, the arm is removed from suspension for the motion exam while maintaining sterility of the sleeve suspension loop.

While working in the glenohumeral joint, the monitor view of the glenoid is typically oriented parallel with the floor. When working in the subacromial space, however, the surgeon may elect to either maintain this orientation (the acromion is vertical) or rotate the camera head to view the acromion in a position parallel with the floor (as it would appear with the patient standing).

If converting to an open procedure through a standard deltopectoral approach for the glenohumeral joint, subscapularis, or biceps tendon, the unsterile portion of the suspension apparatus is removed and the patient's arm is allowed to rest on the ipsilateral hip. The vacuum bag is at least partially inflated (softened) and the patient allowed to roll back into a more supine position. A draw sheet is used to center the patient on the operating table. The table is then configured to a gentle beach chair orientation with acceptable position and support for the head and neck

FIGURE 1.1. A: Patient positioned in the lateral decubitus orientation; anesthesia setup is near the chest. Dual monitors are helpful, particularly for the viewing of an assistant. B: Once draping is complete, easy access to the entire shoulder is afforded; the arm is "suspended" with 10 lb through a disposable arm sleeve.

verified. Although it is unnecessary to completely reprep and redrape, it is prudent to replace the clean, sterile barrier sheet anterior to the shoulder to shield the anesthesiologist and related equipment.

If the surgeon elects to convert to a mini-open approach to the subacromial region, repositioning is unnecessary although some prefer to tilt the table posteriorly toward the surgeon to improve access to the anterior shoulder. An approach to the supraspinatus and infraspinatus is readily obtained by extending the lateral subacromial (LSA) portal proximally. An absorbable suture is introduced transversely through the deltoid at the inferior extent of the portal defect to prevent inadvertent distal extension and iatrogenic injury to the axillary nerve. The deltoid is then divided proximally along its fibers to the level of the acromion.

Beach Chair Orientation

Some surgeons prefer the beach chair position due to its more anatomic orientation, which conforms to the familiar open approach. (1) The patient's thorax is positioned to permit the involved shoulder to overhang the side of the table. Once the hips are flexed 70° to 80° and the legs 30°, the back is then elevated approximately 70°. After padding bony prominences, a vacuum pack supports the hips and thorax, but is displaced from the ipsilateral periscapular region.

Alternatively, a specially designed table with a removable wing for exposure of the operative shoulder may be employed (Fig. 1.2A, B). A relatively more vertical orientation for the back will minimize the dependent position of the camera when the scope is in the posterior portal and also minimize lens fogging. However, a more upright position for the thorax increases the hydrostatic pressure gradient between the head and the brachium. The anesthesiologist sets up near the patient's uninvolved shoulder, and the viewing monitor is placed opposite the surgeon near the foot of the table. A surgical assistant or a sterile, maneuverable mechanical arm holder adjusts the position of the shoulder during the procedure, depending on the access necessary. Somewhat greater mobility of the arm exists when compared with the lateral decubitus position.

FIGURE 1.2. A: Patient positioned in the beach chair orientation; anesthesia setup is near the contralateral shoulder; a table with a removable wing affords easy access to the entire shoulder. B: The anterior and posterior aspects of the shoulder are readily accessed; a sterile arm positioner can be added if desired.

The upright (anatomic) orientation for the arthroscope and monitor view is maintained while working in both the glenohumeral and the subacromial regions. Conversion to an open procedure for all regions of the shoulder is relatively simple and only requires reducing the degree of thorax elevation. The vacuum pack must be at least partially inflated in order to safely change the patient's position without creating pressure points. Alternatively, a relatively more supine position for the thorax can often be accomplished by tilting the entire table into greater Trendelenberg.

A recent case report identified four patients who underwent shoulder surgery in the sitting position, which resulted in one death and three patients with severe brain damage. (2) Cerebral hypoperfusion, rather than cardiovascular risk factors, was believed to be the cause and may be attributable to differences in blood pressure reference points. A blood pressure difference as great as 90 mm Hg between the head and the calf may exist in the sitting position based on hydrostatic factors alone. Potentially catastrophic cerebral hypoperfusion may be avoided by precautions including placing the blood pressure cuff on the brachium rather than the calf, (3) maintaining perioperative blood pressure values at a minimum of 80% of preoperative resting values, and ensuring that the intraoperative blood pressure is at a minimum of 100 mm Hg at the level of the head. Loss of vision and ophthalmoplegia have also been reported following general anesthesia with the patient in the beach chair position, but the exact mechanisms for this pathology are unclear. (3) Thromboembolic events are also possible with the patient in the sitting position and make the use of cyclical pneumatic compression cuffs around the calves prudent.

ANESTHESIA CHOICES

General Anesthesia

Both endotracheal intubation and a laryngeal mask airway (LMA) provide safe, reliable options for maintaining the airway during general anesthesia. No durable analgesia is afforded once the patient awakens, and nausea/vomiting can sometimes be difficult to manage in the perioperative period.

Interscalene Regional Block

Interscalene blocks (ISBs) provide anesthesia, muscle relaxation, and postoperative analgesia although supportive parenteral pain medication may be necessary during the immediate postoperative period. (4) An ISB can be used as the primary means of anesthesia or as an adjunct to general anesthesia. As with any invasive procedure, the risk/benefit ratio determines its use. Proponents note its effectiveness despite the frequent need for some additional narcotic support during the immediate postoperative period and its relatively low risk of serious complications. Dedicated anesthesia teams committed to regional anesthesia and that perform a large number of blocks will help to minimize untoward events. (5) Potential serious complications have been reported including cardiac arrest, grand mal seizures, hematoma, and pneumothorax. Possible neurologic injuries include damage to the recurrent laryngeal, vagal, and axillary nerves. Phrenic nerve dysfunction is common and can give rise to significant respiratory distress. Brachial plexus pathology may include transient paresthesias (which have been reported to be as high as 9% at 24 hours and 3% at 2 weeks post-op), (6) or a brachial plexus palsy, which may be transient, require prolonged recovery or be permanent in a very small number of cases.

It is essential that the block be performed in the awake patient who is able to provide critical feedback during administration of the block. More recently, the use of ultrasound to guide placement of the needles has added a measure of safety. Even with a successful block, the duration of pain relief averages only 9 to 10 hours following surgery, which may make pain management challenging in an outpatient setting. (4) A thorough disclosure of the potential risks should be discussed with the patient, preferably beforehand in an office setting during the preoperative visit.

Adjunctive Pain Management

The suprascapular nerve supplies 70% of the sensation to the shoulder joint. Instillation of 20 cc of 0.25% bupivacaine adjacent to the suprascapular nerve may result in up to a 30% reduction in postoperative narcotic usage and a 5-fold reduction of nausea. (7,8) This block carries a low risk when performed with a blunt-tipped needle, and may be repeated as necessary, even in an office setting on the first postoperative day. (9) In addition, local infiltration of the portal sites with 0.5% bupivacaine leads to further reduction in pain. Pain pumps remain controversial, but have been consistently used in the subacromial space with safety provided that the glenohumeral joint is not exposed to the catheter and infiltrate. Cooling jackets using circulating ice water may also substantially improve patient comfort.

PORTALS

When arthroscopic portals are properly placed, they will provide the necessary field of view and instrument access to desired locations within the glenohumeral joint, acromioclavicular joint, and subacromial space. (10–14) A thorough knowledge of the regional anatomy, particularly the palpable bony landmarks, will improve safety and ensure accuracy in establishing the desired portals. There is a greater margin of safety in creating access to the subacromial space where the use of various accessory portals is routine.

General Technique

Bony landmarks are identified by careful palpation and mapped at the beginning of the case prior to soft tissue distortion from fluid extravasation. Anticipated portal sites are referenced from the landmarks and identified using a surgical marker. All anatomical references and diagrams provided here are for a right shoulder with the patient in the lateral decubitus position. Minor adjustments to the recommended distances from anatomic landmarks may be necessary if the patient is supported in the beach chair orientation or for particularly large or small patients. As experience is gained, surgeon preference may also lead to subtle adjustments in the skin entry site for various portals. The posterior glenohumeral portal is typically established first. It is recommended that all subsequent portals be made in an outside-in manner under direct vision after first establishing the desired tract with a spinal needle. A small skin incision is made at the chosen entry site and a trocar and cannula directed along the path identical to the spinal needle and into the glenohumeral joint or subacromial space.

FIGURE 1.3. Right shoulder in the lateral decubitus orientation viewed from superior (anterior is to the left); bony landmarks are mapped out and the common glenohumeral portals are identified; P, posterior; A, anterior; PI, posteroinferior; PL, posterolateral ("Port of Wilmington"); AS, anterosuperior; AL, anterolateral; MA, midanterior; LC, lateral coracoid.

Glenohumeral Portals (Fig. 1.3)

Posterior (P) serves as the primary intra-articular viewing portal and provides instrument access to the posterior glenoid labrum and rim, posterior capsule, and articular surface of the infraspinatus. The field of view includes the glenoid, posterosuperior humeral head, anterior capsule, biceps, superior subscapularis, glenohumeral ligaments, and articular surface of the supraspinatus and superior subscapularis tendons (Fig. 1.4A, B). The entry site is 1.0 to 1.5 cm inferior and 1.0 cm medial to the posterolateral (PL) corner of the acromion.

After creating a small skin incision, the cannula is introduced and directed toward the coracoid tip. If it is anticipated that this portal will be employed to drill or insert anchors along the posterior glenoid rim, the entry site must be adjusted 1 cm further lateral to account for the anterior glenoid version. This modification will enable the approach to be approximately 45° to the glenoid in the transverse plane. If this lateral modification is not made, the portal will be too "shallow" and create a risk that instruments will either skive off the articular cartilage or be directed too far medial along the glenoid neck.

Anterior (A) enters through the middle of the rotator interval and provides instrument access to the biceps, anterior labrum, glenoid rim, anterior and superior capsule, articular surfaces of the supraspinatus, infraspinatus, and the superior aspect of the subscapularis tendons. The field of view includes the posterior glenoid and labrum,

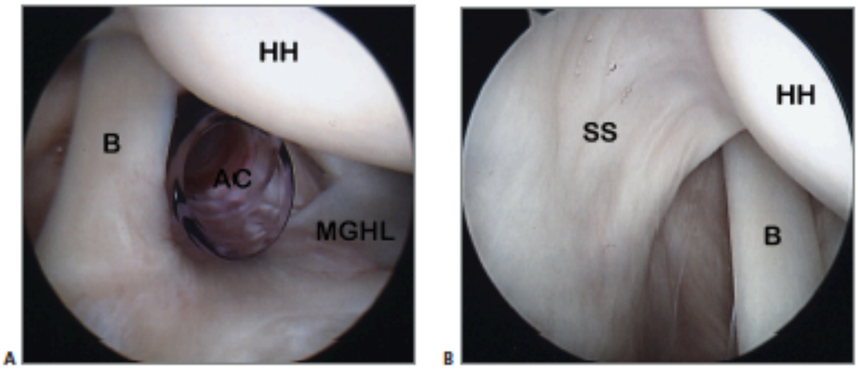


FIGURE 1.4. **A:** All arthroscopic photos are of a right shoulder with the patient in the lateral decubitus position; scope is in the posterior portal viewing anteriorly; HH, humeral head; B, biceps; MGHL, middle glenohumeral ligament; AC, anterior cannula. **B:** Scope is in the posterior portal viewing anteriorly; HH, humeral head; B, biceps; SS, capsule overlying the articular surface of the supraspinatus just posterior to the biceps.

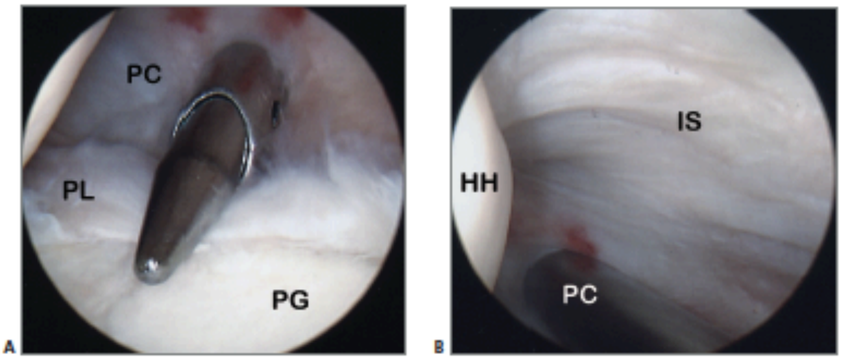


FIGURE 1.5. A: Scope is in the anterior portal viewing posteriorly; PC, posterior capsule; PL, posterior labrum; PG, posterior glenoid. B: Scope is in the anterior portal viewing posterosuperiorly; HH, humeral head; IS, capsule underlying the infraspinatus tendon; PC, posterior cannula.

anterosuperior (AS) humeral head, articular surface of the infraspinatus, posterior capsule, and the biceps origin (Fig. 1.5A, B). The entry site is midway between the coracoid tip and the anterolateral (AL) corner of the acromion. The cannula is directed toward the center of the glenohumeral joint while viewing from the posterior portal.

Midanterior (MA) is the preferred portal to instrument the anterior glenoid rim with drills and anchors in preparing the neck for a Bankart repair. In addition, it affords access to the anterior and inferior capsule for suture-passing instruments. The entry site is 1.5 cm lateral and 1.5 cm inferior to the coracoid tip. A spinal needle identifies the appropriate track, which, after penetrating the skin, is directed somewhat superiorly over the superior boarder of the subscapularis. A small superficial skin incision is made, and an obturator and cannula are initially directed superiorly, then over the top of the subscapularis, and finally inferiorly to enable ready access to the inferior glenoid. Instruments passing through this portal should be able to approach the glenoid at a 45° angle in the transverse plane.

AS provides a tangential view to the anterior glenoid rim and neck (for Bankart repairs), the superior insertion of the subscapularis onto the lesser tuberosity, the superior and posterior capsule, labrum, and glenoid rim (Fig. 1.6A, B). The entry site is 1.0 cm directly lateral to

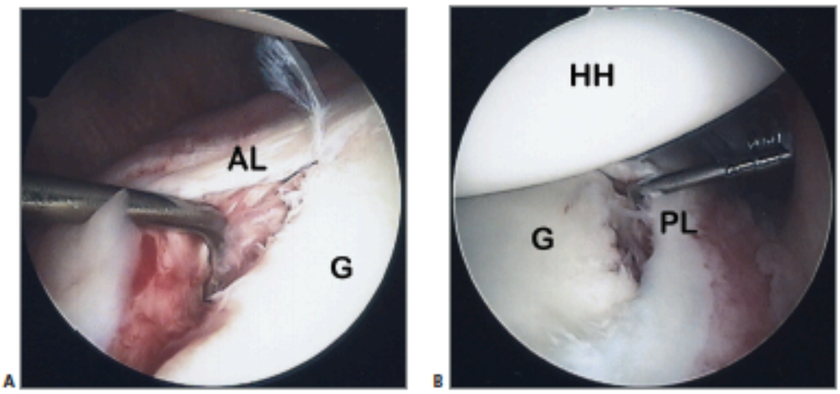


FIGURE 1.6. A: Scope is in the AS portal viewing anteroinferiorly; probe is demonstrating a Bankart lesion; G, glenoid; AL, anterior labrum. B: Scope is in the AS portal viewing posteroinferiorly; probe is inside a posterior labral tear; HH, humeral head; G, glenoid; PL, posterior labrum.

the AL corner of the acromion, and the cannula is directed immediately anterior to the anterior boarder of the supraspinatus and then either anterior or posterior to the biceps tendon, depending on the intended primary use.

AL serves to enable instrument access to the posterior aspect of the coracoid, the anterior, superior, and posterior aspects of the subscapularis for release, and to the lateral boarder of the subscapularis (e.g., for use with antegrade suture-passing instruments). The entry site is 1.0 cm anterior and 1.0 to 1.5 cm lateral to the AL corner of the acromion. The cannula or instrument is directed toward the posterior aspect of the tip of the coracoid or somewhat more inferiorly toward the biceps groove.

Lateral coracoid (LC) enables instrument access to the lesser tuberosity for subscapularis repair from an intra-articular view. The entry site is 1.0 to 1.5 cm directly lateral to the middle of the coracoid tip and the instrument is then directed somewhat laterally toward the lesser tuberosity.

PL (or Port of Wilmington) facilitates placement of anchors at the posterosuperior glenoid rim for labral repair. The portal may penetrate the infraspinatus tendon. Concern has been raised regarding the defect in the tendinous portion of the rotator cuff and it is advisable to limit this portal to the smallest diameter practical for a given anchor and its preparation. The entry site is 1.5 cm anterior and 1.5 cm lateral to the PL corner of the acromion. Viewing from an anterior portal, a spinal needle is directed approximately 45° from lateral to medial to establish the proper track.

Posteroinferior (PI) provides instrument access to the posterior capsule and axillary recess for capsular excoriation and suture plication. The entry site is 2.0 cm inferior and 1 cm lateral to the posterior portal. A spinal needle is used to establish the proper track while viewing from the AS portal. Care must be taken not to err too far inferior and risk injury to the axillary nerve.

Subacromial Portals (Fig. 1.7)

Posterior subacromial (PSA) is a primary viewing portal and offers instrument access to the posterior bursa, cuff, the acromion, and the greater tuberosity. The field of view includes the entire subacromial space, acromioclavicular joint, extra-articular biceps and sheath, the coracoclavicular ligaments, and suprascapular notch (Fig. 1.8A, B). The entry site is the same as the posterior glenohumeral portal. The trocar is directed anterosuperiorly, immediately inferior to the inferior surface of the acromion.

LSA provides a "50-yd line" view of the supraspinatus–infraspinatus insertion onto the greater tuberosity and a lateral view of the acromioclavicular joint, the anterior

FIGURE 1.7. Right shoulder in the lateral decubitus orientation viewed from superior (anterior is to the left); bony landmarks are mapped out and the common glenohumeral portals are identified; PSA, posterior subacromial; PLSA, posterolateral subacromial; LSA, lateral subacromial; LA, lateral acromial; ALSA, anterolateral subacromial; ASA, anterior subacromial; NV, Neviaser portal.

A B

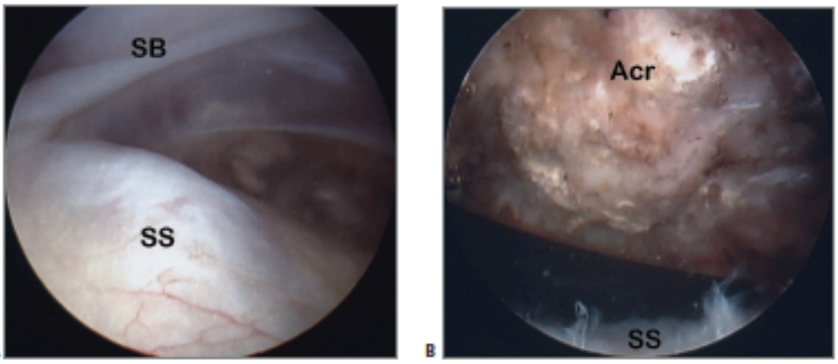


FIGURE 1.8. **A:** Scope is in the PSA portal viewing anteriorly; normal subacromial bursal region; SS, normal supraspinatus with vascular pattern; SB, ASA bursal fold. **B:** Scope is in the PSA portal viewing anteriorly; SS, bursal surface of supraspinatus; Acr, large anterior acromial spur.

acromion, and the posterior bursal curtain. Instruments are able to approach the rotator cuff, greater tuberosity, and acromion. The entry site is 2.5 to 3.0 cm lateral and 0 to 1.0 cm posterior to the AL corner of the acromion. Instruments roughly parallel the inferior surface of the acromion.

Anterolateral subacromial (ALSA) portal is the same as AL glenohumeral portal, but is placed into the subacromial space. When in the anterior subacromial (ASA) space, it provides a view of the extra-articular biceps, the intertubercular groove, the bursal surface of the subscapularis, and the lesser tuberosity (once the clavipectoral fascia has been excised). Instruments can approach the subscapularis tendon for release and suture passage as well as to perform a coracoplasty. The entry site is 1.0 cm anterior and 1.0 to 1.5 cm lateral to the AL corner of the acromion.

ASA is the same as the anterior glenohumeral portal, but enters the subacromial space. It offers a view of most of the subacromial space, but is commonly used for suture management. Instruments can be introduced into the anterior aspect of the rotator cuff for a side-to-side repair. Once through the skin, the trocar is directed immediately beneath the anterior margin of the acromion. When instrument access to the biceps groove is intended, the optimal portal entry site is identified with a spinal needle. While viewing from the AL portal, the needle is directed toward the biceps groove with the humerus internally rotated approximately 20°.

Posterolateral subacromial (PLSA) serves as a primary viewing portal to address rotator cuff pathology. Once established, a 30° scope offers a "50-yd line" view of the rotator cuff and subacromial space (Fig. 1.9). The entry site is approximately 1.0 cm anterior and 1.0 cm lateral to the PL corner of the acromion. An arthroscope in the PL portal may interfere with instruments introduced through the LSA portal if a minimum of 3 cm is not maintained between the two sites.

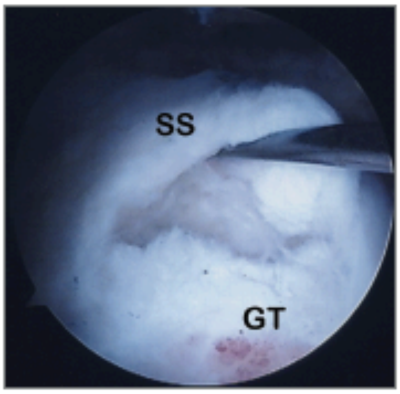


FIGURE 1.9. Scope is in the PLSA portal viewing anteromedially; probe demonstrates a bursal-sided rotator cuff tear; SS, supraspinatus; GT, greater tuberosity.

Lateral acromial (LA) is primarily used for instrument approach to the greater tuberosity (e.g., drill, tap, and anchor insertion for rotator cuff repair). The entry site is immediately lateral to the lateral border of the acromion. The optimal anteroposterior location is identified using a spinal needle. Access to the entire greater tuberosity is possible with internal and external rotation of the humerus. When attempting to place anchors into the medial aspect of the greater tuberosity adjacent to the articular cartilage, it is essential to nearly completely adduct the humerus to avoid approaching the tuberosity at too shallow an angle and potentially violating the articular surface of the humeral head.

Superomedial (SM—Neviaser) is employed to introduce suture-passing and retrieving instruments toward the rotator cuff. The entry site is 1.0 cm medial to the posterior aspect of the acromioclavicular joint. While the arthroscope in the subacromial space and the arm abducted <45°, a spinal needle is directed from medial to lateral at approximately 60° in the frontal plane. If the portal is introduced too close to the acromioclavicular joint, the mobility of the instrument is significantly restricted.

Anterior acromioclavicular (AAC) affords an anterior approach for resection of the distal clavicle. The entry site is 2.0 cm anteroinferior to and in line with the acromioclavicular joint. The optimal path is identified with a spinal needle. Alternatively, when approaching the acromioclavicular joint in direct fashion, two small portals can be established. One is directly AS and a second posterosuperior to the AC joint. A small-diameter arthroscope and shaver are used initially until a greater space can be established.

SUTURE MANAGEMENT

Suture management is one of the most challenging aspects of accurately completing an effective arthroscopic shoulder procedure. By employing a systematic routine, suture can be passed, manipulated, and tied in an efficient manner. Simplifying the steps involved results in time saved and frustration avoided. Suture must be handled carefully to avoid fraying and nicking with the possibility of eventual breakage. Loop rather than jaw-type graspers help maintain this suture integrity. It is optimal to isolate the suture being manipulated whenever possible by placing all other nonworking sutures in a separate portal. Tangling and mistaking various limbs and suture mates can thus be avoided. Once all sutures for a given anchor have been passed, the working cannula is withdrawn and then reinserted placing the sutures outside the cannula, which can then be used to manage a new set of sutures.

In order for sutures to securely reapproximate tissue, they must be optimally placed. When manipulating tissues and suture-passing instruments, efficiency can be gained by having an assistant hold the arthroscope to maintain an acceptable field of view. The surgeon is then able to secure the tissue with graspers in one hand while controlling the suture-passing device with the other, similar to using forceps and a needle driver in an open technique. Antegrade devices, which often simplify suture passage by minimizing the number of steps involved, can be made more efficient by using a counterforce traction suture to control the tissue and prevent it from being pushed away during instrument delivery. When using a penetrating device in retrograde fashion, its mobility can be restricted significantly once it has passed through the tissue. Rather than attempt to "chase" the desired suture with the open jaw, deliver the selected anchor suture to the penetrator with a loop grasper or knot pusher. Various cannulated instruments, with or without an attached suture retrieval loop, do not require the use of a cannula and are able to be introduced through a very small skin nick such as the SM (Neviaser) portal.

Entanglements are avoided by managing sutures in an orderly fashion. When passing sutures through the rotator cuff, it is helpful to pass them from "far to near." Those sutures that are to be passed furthest from the viewing arthroscope are introduced first (Fig. 1.10). Consequently, as subsequent sutures are delivered closer to the arthroscope, the field of view remains unobstructed by previously passed sutures. The suture pairs are then progressively tied in the opposite sequence, that is, those closest to the scope tied first and those furthest tied last. This method permits adherence to the principle of working with sutures in isolation as much as possible.

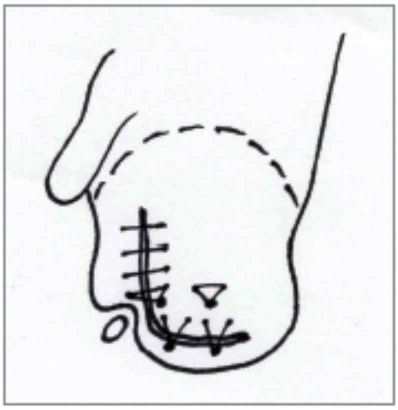


FIGURE 1.10. Diagram of a right shoulder viewed from superiorly depicting a large "L" shaped rotator cuff tear; consider placing most medial (farthest from scope) sutures first and working progressively laterally (closest to scope); consider tying most lateral sutures first and then progressing medial with suture pairs.

When working with sutures that pass through anchors, care must be taken to avoid "offloading" the suture from the anchor. The location of the involved anchor should be kept in view while a limb is being retrieved to verify that the suture is being pulled through the anchor. Stop, reorient yourself, and select the proper limb if the suture is moving through the anchor.

Many methods exist for tying knots. When using a sliding knot, the post limb must pass through the tissue being repaired so that the knot is delivered toward the tissue and away from the anchor. Otherwise, as the knot is introduced, it can become "bound up" at the entry site for the anchor and fail to slide further, compromising loop security. In addition, prominent knots near articular surfaces may generate significant chondral scuffing and abrasion. As half hitches are introduced to secure the knot, the post should be alternated, the throw reversed, and each half hitch seen to "lay down" without inappropriate twists.

TIPS/TRICKS/PEARLS

Accurate portal placement is essential. If the initial portal placement is malpositioned or misdirected, time, frustration, and potential complications can be avoided by establishing a new portal in the optimal location. Using sharp trocars or excessive force to penetrate the capsule can lead to inadvertent damage to the articular cartilage. Once established, screw-in or lock-in cannulas are more secure, particularly when instruments are passed through them frequently. A relatively tight portal of entry through the skin will also help prevent inadvertent withdrawal of the cannula. Clear cannulas improve the visibility of instruments and sutures that are within the tip of the cannula.

It is essential to obtain a clear field of view. Relative hypotensive anesthesia, a hydrodynamic balance of inflow and outflow pressures, irrigation containing epinephrine, and selective radiofrequency cauterization can lead to improved visibility. Repositioning of the joint often improves the view, especially in relatively tight regions (e.g., posterior displacement of the humeral head to improve access to the anterior glenoid or improve working space when addressing subscapularis lesions; adducting the shoulder to safely approach the medial aspect of the lesser tuberosity). Anatomic relationships should be verified prior to resecting or altering any tissue. Motorized instruments and sharp tools must be kept in view to prevent iatrogenic tissue damage.

If a suture is inadvertently offloaded from an anchor with a suture loop eyelet, a new suture can be reintroduced into the anchor (Fig. 1.11A–E). Reposition the suture remaining in the anchor to create asymmetric limbs. Load a free suture in a small atraumatic needle and then pass that needle through the braids of the longer limb of the remaining anchor suture where it exits the working cannula. By placing traction on the short limb of the

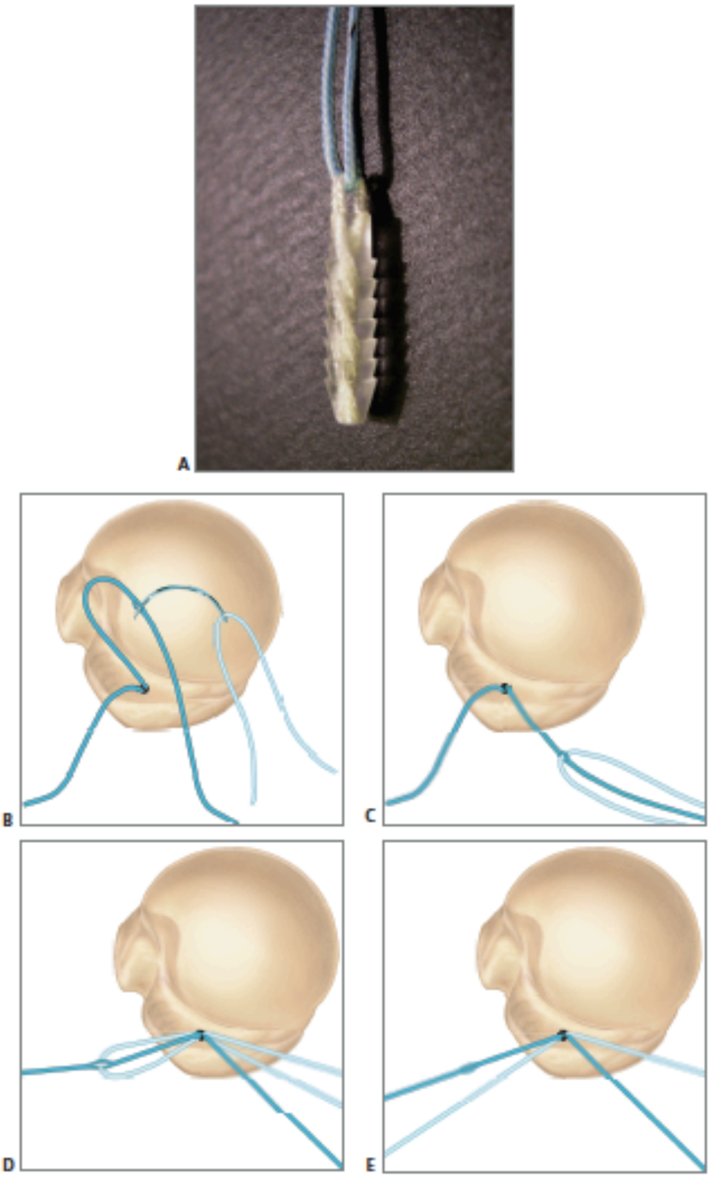


FIGURE 1.11. A: An anchor with a suture eyelet. B: An atraumatic needle delivering a free suture through the braids of the suture, which remains in the anchor. C: A completed pass of the free suture through the anchor suture. D: By pulling on the short limb of the anchor suture, it acts as a shuttle to deliver the free limb through the suture eyelet. E: Both suture limbs are now through the anchor eyelet in a normal fashion.

remaining anchor suture, the free suture can be "shuttled" through the anchor eyelet. Once the two sutures are disengaged, both pass through the suture eyelet.

When a suture is accidently offloaded from an anchor with a rigid eyelet and multiple sutures, a new suture can be secured to the anchor (Fig. 1.12A–C). A simple overhand throw is created outside the cannula with the suture, which still passes through the anchor. A second free suture is passed beneath the loop that has been created. As the half hitch is delivered down the cannula, it draws the second (free) suture to the anchor head. A second and third alternating half hitch are introduced to secure the free strand. Once all limbs are passed through the tissue, the suture passing through the anchor eyelet is tied first, which helps further secure the free strand. Nonsliding knots must be used for both pairs of sutures.

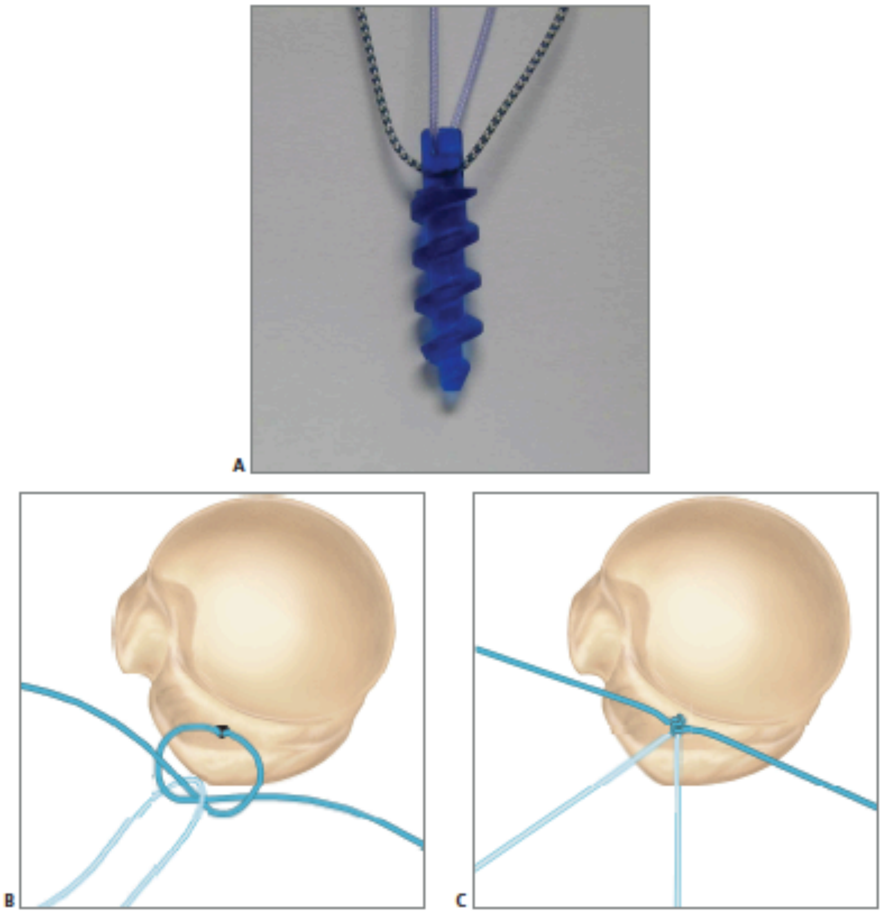


FIGURE 1.12. A: An anchor with a rigid suture eyelet. **B:** An overhand throw is created by the suture remaining through the eyelet and a separate free suture is passed through the loop that is created. **C:** The anchor suture is tied to the anchor head and backed up with two half hitches. Both sutures are now secured to the anchor, but require nonsliding knots to be employed once the sutures are passed through the tissue.

CONCLUSIONS

Either the beach chair or lateral decubitus positions can be used to position patients safely for shoulder arthroscopy. Adequate cerebral blood flow must be maintained when the head and thorax are significantly elevated. General anesthesia is routinely performed and permits greater blood pressure management compared with an ISB. ISBs should be performed by experienced anesthesiologists with a detailed knowledge of the regional anatomy and an opportunity to perform blocks on a routine basis to maintain their skills. The use of ultrasound guidance is recommended.

Accurate portal placement can either greatly facilitate or hinder the performance of any arthroscopic procedure. An 18-G spinal needle will aid in identifying the optimal entry site and path for specific portals. The choice of camera and view orientation is largely surgeon preference, particularly when working in the subacromial space. Manipulating the position and displacement of the shoulder will aid in optimizing the view and working space. A systematic routine for handling sutures will prevent tangling, suture damage, and insecure knots with poor loop security.

REFERENCES

1. Correa MC, Concalves LB, Andrade RP, et al. Beach chair position with instrumental distraction for arthroscopic and open shoulder surgeries. *J Shoulder Elbow Surg*. 2008;17:226–230.
2. Pohl A, Cullen DJ. Cerebral ischemia during shoulder surgery in the upright position: a case series. *J Clin Anesth*. 2005;17:463–469.
3. Papadonikolakis A, Wiesler ER, Olympio MA, et al. Avoiding catastrophic complications of stroke and death related to shoulder surgery in the sitting position. *Arthroscopy*. 2008;24:481–482.
4. Weber S, Jain R. Scalene regional anesthesia for shoulder surgery in a community setting: an assessment of risk. *J Bone Joint Surg Am*. 2002;84:775–779.
5. Bishop J, Sprague M, Gelber J, et al. Interscalene regional anesthesia for shoulder surgery. *J Bone Joint Surg Am*. 2005;87:974–979.
6. Urban M, Urquhart B. Evaluation of brachial plexus anesthesia for upper extremity surgery. *Reg Anesth*. 1994;19:175–182.
7. Ritchie E, Tong D, Chung F, et al. Suprascapular nerve block for postoperative pain relief in arthroscopic shoulder surgery: a new modality? *Anesth Analg*. 1997;84:1306–1312.
8. Matsumoto D, Suenaga N, Oizumi N, et al. A new nerve block procedure for the suprascapular nerve base on a cadaveric study. *J Shoulder Elbow Surg*. 2009;18:607–611.
9. Barber FA. Suprascapular nerve block for shoulder arthroscopy. *Arthroscopy*. 2005;21:1015.
10. Nottage W. Arthroscopic portals: anatomy as risk. *Orthop Clin North Am*. 1993;24:19–26.
11. Stanish W, Peterson D. Shoulder arthroscopy and nerve injury: pitfalls and prevention. *Arthroscopy*. 1995;11:458–466.
12. Lo I, Lind C, Burkhart S. Glenohumeral arthroscopy portals established using an outside-in technique: neurovascular anatomy at risk. *Arthroscopy*. 2004;20:596–602.
13. Meyer M, Graveleau N, Hardy P, et al. Anatomic risks of shoulder arthroscopy portals: anatomic cadaveric study of 12 portals. *Arthroscopy*. 2007;23:529–536.
14. Woolf SK, Guttmann D, Karch MM, et al. The superior-medial shoulder arthroscopy portal is safe. *Arthroscopy*. 2007;23:247–250.

Appendix 3, (Chapter 2.3)

Instructional Course 306

Controversies in Arthroscopic Shoulder Surgery: Arthroscopic Versus Open Bankart Repair, Thermal Treatment of Capsular Tissue, Acromioplasties—Are They Necessary?

Richard L. Angelo, M.D.

A significant debate has involved in the last several years regarding the optimal surgical management of glenohumeral instability. The majority of patients with traumatic anterior-inferior instability sustain a capsulolabral detachment[1,2] described by Perthes[3] and Bankart.[4] The open Bankart repair by Rowe[5] has resulted in a high rate of shoulder stability, although the functional outcome has sometimes been suboptimal, especially for higher-caliber athletes. Although suture anchors have simplified the procedure when compared with bone tunnels,[6] the technique has not changed substantially. In an effort to decrease the morbidity and improve the functional results, Caspari[7,8] developed an arthroscopic capsular shift technique, which began a period of intense interest and study in arthroscopic techniques to stabilize the shoulder. As refinements have been made to the arthroscopic methods, the number of proponents has grown. Current arthroscopic techniques using suture anchors[9] mimic the open Bankart repair and comparable results have been reported.[10-13] Rather than attempt to defend a particular position as to whether an arthroscopic or an open approach to shoulder stabilization is best, the focus here is placed on what we have learned regarding the appropriate indications for each of these techniques.

ARTHROSCOPIC BANKART ADVANTAGES

A number of advantages have been identified for the arthroscopic Bankart repair, including decreased morbidity. Based on visual analog scales, the pain is less than after an open procedure with a decreased need for narcotic medications. The arthroscopic approach requires less disruption of normal anatomy, specifically violation of the subscapularis tendon. During revision of failed open Bankart repairs, the subscapularis tendon, which had been previously divided and repaired, is often quite thin and atrophic. In a small number of cases, a complete failure of the subscapularis repair occurs. Typically, there is less anterior scarring after an arthroscopic repair. It is also possible, although unproven, that there could be a lower incidence of degenerative joint disease of the shoulder after an arthroscopic repair as a result of a greater preservation of normal range of motion.[14,15]

A clear advantage for the arthroscopic technique is the ability to more thoroughly diagnose the extent of pathology throughout the entire glenohumeral joint, including loose bodies, partial articular surface rotator cuff tears, SLAP tears, biceps tendinopathy, and chondral defects. Provided that the joint remains stable, shoulder function is often more normal after an arthroscopic repair. Throwing athletes are more likely to return to the same or higher level of function after an arthroscopic repair.[16] In addition, there is typically less discomfort with overhead activities, including both work and sport pursuits. If an open Bankart repair requires an overnight stay, then an arthroscopic procedure is likely to be less costly. Depending on the experience and skill level of the arthroscopic surgeon, the total operative time may be less than for an open procedure.

OPEN BANKART ADVANTAGES

Open Bankart repairs have been successfully performed for many years and have resulted in a fairly

Address correspondence to Richard L. Angelo, M.D., 12911 120th Ave NE, Suite H-210, Kirkland, WA 98034, U.S.A. E-mail: rlamdortho@aol.com

0749-8063/03/1910-0130$30.00/0
doi:10.1016/j.arthro.2003.10.005

low recurrence rate.[5] Studies of open Bankart repairs, however, often do not include the incidence of subluxations and/or the presence of apprehension, but include only frank dislocations in the reported "failure rate." After an open dissection, it is possible that the subscapularis and capsular repair could add a "buttress" of additional scar tissue aiding in the prevention of recurrent anterior instability. This is particularly true if the subscapularis and capsule have been separated during the exposure and "laminate" together during the healing process. In addition, the open Bankart repair is a technique familiar to most orthopaedic surgeons, requires relatively little special equipment, and results in a reasonably reproducible recurrence rate. It also provides flexibility in addressing large glenoid bone defects with either a coracoid transfer or iliac crests bone graft.

CONSIDERATIONS IN SELECTING AN ARTHROSOPIC VERSUS OPEN REPAIR

Careful patient selection is an essential component in obtaining an optimal result after any shoulder stabilization procedure. There are a number of key issues that must be considered in choosing whether to perform an arthroscopic or open anterior Bankart repair.

It is important to identify the individual patient goals. For one patient with nondominant shoulder instability, a stable shoulder and modest loss of range of motion would be a successful result. For an overhead athlete, a stable shoulder with modest loss of range of motion in their dominant arm would likely preclude a return to their former level of competition and lead to dissatisfaction. Throwing athletes tend to have a somewhat better chance at returning to effective participation after an arthroscopic stabilization. On the other hand, the patient routinely involved in contact and collision sports or very heavy lifting and laboring activities could fare better with an open stabilization. Although some authors do not consider contact activities a contraindication to an arthroscopic stabilization,[17] the majority tend to prefer an open stabilization for these patients. Given the relatively high likelihood of recurrence, select high-demand athletes can be considered for arthroscopic stabilization after a first-time dislocation.[18,19]

The experience of the individual surgeon is also an important consideration. We must each assess our own skill level and ability to arthroscopically treat the entire gamut of glenohumeral pathology contributing to a patient's instability. Usually, experience is gained and skill developed in a "stepwise" fashion. Routinely, different segments of a procedure are mastered until the final surgical procedure is reliable and will withstand careful testing at the completion of the repair. If the arthroscopic procedure is not progressing in a technically sound manner or the time spent becomes excessive, conversion to an open stabilization is prudent.

The patient's history will also help in determining the appropriate approach. For those patients less than 20 years old, there tends to be an increased rate of recurrent instability with both the open and arthroscopic approaches. An open stabilization in this age group is preferred unless the Bankart lesion is traumatic, the capsulolabral tissue robust, and pathologic tissue laxity is absent. The number of instability episodes that the patient reports is likely to loosely parallel the magnitude of the capsular strain and attenuation. Unless the surgeon is confident that the capsular laxity can be effectively addressed arthroscopically, an open procedure should be considered. In general, the specific age and activity level of the patient is somewhat less important than the specific tissue pathology observed, i.e., robust capsular tissue with an intact labral ring versus patulous or pathologically lax capsule with labral obliteration.

On physical examination, indicators of tissue hyperelasticity include elbow hyperextension, excessive patellar laxity, and thumb hyperabduction as well as broadened scars from previous wounds. An open capsular shift procedure is more likely to result in a successful outcome for these individuals. The most accurate assessment of the magnitude of glenohumeral laxity is often obtained during the examination under anesthesia. Translation in the anterior-inferior and posterior directions is determined with the "load and shift" test and inferior laxity determined by the magnitude of the sulcus sign. If posterior laxity is determined to be greater than normal, i.e., the humeral head rides over the posterior glenoid rim, then an arthroscopic posterior capsular augmentation should be considered. If traction is applied to the adducted arm while in 20° of external rotation and the sulcus sign exceeds approximately 0.5 cm, rotator interval closure or plication is recommended.

During the diagnostic portion of the arthroscopy, the character and extent of soft tissue pathology must be carefully evaluated. A well-defined Bankart lesion is likely to be present after a traumatic anterior inferior dislocation. Alternatively, the capsulolabral tissue can also be significantly scarred back to the glenoid rim and the defect not apparent without careful probing. The only indication that an anterior labral periosteal sleeve avulsion (ALPSA) lesion is present could be that the capsulolabral tissue is attached more medial than normal along the neck of the glenoid. The integrity of the midcapsule and inferior glenohumeral ligament should also be evaluated. If they appear markedly stretched, attenuated, or incompetent, consideration should be given for an open repair. Intracapsular tearing must also be sought and could coexist with a labral detachment. In this situation, competency of the anterior capsulolabral tissues will not be restored by simply repairing the labrum to the glenoid rim. A tear of the humeral attachment of the glenohumeral ligaments (HAGL) must always be ruled out during the diagnostic arthroscopy. Although arthroscopic techniques have been described to address this lesion, an open repair is more likely to be reliable.

Bone deficiencies have a significant impact on the selection of the technique. Imaging studies routinely include an anterior-posterior, West Point lateral, and Stryker notch view. If the lateral views suggest a significant anterior glenoid defect, a computed tomography scan assists in operative planning by attempting to quantify the magnitude of the defect. During arthroscopy, the integrity of the bony glenoid rim must be noted. A fracture that comprises less than approximately 20% is not a contraindication to arthroscopic stabilization. If the fragment cannot be fixed, it is excised by operating a motorized burr in reverse to avoid damaging the capsular and periosteal tissues. A significantly higher failure rate after an arthroscopic repair has been reported by Burkhart and DeBeer[11] if >20% to 25% of the glenoid rim is deficient (inverted pear configuration). A coracoid transfer or iliac crests bone graft should be considered in an effort to recreate relatively normal glenoid geometry. If a Hill-Sachs lesion comprises >30% of the articular surface of the humeral head, or engages on the anterior rim of the glenoid with the arm in a position of abduction and external rotation, an open stabilization is recommended. Large humeral head defects could necessitate consideration for an osteochondral graft.

TECHNIQUE OF ARTHROSCOPIC BANKART REPAIR

Patient Orientation

Surgeon preference dictates whether to perform the procedure with the patient in the lateral decubitus or beach-chair position. For the lateral decubitus orientation, the involved forearm and hand should be "suspended" distally and the upper arm laterally with accessory support. Distal "traction" should be avoided because it could oppose retensioning of the glenohumeral ligaments during repair. A standard posterior arthroscopy viewing portal 1.5 cm inferior and 1 cm medial to the posterolateral corner of the acromion is established with a disposable cannula first (this portal will later be used for instrumentation). While viewing the intra-articular space from this portal, a spinal needle establishes the site for the anterosuperior portal. The entry point lies 1 cm lateral and slightly anterior to the anterolateral corner of the acromion. As the portal is established with the arthroscopic sheath, a blunt trocar enters through the capsule immediately anterior to the supraspinatus. Again, while viewing from the posterior portal, the midanterior portal site is located with a spinal needle entering 1.5 cm inferior and 1.5 cm lateral to the inferolateral corner of the coracoid tip. The portal enters immediately superior to the subscapularis tendon. This portal should allow instruments to reach the most inferior aspect of the glenoid without an excessively acute angle in the plane of the glenoid. In addition, an angle of approximately 45° in the transverse plane is desired to provide proper approach to the glenoid rim. If this portal is "too shallow," the drill bit or anchor can skive across the glenoid surface or enter the glenoid neck too medially. The anterosuperior portal provides an excellent view of the anterior glenoid rim and the most inferior aspect of the Bankart lesion. Some surgeons prefer the view from a posterior portal while performing the repair.

GLENOID PREPARATION

Marginal articular cartilage debris is removed with a motorized shaver introduced through the midanterior working portal. It is important when working around the capsule or labrum with a motorized instrument that the suction is turned off to avoid inadvertent capture and damage to those tissues. A liberator-elevator or similar tool is then used to release the capsulolabral tissue from its scarred position and thus expose the lateral 1 cm of the glenoid neck and subscapularis muscle fibers. It is important that the capsular tissue be released around to the 6-o'clock position on the glenoid. If the release is inadequate, the tissues can be tethered medially and inferiorly, preventing adequate superior and lateral retensioning of the tissues. With the suction turned off, a motorized shaver or hooded burr is used to lightly excoriate the glenoid neck surface to prepare for optimal healing.

ANCHOR/SUTURE DELIVERY

Drill holes for the appropriate anchor are placed 2 to 3 mm onto the glenoid surface at the 5:00, 3:30, and 2:00 positions. The anchor of choice is delivered into the inferior hole and tested for security by placing traction on the sutures. When the anchor is inserted, the eyelet should be oriented to permit the sutures to exit anteroinferior and posterosuperior to allow the repair suture to remain collinear with the eyelet and slide easily during knot tying. Using a loop grasper or crochet hook, the anteroinferior suture is retrieved out the posterior cannula. A cannulated suture hook or similar suture-passing device is then introduced through the midanterior portal. A pass is made through the capsule beginning 1 cm inferior and 1 cm lateral to the capsular rim. The instrument is then brought up under the labrum adjacent to the rim of the glenoid. Superior tension is applied to the instrument; if the exiting tip of the suture delivery device is superior to the anchor drill hole, the tissues will not be adequately retensioned and a lax inferior pouch will result. The instrument should be removed and another pass made through the capsule more inferiorly. A No. 0 or 1 polydioxanone suture (PDS) or other suture shuttle is then delivered through the tissues and retrieved out the posterior cannula. Outside the posterior cannula, the PDS is tied near the end of the previously retrieved anteroinferior suture with a simple overhand loop. The PDS shuttles the permanent suture limb through the capsulolabral tissue from posterior to anterior. As this newly passed limb of suture is tensioned, the rim of tissue should be observed to ride up onto the glenoid rim and tighten the inferior glenohumeral ligament. Verify that significant distal traction on the humeral head is not present before knot tying. If tying a sliding not, the limb that passes through the tissue must be the post to deliver the knot away from rather than toward the articular surface. As the knot is thrown, delivered, and secured with alternating half-hitches, the "pursed" tissue creates a pseudolabrum. These steps are repeated for the second and third anchors. Occasionally, four anchors are necessary and are evenly spaced along the glenoid rim from 1:30 to 5:30.

AUGMENTATION

In the majority of cases exhibiting a Bankart lesion, posterior capsular augmentation is unnecessary. If, however, excessive posterior capsular laxity was determined during the examination under anesthesia and the appearance at the diagnostic arthroscopy, it should be addressed with either suture plication or thermal capsulorrhaphy. Suture plication is simple, low risk, and affords the opportunity to be somewhat quantitative in the amount of capsular tightening that is created.[20-22] While viewing from the anterosuperior portal, a rasp or whisker blade shaver is introduced through the posterior portal. The posterior and inferior capsule is slightly excoriated for 2 cm adjacent to the posterior glenoid rim to stimulate healing. A cannulated suture hook is then used to create a small pleat in the capsular tissue. The suture hook tip enters the capsule 1 to 1.2 cm lateral to the intact labral rim and exits approximately 5 or 6 mm lateral to the rim. The tip of the instrument then enters and passes beneath the intact labrum so as to create a small pleat of tissue. The PDS suture is delivered and using the most lateral limb of the suture as a post, a sliding not is securely tied and backed up. Two or 3 additional sutures can be placed to complete the augmentation. Care must be exercised so that an increasingly larger pleat of tissue is not created as the repair progresses superiorly. Otherwise a significant restriction of motion may be created.

An alternative is to perform a thermal capsulorrhaphy. The amount of shrinkage and its effect on the mechanical properties of the capsule are both time- and temperature-dependent. The thermal tip should slightly indent the capsule and must be constantly moving. A "striped" or grid pattern should be created.[23] It is advisable to leave as much normal tissue as that which is thermally shrunk. The magnitude of visible shrinking is quite variable and should not be used as the sole end point for shrinkage, otherwise overheating of the tissues can occur.

A rotator interval closure is considered depending on the magnitude of the sulcus sign and the posterior inferior laxity determined on examination. Two methods have been described. An outside-in approach involves passing the sutures and tying them in the subacromial bursa while imbricating the coracohumeral ligament. A spinal needle is introduced through the anterosuperior skin incision, transgresses the subacromial space, and penetrates the superior glenohumeral ligament as it enters the joint. A sharp-tipped suture retriever is then passed through the midanterior cannula (which has been withdrawn outside the capsule), and passes through the middle glenohumeral ligament to retrieve the PDS suture limb in the joint. The limbs are retrieved and the knot tied in the subacromial space. If this method is chosen, the arm should be placed in external rotation during knot tying to avoid overtightening the shoulder. A second technique involves closing the middle to the superior glenohumeral ligament. A simple method to accomplish this is to introduce a sharp-tipped suture grasper/retriever through the midanterior portal while viewing from the posterior portal. The instrument is used to grasp the PDS suture 3 cm from the end and deliver it through the middle glenohumeral ligament from outside-in. The limb is then retrieved out through the anterosuperior cannula. The end of the remaining limb, which exits the midanterior cannula, is then delivered from outside-in through the superior glenohumeral ligament and retrieved out the anterosuperior cannula. Arthroscopic knots can then be delivered through the anterosuperior portal to affect closure of the rotator interval. Additional sutures can be placed as necessary.

TEST THE REPAIR

It is important to verify that the repair is secure. While viewing arthroscopically, the entire repair should be carefully palpated. In addition, the arm should be removed from any traction or suspension and tested in the jeopardy position of abduction and external rotation. If the repair is inadequate, further measures must be taken to address the residual instability.

REFERENCES

1. Taylor DC, Arciero RA. Pathologic changes associated with shoulder dislocations: arthroscopic and physical examination findings in first time anterior dislocations. *Am J Sports Med* 1997;25:306-311.
2. Baker CL, Uribe JW, Whitman C. Arthroscopic evaluation of acute initial anterior shoulder dislocations. *Am J Sports Med* 1990;18:25-28.
3. Perthes G. Uber operationen bei haituellar schulterluxation. *Dtsch Zischierchi* 1906;85:199.
4. Bankhart ASB. The pathology and treatment of recurrent dislocation of the shoulder joint. *Br J Surg* 1938;26:23-29.
5. Rowe CR, Patel D, Southmayd WW. The Bankart procedure: a long-term end-result study. *J Bone Joint Surg Am* 1978;60:1-16.
6. Levine WN, Richmond JC, Donaldson WR. Use of suture anchor in open Bankart reconstruction. A follow-up report. *Am J Sports Med* 1994;22:723-726.
7. Caspari RB. Arthroscopic reconstruction of anterior shoulder instability. *Tech Orthop* 1988;3:59-66.
8. McIntyre LF, Caspari RB. The rationale and technique for arthroscopic reconstruction of anterior shoulder instability using multiple sutures. *Orthop Clin North Am* 1993;24:55-58.
9. Wolf EM, Wilk RM, Richmond JL. Arthroscopic Bankart repair with suture anchors. *Oper Tech Orthop* 1991;1:187-191.
10. Bacilla P, Field LD, Savoie FH. Arthroscopic Bankart repair in a high-demand patient population. *Arthroscopy* 1997;13:51-60.
11. Burkhart SS, De Beer JF. Traumatic glenohumeral bone defects and their relationship to failure of arthroscopic Bankart repairs: significance of the inverted-pair glenoid and the humeral engaging Hill-Sachs lesion. *Arthroscopy* 2000;16:677-694.
12. Sperber A, Hamberg P, Karlsson J, et al. Comparison of an arthroscopic and an open procedure for posttraumatic instability of the shoulder: a prospective, randomized multicenter study. *J Shoulder Elbow Surg* 2001;10:105-108.
13. Warner JJP. Treatment options for anterior instability: open versus arthroscopic. *Oper Tech Orthop* 1995;5:233-237.
14. Cole BJ, Warner JJ. Arthroscopic vs open Bankart repair for traumatic anterior shoulder instability. *Clin Sports Med* 2000;19:19-48.
15. Kim SH, Ha KI, Kim SH. Bankart repair in traumatic anterior shoulder instability: open vs arthroscopic technique. *Arthroscopy* 2002;18:755-763.
16. Weber SC. A prospective evaluation comparing arthroscopic and open treatment of recurrent anterior glenohumeral dislocation. *Orthop Trans* 1991;15:763 (abstr).
17. Mologne TS, McBride MT, Lapoint JM. Assessment of failed arthroscopic labral repairs. Findings as open surgery. *Am J Sports Med* 1997;25:813-817.
18. Kirkley A, Griffin S, Richards C, Miniaci A, Montadi N. Prospective randomized clinical trial comparing the effectiveness of immediate arthroscopic stabilization versus immobilization in first traumatic anterior dislocations of the shoulder. *Arthroscopy* 1999;15:507-514.
19. Bottoni CR, DeBernardino TM, Wilckens JH, D'Alleyrand JC. A prospective, randomized evaluation of arthroscopic stabilization versus nonoperative treatment of acute, traumatic, first-time shoulder dislocation. *Arthroscopy* 2000;16:432 (abstr).
20. Abrams JS, Savoie FS III, Tauro JC, Bradley JP. Recent advances in the evaluation and treatment of shoulder instability: anterior, posterior, and multidirectional. *Arthroscopy* 2002;18:1-13 (suppl).
21. Wolf EM, Eakin CL. Arthroscopic capsular plication for posterior shoulder instability. *Arthroscopy* 1998;14:153-163.
22. Antoniou J, Duckworth DT, Harryman DT III. Capsulolabral augmentation for the management of posteroinferior instability of the shoulder. *J Bone Joint Surg Am* 2000;82:1220-1230.
23. Lu Y, Hayashi K, Edwards RB, Fanton GS, Thabit G, Markel MD. The effect of monopolar radiofrequency treatment patten on joint capsular healing: in vitro and in vivo studies using an ovine model. *Am J Sports Med* 2000;28:711-719.

Appendix 4, (Chapter 2.4)

31

SYMPOSIUM

Arthroscopic Bankart Repair for Unidirectional Shoulder Instability

*Richard L. Angelo, MD

Abstract

A successful arthroscopic Bankart repair for unidirectional shoulder instability requires careful patient selection and, to the extent possible, the restoration of normal anatomy. The patient's goals and anticipated demands are important considerations. A patient who participates in an overhead sport requires not only a stable shoulder but also a full range of shoulder motion. An athlete who engages in a contact or collision sport, however, may tolerate a mild loss of motion provided the shoulder is stable. Compared with an open procedure, an arthroscopic repair provides the opportunity to retain the most normal postoperative range of motion and function. Other considerations include patient age, which often relates to overall tissue laxity, and the number of previous instability episodes, which correlates with the severity of pathology (in particular, capsulolabral strain, glenoid chondromalacia, and bony deficiency of the glenoid or posterior humeral head). The magnitude of bone loss, particularly for the anterior glenoid, can make an arthroscopic repair inadvisable. Accurate portal placement, glenoid preparation, anchor insertion, and suture passage are key components of the arthroscopic technique, but the most important overall goal is the secure restoration of capsulolabral tissue tension. Secondary posteroinferior laxity, partial rotator cuff tears, labral disorders, and articular cartilage pathology may also require treatment.

Instr Course Lect 2009;58:305-313.

The choice of surgical procedure to correct instability is determined in part by the patient's anticipated shoulder demands and functional goals. Patient history (including such factors as age, tissue quality, and previous instability episodes), and clinical and radiographic findings also are important considerations.

Considerations in Decision Making

Patient Goals

Shoulder instability recurs in 7% to 10% of patients after an open or arthroscopic suture anchor repair.[1-4] The choice of surgical procedure to correct instability is determined in part by the patient's goals, including anticipated shoulder demands. Although an open Bankart repair is a reliable method for eliminating clinical instability, range of motion and overall function may be unacceptably compromised in some patients who engage in high-demand activities.[5] In particular, overhead athletes, including throwers, cannot tolerate restrictions in flexion or external rotation. The patient's range of motion usually returns more rapidly and completely after an arthroscopic rather than an open repair, and the patient is more likely to be able to return to competitive throwing.[6] However, overhead athletes have a lower overall rate of functional success (70%) than other athletes (90%) following an arthroscopic repair.[7]

A comparison of 30 open and 30 arthroscopic Bankart repairs found that muscle strength for forward elevation was markedly weaker after open repair for as long as 3 months; the difference, however, was only 5% after 6 months.[6] Muscle strength for external and internal rotation was significantly weaker 6 weeks after open repair but also approached 5% after 6 months. In a biomechanical investigation of arthroscopic anterior repair, a traumatic dislocation

Richard L. Angelo, MD is a consultant for or an employee of DePuy and Mitek.

was created in 12 cadaver specimens, and an arthroscopic suture anchor repair of the Bankart lesion was performed. Glenohumeral translation and rotation were found to approach normal predislocation values.[8]

Arthroscopic Bankart repair remains controversial for athletes who participate in contact and collision sports. Several recent studies have concluded that there is no increased risk of recurrent instability for these athletes after an arthroscopic procedure.[2,7,9] The overall recurrence rate was 10% in 85 patients treated with a suture anchor arthroscopic anterior repair.[7] Two patients had a recurrence in a subset of 18 collision sport athletes (a similar 11% recurrence rate). In a review of contact or collision sport athletes who were younger than 20 years, 2 of 18 patients (11%) had a recurrent dislocation at a minimum 2-year follow-up but did not require further treatment.[9] In another evaluation of suture anchor repairs, the 9.5% recurrence rate for contact sport athletes (2 of 21 patients) was not significantly different from the 6% overall failure rate.[2]

A review of 48 shoulders in 46 collision sport athletes reached a different conclusion. Sixteen of the shoulders were arthroscopically stabilized (4 using Suretacs [Smith and Nephew Endoscopy, Mansfield, MA] and 16 using suture anchors) and 32 underwent open repair. Instability recurred after 25% of the arthroscopic repairs (one using Suretacs and three using suture anchors) compared with 12.5% of the open repairs. The authors concluded that open stabilization is a more reliable method of repairing anterior shoulder instability in contact athletes.[6]

Patient History

Patient age is believed to affect the probability of failure after an open or arthroscopic anterior stabilization procedure. Patients in their teens generally have greater tissue and collagen elasticity, which may predispose them to a higher likelihood of repair failure. In addition, younger patients are more likely to be attracted to high-risk activities, including so-called extreme sports such as snowboarding and aggressive mountain biking. Few studies have specifically evaluated pediatric patients with shoulder instability. In a review of 32 arthroscopic Bankart repairs in patients age 11 to 18 years, 16 shoulders were repaired after unsuccessful nonsurgical treatment and 16 were repaired after the initial instability episode.[10] At an average 25-month follow-up, three redislocations had occurred in two patients from the first group (18.5%), and two redislocations had occurred in two patients (12.5%) from the second group. The small size of the study, however, does not permit a conclusion as to the optimal treatment of pediatric patients with shoulder instability. The choice of an open or arthroscopic procedure should be based on tissue quality and capsulolabral integrity rather than solely on the patient's age. If traumatic pathology is identified and the patient's soft tissues are reasonably robust, an arthroscopic suture anchor technique can provide a reliable repair.

The number of previous instability episodes also should be considered because capsular strain, labral tearing, and glenoid erosion tend to increase with each occurrence of instability.[11] The severity of the accrued pathology must be evaluated during the diagnostic arthroscopy and helps determine whether an arthroscopic repair is suitable.

Physical Examination

The findings of the physical examination must support the clinical impression of unidirectional shoulder instability. Once the dislocated shoulder is reduced, any deformity usually disappears. There may be diffuse tenderness over the anterior capsular tissues and, less frequently, along the posterior glenohumeral joint line. The patient's range of motion is often restricted following an instability event secondary to pain. In those with chronic instability, acquired anterior capsular laxity may result in an increase in external rotation over their norm. Excessive anterior translation typically appears on the load-and-shift test unless involuntary muscular guarding is present. The magnitude of laxity in the posterior and inferior directions can help determine the need for accessory posterior plication and closure of the rotator interval, respectively. Most patients exhibit apprehension when the shoulder is placed in a position of abduction and external rotation. The relocation test is positive if the apprehension sign is minimized or eliminated when the examiner's hand is placed over the anterior aspect of the proximal humerus and prevents anterior subluxation of the humeral head. With the arm in full adduction and 30° of external rotation, a sulcus sign of more than 1 cm suggests that significant multidirectional laxity is present. In patients with true multidirectional instability, however, clinical symptoms must also be present in more than one direction. Hyperelasticity findings (the ability of the thumb to be passively placed against the forearm, elbow hyperextension, and marked medial/lateral translation of the patellas) suggest that the collagen tissue is pathologic, which is a known risk factor for failure after arthroscopic stabilization.[12]

Imaging

Routine radiographs should be obtained. The AP view may show a fracture of the inferior glenoid margin; a West Point axillary lateral view is more sensitive for anterior and inferior rim fracture fragments; the Stryker notch view identifies the presence and size of a Hill-Sachs defect of the posterior humeral head. A CT scan, especially with three-dimensional reconstruction, is useful for assessing the size of a glenoid rim deficiency or fracture fragment, and a Hill-Sachs lesion of the humerus.

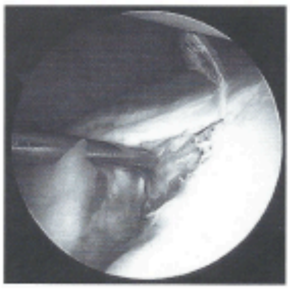


Figure 1 Probe entering a right-shoulder acute Bankart lesion through the posterior portal.

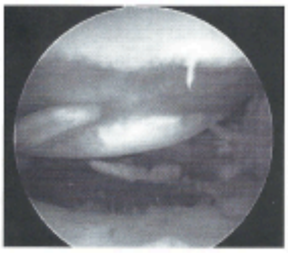


Figure 2 An acute Bankart lesion of the humeral head with chondral damage and loose bodies (posterior portal, right shoulder).

Diagnostic Arthroscopy

A thorough arthroscopic assessment of the instability pathology is imperative. The extent and nature of an acute Bankart lesion are usually apparent (Figure 1). The humeral head must often be displaced posteriorly to detect the inferior extent of capsulolabral detachment from the glenoid. The true capsular margin may be difficult to identify if the labrum has been obliterated. An anterior labroligamentous periosteal sleeve avulsion may be difficult to detect in the chronically unstable shoulder; the most reliable clue to its presence is that the capsulolabral tissue appears to be attached too far medially along the glenoid neck (3 to 4 mm medial to the rim). Chondral damage may also have occurred in a shoulder with chronic instability (Figure 2).

Bone loss can result from a fracture or progressive erosion. Thin glenoid rim fracture fragments often remain securely affixed to the capsule and can be detected only by palpation of the tissues with a hook probe. These small wafer fragments can be repaired or simply excised. Evidence suggests that rim fracture fragments larger than 10% of the glenoid diameter should be preserved. In a review of 42 shoulders with posttraumatic recurrent anterior instability, CT was used to estimate the glenoid defect size, which ranged from 11.4% to 38.6%.[13] The bony rim fragment was incorporated during arthroscopic Bankart repair with 39 of the shoulders rated good or excellent on the University of California Los Angeles Shoulder Scale at an average 39-month follow-up. Two reinjuries were reported. In another study, 21 patients with a bony deficiency of the glenoid, including 11 with a traumatic rim fracture and 10 without an identifiable fragment but with attritional bone loss, had a suture anchor arthroscopic Bankart repair.[14] At a mean 34-month follow-up, 2 of the 21 patients had recurrent subluxation, and 1 had a recurrent dislocation. None of the patients with repair of a rim fracture fragment had an episode of postsurgical instability. In a study of 65 patients (41 with acute instability and 24 with chronic instability) who had undergone an arthroscopic suture anchor repair of a bony Bankart lesion, 2 patients (1 with acute and 1 with chronic instability) experienced a redislocation at a minimum 4-year follow-up.[15] The average Rowe score of the patients with acute instability improved from 59 to 92, and the score of those with chronic instability improved from 43 to 61. Glenoid rim erosion may increase with recurrent episodes of instability. Arthroscopic Bankart repair has an unacceptably high failure rate if there is significant anteroinferior glenoid bone loss. In a retrospective review of 194 consecutive arthroscopic Bankart repairs using suture anchors, two groups of patients were identified based on whether or not significant glenoid or humeral bone loss was present.[16] Glenoid loss was considered significant if the normal pear-shaped configuration of the glenoid had changed to an "inverted pear" shape, in which bone loss resulted in the inferior one half of the glenoid becoming narrower than the superior half. A 67% failure rate following arthroscopic Bankart repair was found in patients with an inverted pear-shaped glenoid or a significant Hill-Sachs lesion, compared with 4% for patients without a significant bony defect. The size of larger glenoid defects can be estimated using the central bare spot as a reference. The normal radius of the glenoid (inferior two thirds) is the distance

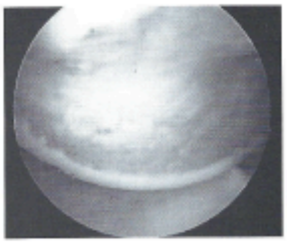


Figure 3 Posterior view of a Hill-Sachs lesion (posterior portal, right shoulder)

from the central bare spot to the intact posterior glenoid rim. The difference of the normal radius and the distance from the bare spot to the remaining anteroinferior glenoid margin can be used to estimate the percentage defect (for example, a defect of one half the length of the radius is approximately 25%).

Hill-Sachs lesions are common after shoulder instability, especially if the episodes are recurrent (Figure 3). Most of these lesions are relatively small and can be ignored without compromising the success of the repair. However, a specific subgroup is associated with a higher failure rate after arthroscopic stabilization.[17,18] In the study cited above regarding glenoid rim deficiency, significant Hill-Sachs lesions were also defined.[16] Three of 21 failures in that study were deemed to have been caused by an "engaging" Hill-Sachs lesion, wherein the posterior humeral defect engaged on the anterior glenoid rim with the arm in a functional position of abduction and external rotation. If the loss is greater than 30% to 35% of the articulating surface of the humeral head, an open osteoarticular allograft or arthroscopically assisted transhumeral impaction grafting may be considered.[17-19]

Translation of the humeral head can be difficult to quantify. However, a qualitative estimation of anterior capsular laxity or strain is useful in determining how much to plicate the capsule with each anchor suture. In addition, the magnitude of posterior laxity helps determine whether or not to augment the repair with several posterior "pinch-tuck" capsular plication sutures. A traumatic midcapsular rent or tear can exist, even in the presence of a separate Bankart lesion. In a prevalence study, 12 of 303 shoulders undergoing stabilization (4%) had a midcapsular tear in addition to a Bankart lesion.[20] Eleven of the 12 tears were repaired arthroscopically, with the average Rowe score of those patients improving from 30.4 to 90.4 at 31-month follow-up. In a review of 21 patients with a midcapsular tear, 7 tears were isolated and 14 were accompanied by a Bankart lesion.[21] More than 90% of the patients had a good or excellent Rowe score after an open or arthroscopic capsular repair along with a Bankart repair when indicated. The average loss of external rotation was 8° for patients with an isolated capsular closure and 16° for those who also had a Bankart procedure.

Humeral avulsion of the glenohumeral ligaments is recognized as a cause of recurrent shoulder instability. The lesion may not be readily apparent and must be carefully sought during diagnostic arthroscopy by examining the anterior and posterior capsular insertions onto the humeral neck. These avulsions can be repaired arthroscopically, although the procedure is technically demanding.[22] Considerably less morbidity is generated with a posterior arthroscopic repair than an open posterior approach. For an anterior avulsion, however, a standard open deltopectoral approach provides ready access to the anterior neck of the humerus with only partial subscapularis detachment.

Any associated lesion (superior or posterior labral lesion, chondral injury, or partial-thickness rotator cuff tear) should be identified and treated.

Technique for a Suture Anchor Arthroscopic Bankart Repair

A successful arthroscopic Bankart repair requires careful patient selection, a thorough understanding of normal and pathologic anatomy, skill in using arthroscopic tools and implants to approximate normal anatomy, and discernment in guiding the postsurgical rehabilitation program. The technique is not inordinately difficult, but mastery requires study and practice. The necessary skills can be honed in dry model and cadaver laboratories. Thorough planning and the ability to mentally rehearse the procedure are invaluable preparations. The operating room staff must be oriented to the sequence of steps and instruments to minimize miscues and optimize efficiency.

Optimal Visualization

A clear arthroscopic view of the intra-articular structures is essential and is improved by using mildly hypotensive anesthesia (approximately 100 mm Hg systolic pressure). Epinephrine can be introduced into the irrigant to help control bleeding (1 mL of 1/1000 epinephrine per 3 L). Although the procedure can be satisfactorily performed using gravity inflow, a pump allows blood pressure spikes to be offset with a temporary increase in inflow pressure. The pump pressure must be carefully monitored to prevent excessive fluid extravasation.

Patient Positioning

The patient's position must allow access to all areas of pathology. In the lateral decubitus position, the pelvis and thorax are supported by a vacuum pack (bean bag) with a 20° posterior tilt of the chest to facilitate access and orient the glenoid approximately parallel to the floor. A 5- to 10-lb weight attached to the arm sleeve permits suspension rather than traction of the arm. Excessive traction may compromise the ability to adequately retension the soft tissues in a superior direction and can cause undue tension on the brachial plexus during a prolonged procedure. A 5- to 10-lb accessory lateral traction can be oriented perpendicular to the humerus to aid in separating the humeral head from the glenoid and provide additional working space. Alternatively, an assistant can manually displace the humeral head posteriorly as the need arises. The standard beach chair position may be preferred because of the normal, upright orientation of the shoulder anatomy. However, the posterior aspect of the shoulder is relatively difficult to access arthroscopically with the patient sitting. For a patient in the beach chair position, a higher systolic blood pressure is necessary to maintain adequate cerebral perfusion; the normal compensatory mechanisms for cerebral blood flow may be compromised with the patient in a sitting position under general anesthesia. Conversion to an open procedure is simplest if the patient is in the beach chair position but also can be accomplished relatively easily from the lateral decubitus position.

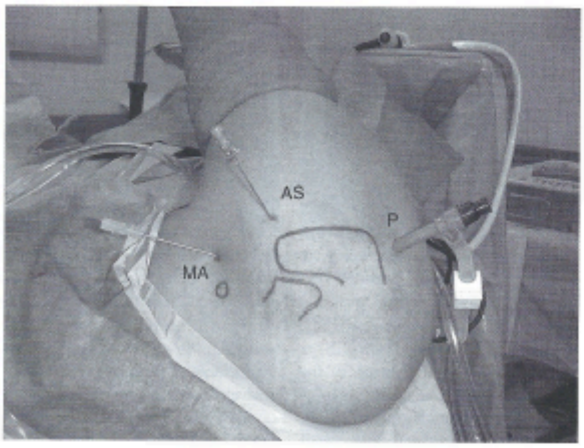


Figure 4 The standard arthroscopic portals. AS = anterosuperior, MA = midanterior, P = posterior.

Portals

Accurate portal placement facilitates identification and treatment of all intra-articular pathology. Poorly placed portals create difficulty in preparing the tissues for repair, placing sutures and anchors, and tying knots. The entry site for the posterior portal is 1.5 cm inferior and 1.0 cm medial to the posterolateral corner of the acromion. The cannula and trocar are directed toward the coracoid tip anteriorly (Figure 4). If drills and anchors may be required for the posterior glenoid, the entry site should be adjusted 1.0 cm farther lateral to provide an acceptable approach to the narrow posterior rim of the anteverted glenoid.

The midanterior portal provides access to the anterior glenoid for instruments. The entry site is located 1.5 cm lateral and 1.5 cm inferior to the coracoid tip. A spinal needle should be used to verify accurate placement; it is initially directed slightly superior, over the superior border of the subscapularis. When the arm is relatively adducted, the subscapularis is fairly lax and can be depressed inferiorly by the incoming needle or cannula and permit ready access to the anteroinferior glenoid. This portal should provide a 30° to 45° approach to the glenoid rim in the transverse plane, which is essential for safe drilling and anchor insertion. An 8.5-mm clear threaded cannula is optimal for this critical working portal.

An anterosuperior portal can be used as a working portal, with viewing from the posterior portal. Alternatively, the anterosuperior portal can be used as the primary viewing portal, and only the arthroscopic sheath needs be inserted. The optimal entry site is 1.0 cm lateral and slightly anterior to the anterolateral corner of the acromion. The proper path is established using a spinal needle that enters immediately anterior to the supraspinatus tendon and through the rotator interval, either anterior or posterior to the biceps tendon.

Glenoid Preparation

A full-radius synovial resector is used to débride the articular glenoid margins, removing ragged or unstable articular cartilage (Figure 5). To

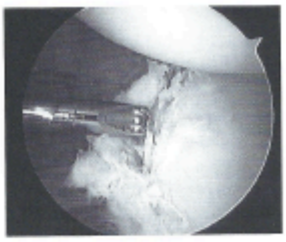


Figure 5 Shaver débridement of chondral and labral damage (anterosuperior portal, right shoulder).

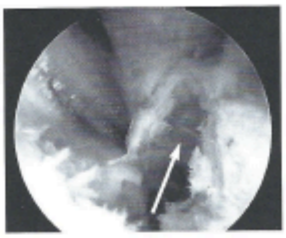


Figure 6 Exposed subscapularis muscle deep to the capsule *(arrow)* (anterosuperior portal, right shoulder).

Figure 7 Orientation of the glenoid drill hole in the transverse plane *(long arrow)*; intact bone lateral to the drill hole *(short arrow)* (anterosuperior portal, left shoulder). G = glenoid, HH = humeral head.

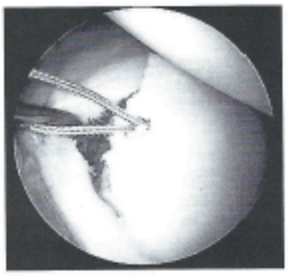


Figure 8 Suture anchor placed 2 to 3 mm onto the face of the glenoid, with intact anterior bone margin (anterosuperior portal, right shoulder).

prevent inadvertent damage to the sometimes fragile adjacent capsulolabral tissue, suction on the shaver should be turned off. The repair must restore normal capsulolabral tissue tension to prevent future glenohumeral instability. Adequate retensioning requires mobilization of the capsular tissue, which may have scarred medially along the glenoid neck. A liberator elevator is introduced through the midanterior portal and used to free the scarred capsular tissue from bone. When the release is complete, subscapularis muscle tissue should be visible medial to the capsule (Figure 6). Adequate mobilization allows the capsule to be advanced both superiorly and laterally onto the glenoid rim. An aggressive shaver or a 4.0-mm burr run in reverse is used to lightly excoriate the anterior neck of the glenoid to provide a bed for tissue healing.

A thin rim fracture fragment (1 to 2 mm wide) can be repaired or excised. If the fragment is removed with a burr, the reverse setting should be used to protect the underlying periosteum and thus enhance the integrity of the repair. Reduction and repair should be performed for a fragment larger than 10% of the glenoid diameter. The fragment can be encircled with anchor sutures, or alternatively, sutures can be passed through the fragment by predrilling with a small Kirschner wire. It is often necessary to introduce the Kirschner wire from the posterior portal to safely approach the separate anterior bone segment. Great care must be exercised to avoid inadvertent wire penetration into the soft tissues anterior to the capsule. Overreducing the rim fragment and creating articular incongruity should be avoided.

Anchor Placement

Anchor holes should be drilled 2 to 3 mm onto the articular surface of the glenoid (Figures 7 and 8). The anterior wall of the drill hole must have sufficient integrity to prevent the strong repair sutures from cutting out anteriorly and medially during the healing period, rendering the repair ineffective. The drill bit should approach the glenoid at approximately 45° in the transverse plane. If the approach angle is too shallow, there is a risk that the bit will skive onto the articular cartilage or that the anchor hole will be located too medial along the glenoid neck. Anchors should be evenly spaced between the 5-o'clock and 2-o'clock positions on the glenoid. After an anchor is implanted, its security should be tested by firmly pulling on the suture strands. Higher rates of repair failure have been reported when fewer than three or four anchors are used.[12] Whether made of metal or absorbable material, loose or prominent anchors may cause significant articular cartilage damage. Nonmetallic, nonresorbable anchors made of polyether ether ketone material have been introduced to avoid the cavitary bone cysts sometimes associated with resorbable anchors.

Reports on the use of knotless anchors are conflicting. In one study, experienced users documented excellent results at a minimum 2-year follow-up; knotless anchors failed in only 5 of 72 patients (6.9%), all of whom were younger than age 22 years.[23] After a similar follow-up period, another retrospective review reported failure of 5 of 21 knotless anchor repairs (23.8%) compared with 3 of 61 repairs using knot-tying anchors (4.9%).[24]

Capsulolabral Retensioning

Restoration of normal anatomy during the repair is essential, to the extent possible. A review of 24 patients who underwent revision surgery after an unsuccessful open anterior repair reported a persistent or recurrent Bankart lesion in 16, with capsular redundancy in 4.[25] Viewing is from the anterosuperior portal. A serrated drill guide provides a more secure purchase on the glenoid face than a fishmouth style tip. After drilling a hole in the appropriate orientation, an anchor is inserted through the midanterior portal and the sutures are retrieved out the posterior cannula. A clamp is used to identify the limb that exits the anchor inferiorly so that when this suture is shuttled back through the tissue, a 180° twist of the suture at the anchor eyelet is avoided. Glenohumeral reduction should be maintained during suture placement, and arm traction should not be permitted to cause inferior subluxation of the humeral head. If the capsulolabral tissue is markedly displaced inferiorly, it may be necessary to pass a suture through the capsule and apply superior traction to appropriately tension the capsule while introducing the repair sutures (Figure 9).

The goal is to introduce the inferior limb of the anchor suture approximately 1.0 cm inferior and 1.0 cm lateral to the anchor exit point from the glenoid rim. This placement will permit superior advancement as well as mediolateral plication of the capsule when the suture is tied. A curved, cannulated suture hook is passed down the midanterior portal, through the capsule, and then up beneath any remnant of the labrum. Once the hook has been placed through the tissue, it is brought superiorly with moderate tension to check for proper placement; if the hook can be displaced superior to the anchor site, the capsular tissues will not be appropriately tensioned when the suture is tied, and the hook must be replaced farther inferior. When the hook is correctly placed, a monofilament shuttle suture is delivered and

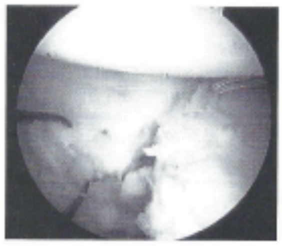


Figure 9 Monofilament traction suture used to retension the capsule superiorly (anterosuperior portal, right shoulder).

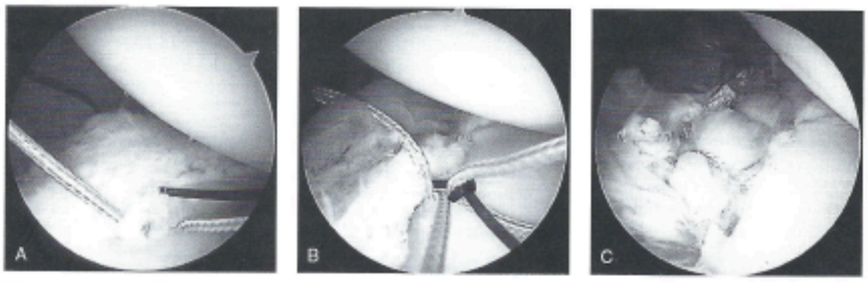


Figure 10 **A,** Monofilament shuttle suture passing through the capsule inferior and lateral to the glenoid anchor site. **B,** Monofilament suture shuttling the inferior limb of the anchor suture through capsulolabral tissue. **C,** The completed Bankart repair, with secure reattachment of capsulolabral tissue to the glenoid rim (All views are from the anterosuperior portal, right shoulder.)

retrieved out the posterior cannula, using a loop grasper (Figure 10, *A*). Using a simple overhand throw, this posterior limb of monofilament suture is tied around the tail of the inferior anchor suture limb (previously identified with a clamp), which is then shuttled from posterior to anterior through the capsule (Figure 10, *B*). This limb, which passes through the tissue, becomes the post for a sliding knot that is delivered laterally away from the glenoid as it is introduced. Sliding knots are backed up with three or four half hitches. These steps are repeated for each anchor and suture pair (Figure 10, *C*).

Test Repair

The repair should be palpated with a nerve hook to ensure that the sutures are tight, the capsulolabral tissue is securely fixed to the glenoid rim, and appropriate tension has been restored. The arm is then removed from suspension and shoulder stability and acceptable range of motion is confirmed. Absorbable subcutaneous sutures and adhesive strips complete the portal closures.

Summary

For an appropriately selected patient, arthroscopic Bankart repair can effectively restore capsulolabral tension and functional shoulder stability while optimizing postsurgical range of motion. The patient's anticipated demands and the surgeon's familiarity and experience with arthroscopic techniques may affect the choice of an open or arthroscopic stabilization procedure. Although youth is not an absolute contraindication to an arthroscopic anterior repair, if significant tissue laxity is present in a pediatric patient, the shoulder should be stabilized using an open procedure. If many instability episodes have led to marked capsular attenuation or significant bony loss, an open repair also is advisable. A CT scan is particularly valuable in assessing the extent of anterior glenoid bone loss and magnitude of posterolateral humeral head impression defects. An arthroscopic Bankart repair requires adequate mobilization of the capsulolabral tissue, careful glenoid preparation, secure anchor placement, and accurate suture delivery. These steps, when properly performed and followed by appropriate rehabilitation, will lead to a successful arthroscopic shoulder stabilization.

References

1. Marquardt B, Witt KA, Liem D, Steinbeck J, Pötzl W: Arthroscopic Bankart repair in traumatic anterior shoulder instability using a suture anchor technique. *Arthroscopy* 2006;22: 931-936.
2. Carreira DS, Mazzocca AD, Oryhon J, Brown FM, Hayden JK, Romeo AA: A prospective outcome evaluation of arthroscopic Bankart repairs: Minimum 2 year follow-up. *Am J Sports Med* 2006;34:771-777.
3. Kim SH, Ha KI, Cho YB, Ryu BD, Oh I: Arthroscopic anterior stabilization of the shoulder: Two to six year follow-up. *J Bone Joint Surg Am* 2003; 85:1511-1518.
4. Kim S-H, Ha K-I, Kim Y-M: Arthroscopic revision Bankart repair: A prospective outcome study. *Arthroscopy* 2002;18:469-482.
5. Fabbriciani C, Milano G, Demontis A, Fadda S, Ziranu F, Mulas PD: Arthroscopic versus open treatment of Bankart lesion of the shoulder: A prospective randomized study. *Arthroscopy* 2004;20:456-462.
6. Rhee Y, Ha J, Cho N: Anterior stabilization in collision athletes: Arthroscopic versus open Bankart repair. *Am J Sports Med* 2007;35:148.
7. Ide J, Maeda S, Tagaki K: Arthroscopic Bankart repair using suture anchors in athletes. *Am J Sports Med* 2004;32:1899-1905.
8. Marquardt B, Hurschler C, Schneppendahl J, Witt KA, Pötzl W, Steinbeck J: Quantitative assessment of glenohumeral translation after anterior shoulder dislocation and subsequent arthroscopic Bankart repair. *Am J Sports Med* 2006;34:1756-1762.
9. Mazzocca AD, Brown FM Jr, Carreira DS, Hayden J, Romeo AA: Arthroscopic anterior shoulder stabilization of collision and contact athletes. *Am J Sports Med* 2005;33:52-60.
10. Jones KJ, Wiesel B, Ganley TJ, Wells L: Functional outcomes of early arthroscopic Bankart repair in adolescents aged 11-18 years. *J Pediatr Orthop* 2007;27:209-213.
11. Habermeyer P, Gleyze P, Rickert M: Evolution of lesions of the labrum-ligament complex in post-traumatic anterior shoulder instability: A prospective study. *J Shoulder Elbow Surg* 1999;8:66-74.
12. Boileau P, Villalba M, Hery JY, Balg F, Ahrens P, Neyton L: Risk factors for recurrence of shoulder instability after arthroscopic Bankart repair. *J Bone Joint Surg Am* 2006;88: 1755-1763.
13. Sugaya H, Moriishi J, Kanisawa I, Tsuchiya A: Arthroscopic osseous Bankart repair for chronic recurrent traumatic anterior glenohumeral instability. *J Bone Joint Surg Am* 2005;87: 1752-1760.
14. Mologne TS, Provencher MT, Menzel KA, Vachon TA, Dewing CB: Arthroscopic stabilization in patients with an inverted pear glenoid: Results in patients with bone loss of the anterior glenoid. *Am J Sports Med* 2007;35: 1276-1283.
15. Porcellini G, Paladini P, Campi F, Paganelli M: Long-term outcome of acute versus chronic bony Bankart lesions managed arthroscopically. *Am J Sports Med* 2007;35:2067-2072.
16. Burkhart SS, De Beer JF: Traumatic glenohumeral defects and their relationship to failure of arthroscopic Bankart repairs: Significance of the

inverted-pear glenoid and the humeral engaging Hill-Sachs lesion. *Arthroscopy* 2000;16:677-694.

17. Kropf EJ, Sekiya JK: Osteoarticular allograft transplantation for large humeral head defects in glenohumeral instability. *Arthroscopy* 2007;23:322-325.
18. Kazel MD, Sekiya JK, Greene JA, Bruker CT: Percutaneous correction (humeroplasty) of humeral head defects (Hill-Sachs) associated with anterior shoulder instability: A cadaveric study. *Arthroscopy* 2005;21:1473-1478.
19. Re P, Gallo R, Richmond J: Transhumeral head plasty for large Hill-Sachs lesions. *Arthroscopy* 2006;22:798-804.
20. Mizuno N, Yoneda M, Hayashida K, Nakagawa S, Mae T, Izawa K: Recurrent anterior shoulder dislocation caused by a midsubstance complete capsular tear. *J Bone Joint Surg Am* 2005;87:2717-2723.
21. Rhee YG, Ha JH, Park KJ: Clinical outcome of anterior shoulder instability with capsular midsubstance tear: A comparison of isolated midsubstance tear and midsubstance tear with Bankart lesion. *J Shoulder Elbow Surg* 2006;15:586-590.
22. Richards DP, Burkhart SS: Arthroscopic humeral avulsion of the glenohumeral ligaments (HAGL) repair. *Arthroscopy* 2004;20:134-141.
23. Thal R, Nofziger M, Bridges M, Kim JJ: Arthroscopic Bankart repair using knotless or bioknotless suture anchors: 1-7 year results. *Arthroscopy* 2007;23:367-375.
24. Cho N, Lubis M, Ha J, Rhee Y: Clinical results of arthroscopic Bankart repair with knot-tying and knotless suture anchors. *Arthroscopy* 2007;23:566-567.
25. Marquardt B, Garmann S, Schulte T, Witt KA, Steinbeck J, Pötzl W: Outcome after failed traumatic anterior shoulder instability repair with and without surgical revision. *J Shoulder Elbow Surg* 2007;16:742-747.

Appendix 5, (Chapter 3)

Metric Development for an Arthroscopic Bankart Procedure: Assessment of Face and Content Validity

Richard L. Angelo, M.D., Richard K. N. Ryu, M.D., Robert A. Pedowitz, M.D., Ph.D., and Anthony G. Gallagher, Ph.D., D.Sc.

Purpose: To establish the metrics (operational definitions) necessary to characterize a reference arthroscopic Bankart procedure, and to seek consensus from experienced shoulder arthroscopists on the appropriateness of the steps, as well as errors identified. **Methods:** Three experienced arthroscopic shoulder surgeons and an experimental psychologist (comprising the Metrics Group) deconstructed an arthroscopic Bankart procedure. Fourteen full-length videos were analyzed to identify the essential steps and potential errors. Sentinel (i.e., more serious) errors were defined as either (1) potentially jeopardizing the procedure outcome or (2) creating iatrogenic damage to the shoulder. The metrics were stress tested for clarity and the ability to be scored in binary fashion during a video review as either occurring or not occurring. The metrics were subjected to analysis by a panel of 27 experienced arthroscopic shoulder surgeons to obtain face and content validity using a modified Delphi Panel methodology (consensus opinion of experienced surgeons rendered by cyclical deliberations). **Results:** Forty-five steps and 13 phases characterizing an arthroscopic Bankart procedure were identified. Seventy-seven procedural errors were specified, with 20 designated as sentinel errors. The modified Delphi Panel deliberation created the following changes: 2 metrics were deleted, 1 was added, and 5 were modified. Consensus on the resulting Bankart metrics was obtained and face and content validity verified. **Conclusions:** This study confirms that a core group of experienced arthroscopic surgeons is able to perform task deconstruction of an arthroscopic Bankart repair and create unambiguous step and error definitions (metrics) that accurately characterize the essential components of the procedure. Analysis and revision by a larger panel of experienced arthroscopists were able to validate the Bankart metrics. **Clinical Relevance:** The ability to perform task deconstruction and validate the resulting metrics will play a key role in improving surgical skills training and assessing trainee progression toward proficiency.

The intent of any surgical training program, both for residents and for established surgeons acquiring a new procedural skill, is to enable the trainee to acquire the requisite skill sets necessary to perform the designated surgical procedure well and safely. To accomplish this mission, a clearly defined endpoint or set of skill proficiencies must be identified. Furthermore, it must be verified that mastery of those skill sets can accurately be measured during the trainee's progress. It must also be confirmed that the acquisition of those skills is predictive of the ability to perform an effective surgical procedure. Many experienced surgeons who are "proficient" in the performance of a specific procedure and are able to perform it well are also able to identify and agree on the essential "steps" to be completed, as well as the "errors" to be avoided, for that procedure. The reader is referred to Table 1 for a glossary of terms used throughout the article. One challenge, however, in identifying those key features is that surgeons rarely think about the procedures they perform with this level of detail. Surgeons who are proficient in the performance of a specific operation will exhibit many if not all of the important "performance characteristics" (Table 1) that contribute to actually performing the procedure well. They may, however, have automated to many of these steps and how they are performed and, as a consequence, may be less cognizant of the details and more granular elements of the techniques they use.[1-3]

From ProOrtho Clinic (R.L.A.), Kirkland, Washington; The Ryu Hurvitz Orthopedic Clinic (R.K.N.R.), Santa Barbara, California; Professor Emeritus, University of California (R.A.P.), Los Angeles, California, U.S.A.; and University College Cork (A.G.G.), Cork, Ireland.

The authors report the following potential conflict of interest or source of funding: R.L.A. receives support from DePuy Mitek. R.K.N.R. receives support from MedBridge and Mitek.

Received December 26, 2013; accepted April 15, 2015.

Address correspondence to Richard L. Angelo, M.D., ProOrtho Clinic, 12911 120th Ave NE, Ste H-210, Kirkland, WA 98072, U.S.A. E-mail: rlangelo@comcast.net

0749-8063/13907/$36.00
http://dx.doi.org/10.1016/j.arthro.2015.04.093

Table 1. Glossary

	Definition
Behavioral scientist	A professional who engages in any discipline concerned specifically with the subject of human actions and behavior
Concurrent validity	A type of evidence in which there is a positive relation between the test scores from one instrument and the scores from another instrument purporting to measure the same construct
Construct validity	A type of evidence that supports that specific test items identify the quality, ability, or trait they were designed to measure
Content validity	An estimate (by expert/experienced opinion) of the validity of a testing instrument based on a detailed examination of the contents of the test items
Definition	A definite, distinct, and clear objective characterization providing an accurate and reliable identification of whether an event was or was not observed to have occurred
Delphi Panel (modified)	A structured communication technique originally developed as a systematic, interactive forecasting method that relies on the opinion of an experienced panel; in the modified form, the members of the panel answer queries/vote in 2 or more rounds (cycles) on the appropriateness of the metric-based operational definitions of detailed aspects of procedure performance with the goal of achieving consensus—voting is not anonymous
Description	A qualitative characterization of certain or salient aspects or features of an event
Error	A deviation from optimal performance
Face validity	An estimate by an experienced panel that reviews the content of an assessment or tool to see if it seems appropriate and relevant to the concept it purports to measure
Iterative process	A process for calculating or progressing toward a desired result by means of repeated cycles of operations (deliberations); an iterative process should be convergent, that is, it should come closer to the desired result as the number of iterations increases
Likert-like scale	A method of ascribing a quantitative value to qualitative data to make them amenable to statistical analysis
Metric	A standard of measurement of quantitative assessments used for objective evaluations to make comparisons or to track performance
Metric stress testing	A method for determining how specific metric definitions fare during their application and use in scoring in vivo or video-recorded performances
Metric unit	A method of measurement in which the basic parts or components are discrete performance elements
Operational definition	Terms used to define a variable or event in terms of a process (or set of validation tests) needed to determine its existence, quantity, and duration
Performance characteristics	The features determining the accomplishment of a given task measured against preset known standards of accuracy and completeness
Predictive validity	A type of evidence that determines the extent to which the scores on a test are predictive of actual performance
Procedure phase	A group or series of integrally related events or actions that, when combined with other phases, make up or constitute a complete operative procedure
Proficiency/proficient	A specific level of performance defined by a quantitative score (benchmark) or scores on a standardized test or other form of assessment
Reference procedure	A straightforward operative procedure; an agreed on/accepted approach to the performance of an uncomplicated surgical procedure
Reliability of identification (inter-rater reliability)	The extent of agreement between 2 raters on the occurrence of a series of observed events; it ranges between 0, no agreement, and 1.0, complete agreement
Sentinel error	An event or occurrence involving a serious deviation from optimal performance during a procedure that either (1) jeopardizes the success/desired result of the procedure or (2) creates iatrogenic insult to a patient's tissues
Step	A component task, the series aggregate of which constitutes the completion of a specific procedure
Task analysis	An assessment of how a procedure is accomplished, including a detailed (functional) description of the manual activities or tasks along with their duration, frequency, and complexity and any other unique and distinguishing factors
Task deconstruction	To break down a procedure into constituent tasks, steps, or components

The units of performance that constitute skill can be elucidated with a "task analysis" (Table 1), or breakdown and detailed description of the steps or actions necessary to perform the procedure. In attempting to characterize specific skills, psychologists have subjected them to a detailed task analysis and then "operationally defined" (Table 1), rather than simply described, the resulting steps. A definition specifies the order, duration, and result of the specific action and provides precise parameters such that it can be unambiguously determined whether that specific event did or did not occur. A "description" (Table 1), on the other hand, only offers a general characterization of an event or behavior in qualitative terms. Definitions are the preferred foundations of measurement science. The definitions, or "metrics" (Table 1), for a specific procedure provide a quantitative standard of measurement that can be used to objectively assess performance. These metrics must then be validated with respect to whether their characterization fits with what is known about the skill being analyzed. The task analysis–derived characterizations, or "metric units" (Table 1), of skilled performance do not have to capture every aspect of performance but should at least allow for ordinal differentiation between different levels of performance as described by Dreyfus and Dreyfus.[4] The metrics created from this analysis can serve as a tool to evaluate the effectiveness of different training protocols for a particular surgical procedure.

"Face validity" (Table 1) is verified by the opinion of an experienced panel that reviews the content of an assessment or tool to determine if it is appropriate and relevant to the concept it purports to measure. "Content validity" (Table 1) of a testing instrument is similarly obtained and based on the opinion of an experienced panel that performs a detailed examination of the contents of the test items. Thus the face validity and content validity of tools assessing procedural skill are not verified by statistical analysis but, rather, by the summary opinion of an experienced panel of surgeons. An additional question that relates to establishing the validity of the metric definitions for a particular procedure asks, "Do more skilled individuals perform better on the defined metrics than less skilled or experienced individuals, and do the specific metrics identify the quality, ability, or trait they were designed to measure ('construct validity' [Table 1])"? In contrast to face and content validity, the establishment of construct validity requires sufficient data and statistical analysis to prove that it exists.

Task analysis for a particular operation should be performed initially for a "reference procedure" (Table 1)[5-7]—one that is straightforward with a generally accepted or agreed on method that is uncomplicated under ideal circumstances. An optimal approach to learning should ensure that trainees are capable of performing a routine procedure before they have to deal with the technique variations necessary to address more complex pathology.

We sought to study the effectiveness of "proficiency-based progression" (Table 1) training plus simulation for the acquisition of surgical skills. Proficiency-based progression dictates that the learner must demonstrate the ability to meet specific performance benchmarks before he or she is permitted to progress in training. This investigation required the development and validation of specific tools to conduct the analysis. The first component needed was a "metric tool" (Table 1) that could objectively and accurately characterize an arthroscopic Bankart repair. The development of this tool is the focus of this study. Future investigations will report on the establishment of additional tools.

The purpose of this study was to establish the metrics (operational definitions) necessary to characterize a reference arthroscopic Bankart procedure and to seek consensus from experienced shoulder arthroscopists on the appropriateness of the steps, as well as errors identified. The null hypothesis was that face and content validity for the step and error metrics derived from task deconstruction of an arthroscopic Bankart procedure would not be demonstrated.

Methods

Arthroscopic Bankart Metric Development

Three experienced arthroscopic shoulder surgeons (R.L.A., R.K.N.R., R.A.P.), each with over 25 years of clinical practice, and an experimental psychologist (A.G.G.) formed the Metrics Group that characterized an arthroscopic Bankart repair. A detailed task analysis and deconstruction process (described in detail elsewhere)[6] was used to identify the units of performance that are integral to the skilled performance of the instability repair. The goal was to characterize a "reference" arthroscopic Bankart repair and not one attempting to manage unusual or complex instability pathology. Procedure performance characterization (task deconstruction) was guided by (1) decades of practice and teaching experience by the Metrics Group, (2) published studies on arthroscopic Bankart repair,[8-10] and (3) manufacturer guidelines on device use. Two 2.5-day face-to-face meetings and eight 1.5- to 2-hour online conferences were conducted, along with countless E-mail exchanges, to craft the procedural metrics. For the online sessions, the use of Skype videoconferencing (Microsoft, Redmond, WA; available at www.skype.com) enabled the investigators (R.L.A., R.K.N.R., R.A.P., A.G.G.; who reside in different geographic locations) to simultaneously review arthroscopic videos in real time with acceptable resolution. One investigator initiated a standard Skype video connection for a group call using a laptop computer. A second computer (desktop) with a high-resolution screen was used to play the arthroscopic video being studied. An independent USB camera (Ipevo, Sunnyvale, CA) was connected to the USB port of the laptop to which the Skype video input was directed, instead of the resident camera on the laptop screen ("settings" tab in Skype). Thus all of the members on the group Skype call viewed the arthroscopic image rather than the call initiator's image.

Fourteen video recordings of a complete in vivo Bankart procedure, performed by surgeons with varying levels of experience (Table 2), were reviewed by the Metrics Group in detail to assist in the creation and stress testing of the metrics. The videos represented surgeons with practice experience ranging from 3 to 33 years. Both the lateral decubitus ($n = 10$) and

Table 2. Source Videos for Bankart Metric Creation and Stress Testing

Video No.	Surgeon Time in Practice, yr	Patient Orientation
1	25	LD
2	17	LD
3	25	LD
4	26	LD
5	25	LD
6	17	BC
7	18	BC
8	26	LD
9	28	LD
10	21	LD
11	24	LD
12	25	LD
13	3	BC
14	4	BC

BC, beach chair; LD, lateral decubitus.

beach-chair (n = 4) orientations of the patients were represented. All metrics were constructed to be applicable to and able to be scored for surgical procedures performed with patients in both the lateral decubitus and beach-chair orientations. During the series of video reviews, each metric unit was identified and the definition refined so that it could be unambiguously scored as either occurring or not occurring, with a high degree of reliability, by an independent group of raters. Each step was further defined by identifying beginning points and endpoints during the procedure for that metric. The aim was that these detailed metric units would accurately capture the essence of procedure performance, as well as serve as a sound and comprehensive training guide for persons learning the procedure. The metrics included the specific operative steps, the general order in which they should be accomplished, and the instruments and the manner in which they should be used. "Procedural phases" (Table 1) were specified for groups of related steps. In addition to specifying each procedural step, metrics were also created to identify potential "errors" (Table 1), or actions that deviate from optimal performance and should not be done.[11] The intent, again, was to create unambiguous operational definitions (rather than descriptions) for each metric error. A special designation was made for more serious, or "sentinel," errors defined by events that, by themselves, could either (1) jeopardize the outcome of the procedure or (2) lead to significant iatrogenic damage to the shoulder joint. An additional error characterization was termed "damage to non-target tissue." This occurrence defined an event that was injurious to tissues not intentionally being addressed during the defined task, such as "scuffing of articular cartilage by an instrument" or "lacerating the intact labrum."

By agreed on convention, an event (step or error) must have been observed on the video to be scored. Thus inference that an event was "likely to have occurred" was eliminated. For example, if comparable views of the anterior humeral head showed relatively healthy or pristine articular cartilage early in the procedure, with scuffing and abrasion later during the repair, but the injurious event was not observed on the video, it was not scored as an error (or damage to non-target tissue).

Metric Stress Testing and Reliability of Identification

After the 4 members of the Metrics Group were satisfied that the entirety of the procedure had been well characterized, they "stress tested" (Table 1) the metrics by subjecting them to a robust assessment of how reliably they could be independently scored in blinded fashion. Eight video recordings of complete arthroscopic Bankart procedures that were performed by surgeons possessing a wide range of technical skill were independently reviewed and scored. Both the lateral decubitus and beach-chair orientations were represented by the videos studied. Each metric was scored in binary fashion as either yes or no (occurring or not occurring). After each video review, differences in the scoring of each metric by the reviewers (R.L.A., R.K.N.R., R.A.P.) were compared and discussed. Where necessary, operational definitions were clarified, modified, or dropped and new ones added to optimize the functionality of the characterizations as a whole. This process of independent viewing, scoring, and revising the step and error metrics was continued until the Metrics Group was satisfied that the metrics accurately and unambiguously characterized the specifics of an arthroscopic Bankart procedure and could be "reliably identified" (Table 1) by independent reviewers. The extent of agreement between 2 raters for the entire group of step and error Bankart metrics could potentially range between 0, no agreement, and 1.0, complete agreement.

Face and Content Validation of Bankart Metrics by Modified Delphi Panel

The Delphi Panel method[12] (Table 1) is a process that provides an interactive communication structure between researchers (i.e., the Metric Group authors) and an experienced panel (as described later) in a field or discipline to provide systematic feedback on a given topic (i.e., the accuracy of the metrics developed for a reference approach to a Bankart procedure). The Delphi method uses an "iterative process" (Table 1) for progressing toward a desired result by means of repeated cycles of deliberations. The iterative process should be convergent, that is, it should come closer to the desired result as the number of iterations or cycles of review increases. For the Bankart characterization, the desired result (consensus on the appropriateness of a particular metric) was obtained by means of repeated cycles of questioning, deliberation, metric modification, and voting on the appropriateness of each refined metric definition. The methodology assumes that good-quality knowledge evolves from the process. The Delphi method was modified to the extent that the voting cycles, with each new iteration, were not anonymous.

The determination of face and content validity for the Bankart characterization was made by subjecting each metric to an appraisal by a group of surgeons who were very experienced in the performance of an arthroscopic Bankart repair. Twenty-seven board-certified orthopaedic surgeons (the 3 Metrics Group surgeons and 24 additional Arthroscopy Association of North America shoulder faculty instructors) with an average of over 23 years in clinical practice involving shoulder arthroscopy attended a Delphi Panel. Four of the panelists are full-time academicians, 9 are in private group practice and have direct involvement in teaching fellows, and 14 are in private practice with a clinical affiliation with a university orthopaedic department. Each member of the Delphi Panel is a master or associate master faculty member of the Arthroscopy Association of North America and has taught the technique for arthroscopic Bankart repair during shoulder courses conducted at the Orthopedic Learning Center (Rosemont, IL). An experimental psychologist facilitated the meeting.

Delphi Panel Procedure

An overview of the project and meeting objectives was presented. Background information regarding proficiency-based progression training, prior literature showing the validity of this training approach for procedural specialties, and the specific objectives of the current Delphi Panel[12] were reviewed. It was explained that the Bankart metrics had been developed by the Metrics Group for a reference approach to arthroscopic anterior shoulder stabilization for unidirectional anterior glenohumeral instability.[6,7] It was acknowledged that the designated reference procedure might not reflect the exact techniques used by individual panelists but that the operative steps that were presented accurately embodied the essential and key components of the procedure. An affirmative vote by a panel member indicated that the metric definition presented was accurate and acceptable as written but not necessarily that it was the manner in which that particular panelist might have chosen to complete the step. "Consensus" meant that there was unanimity in voting among the panelists and that a particular metric definition was "not wrong or inappropriate." Each of the procedural steps and potential errors were evaluated individually. After each metric definition was presented, panel members voted on whether or not the metric was acceptable as written. If the panel could not achieve consensus because of lack of clarity or differences in opinion, the metric definition was revised accordingly and a new vote conducted on the acceptability of the modified metric. This process was repeated until the metric was accepted. If consensus could not be achieved through a series of modifications, the metric was deleted. When it was deemed necessary, a new metric was defined and added.

Results

Bankart Procedure Metrics

The step metrics resulting from task deconstruction were grouped into 13 separate phases of the procedure (in Roman numerals). Each phase (e.g., "arthroscopic instability assessment" or "inferior anchor preparation/insertion") contains a series of related, unambiguously defined, observable procedure events (steps) with specific beginning and ending points. All potential errors identified had been noted to occur during the stress testing of the metrics. Some of the identical errors and sentinel errors could occur during different phases of the procedure that recurred during the 3-anchor repair (e.g., "uncorrected entanglement of shuttling device or suture").

Modified Delphi Panel

All phases of the procedure were accepted as identified. Only a minority of procedure phases and their associated metrics were accepted without discussion. At the conclusion of the deliberations, consensus among the Delphi Panel was reached for 45 steps, 77 errors (29 unique), and 20 sentinel errors (8 unique) (Tables 3 and 4). During the panel deliberations, 2 metrics were deleted, 1 was added, and 5 were modified before consensus was achieved (Table 5).

Summary of Points Raised and Voting Outcomes of Bankart Delphi Panel

The minutes of the metric validation meeting (Copernicus Study/Delphi Panel), held November 18, 2011, are presented herein (recorded by Robert Pedowitz, M.D., Ph.D.). The meeting chair was Rick Angelo, M.D. The Project Leadership Team comprised Rick Angelo, Rick Ryu, Rob Pedowitz, and Tony Gallagher. The attendees comprised R. Angelo, R. Ryu, R. Pedowitz, J. Tokish, R. Bell, R. Hunter, K. Nord, V. Goradia, A. Barber, S. Snyder, B. Beach, J. Abrams, B. Shaffer, J. Tauro, L. Higgins, S. Weber, S. Koo, D. Richards, J. Esch, J. Dodds, J. Randle, J. Richmond, A. Curtis, J. Burns, N. Sgaglione, J. Kelly, and S. Powell (27 voting attendees), as well as T. Gallagher (meeting facilitator)

The meeting overview is as follows:

1. Dr. Angelo presented a brief overview of the project and meeting objectives.
2. Dr. Gallagher presented the background of proficiency-based training, as well as some prior literature demonstrating the validity of this training

Table 3. Thirteen Phases of Arthroscopic Bankart Procedure (in Roman Numerals) and Brief Summary of 45 Steps of Procedure

I. Portals
- 1. Posterior portal established
- 2. View posterior humeral head and extent of the Hill-Sachs when present
- 3. Introduce mid-anterior spinal needle immediately superior to the subscapularis and direct it toward the anteroinferior glenoid and labrum
- 4. Establish a cannula that abuts the superior border of the subscapularis near the lateral subscapularis insertion
- 5. Demonstrate instrument access to the anteroinferior glenoid/labrum
- 6. Introduce anterosuperior spinal needle at the superolateral aspect of the rotator interval and direct it toward the anterior glenoid
- 7. Establish an anterosuperior cannula, arthroscopic sheath, or switching stick

II. Arthroscopic instability assessment
- View from posterior portal
  - 8. View or probe the superior labral attachment onto the glenoid
  - 9. View or probe articular surface of the rotator cuff
  - 10. Probe anteroinferior glenoid/Bankart pathology including rim fracture, articular defect
- View from anterosuperior portal
  - 11. View or probe the midsubstance of the anterior-inferior glenohumeral ligaments
  - 12. View or probe the insertion of the anterior glenohumeral ligaments onto the anterior humeral neck

III. Capsulolabral mobilization/glenoid preparation
- 13. Elevate the capsulolabral tissue from the glenoid neck and articular margin
- 14. View the subscapularis muscle superficial to the mobilized capsule
- 15. With an instrument, grasp and perform an inferior to superior shift of the capsulolabral tissue (to show tension is restored)
- 16. Obtain a view of the anterior glenoid neck
- 17. Mechanically abrade the glenoid neck

IV. Inferior anchor preparation/insertion
- 18. Seat the guide for the most inferior anchor hole at the inferior region of the anteroinferior quadrant
- 19. Drill anchor hole oblique to the glenoid articular face
- 20. Insert anchor
- 21. Test anchor security by pulling on suture tails

V. Suture delivery/management
- 22. Pass a cannulated suture hook or suture retriever through the capsular tissue inferior to the anchor
- 23. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula

VI. Knot tying
- 24. Deliver an arthroscopic sliding knot
- 25. Back up with 3 or 4 half-hitches
- 26. Cut suture tails

VII. Second anchor preparation/insertion
- 27. Seat the drill guide for the second anchor superior to the first anchor and inferior to the equator of the glenoid
- 28. Drill anchor hole oblique to the glenoid articular face
- 29. Insert suture anchor
- 30. Test anchor security by pulling on suture tails

VIII. Suture delivery/management
- 31. Pass a cannulated suture hook or suture retriever through the capsular tissue inferior to the suture anchor
- 32. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula

IX. Knot tying
- 33. Deliver an arthroscopic sliding knot
- 34. Back up with 3 or 4 half-hitches
- 35. Cut suture tails

X. Third anchor preparation/insertion
- 36. Seat the drill guide for the third anchor at or superior to the equator
- 37. Drill anchor hole oblique to the glenoid articular face
- 38. Insert suture anchor
- 39. Test anchor security by pulling on suture tails

XI. Suture delivery/management
- 40. Pass a cannulated suture hook or suture retriever through the capsular at or inferior to the suture anchor
- 41. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula

XII. Knot tying
- 42. Deliver an arthroscopic sliding knot
- 43. Back up with 3 or 4 half-hitches
- 44. Cut suture tails

XIII. Procedure review
- 45. View and/or probe final completed repair

Table 4. Summary of 29 Different Bankart Procedure Metric Errors

1. Failure to maintain intra-articular position of the posterior cannula
2. Failure to maintain intra-articular position of the mid-anterior cannula
3. Failure to maintain intra-articular position of the anterosuperior cannula
4. Damage to the superior border of the subscapularis during creation of the midanterior portal
5. Damage to the anterior border of the supraspinatus during creation of the anterosuperior portal
6. Loss of intra-articular position of arthroscope/sheath or operating cannula (loss of each portal is scored only once for each Roman numeral, i.e., up to a total of 3 for arthroscope + 2 portals)
7. Laceration of intact capsulolabral tissue (sentinel error)
8. Failure to maintain control of a working instrument (sentinel error)
9. Guide is not located in the inferior region of the anteroinferior quadrant of the glenoid for the most inferior anchor
10. Entry of the completed tunnel lies outside safe zone of 0 to 3 mm from the bony glenoid rim (sentinel error)
11. Shallow undermining and deformation of articular cartilage (sentinel error)
12. Failure to maintain secure seating of the drill guide during anchor insertion
13. Breakage of the implant
14. Implant remains visibly proud (sentinel error)
15. Failure to insert the anchor with the inserter laser line (when present) to or beyond the laser line on the drill guide
16. Anchor fails to remain securely fixed within bone at the appropriate depth
17. Capsular penetration is at or superior to anchor hole (sentinel error)
18. Capsular penetration is not at or peripheral to the capsulolabral junction
19. Instrument breakage
20. Tearing of capsulolabral tissue
21. Uncorrected entanglement of shuttling device or suture
22. Off-loading of suture anchor
23. Breakage of suturing device
24. Failure to create and maintain indentation of the capsule or labral tissue on knot completion (sentinel error)
25. Visible void is present between throws of the completed primary knot (sentinel error)
26. Completed knot abuts articular cartilage
27. Visible void is present between throws of the completed half-hitches
28. Suture breakage
29. Guide is inferior to the equator of the glenoid for the third anchor position

NOTE. Metric errors can be associated with multiple phases and steps of the procedure.

approach for procedural specialties, and he explained the specific objectives of the current Delphi meeting.

3. Dr. Angelo presented each procedural step and explained the associated metrics that have been developed by the Metrics Group for a reference approach to anterior shoulder stabilization for glenohumeral instability. The comments and recommendations for each of the steps, with associated vote, are presented in Table 6.

Discussion

The principal findings of this study are that (1) an arthroscopic Bankart procedure can be deconstructed

Table 5. Delphi Panel Metric Changes

Modification	Issue	Deliberation	Resolution
Deleted (2)	Should failure of anchor purchase remain an error?	Cadaveric bone variability too great to score accurately	Delete error
	Alternating posts for knot tying	Arthroscope views not consistent enough to score reliably	Delete error
Added (1)	Completed knot position	May cause iatrogenic damage if it abuts articular cartilage	Add error to each of the 3 knots tied
Modified (5)	Diagnostic steps—probe or view (how long?)	Does "looking" equal "ascertaining"?	View must be held long enough to determine pathology
	Adequacy of capsular mobilization	Should demonstrate effort at capsular mobility	Must take instrument (grasper) to demonstrate
	Whether to ascribe "critical" to "laceration of labrum"	Labral variability too great in cadavers; can still see violation of "hoop" integrity	Retain as a sentinel error
	Consider deleting the term "sliding" from knot description	Sliding knot would be acceptable for a reference procedure	Retain the term "sliding"
	Should diagnostic steps be included in the procedure metrics?	May skew results if there is excessive influence of diagnostic steps in procedure	Include only steps directly related to instability assessment

Table 6. Comments, Recommendations, and Associated Vote

Comments and Recommendations Regarding Procedural Steps and Metrics	Vote on Steps and Metrics
I. Portals (steps 1-7)	Unanimous affirmative
Agreement that this is an outside-in reference approach for portal placement, though some surgeons use an inside-out approach	
Importance of pre-surgery setup, though assessment of this phase would be very difficult using arthroscopic videotapes	
II. Diagnostic arthroscopy (steps 8-12)	Unanimous affirmative
Clarified metric "view or probe," not "view and probe"	
Discussion about whether we should include metrics for view or probe of posterior labrum, superior labrum, biceps, and rotator cuff	
Recommendation: Limit diagnostic elements for the current procedural assessment, and consider creation of a diagnostic arthroscopy reference procedure (vote to drop diagnostic elements failed—because one component of the metrics is teaching essential components of the procedure)	
III. Capsulolabral mobilization (steps 13-15)	Unanimous affirmative
Should a laceration of the labrum be defined as a "critical error"?	
Should the pre-existing tissue quality of the labrum be assessed so that laceration of poor tissue does not qualify as a critical error?	
Consider adjustment of metric definition to describe grasping of the anatomic structure of the anterior inferior glenohumeral ligament	
IV. Glenoid neck preparation (steps 16-17)	Unanimous affirmative
V. Insertion of first anchor (steps 18-21)	Unanimous affirmative
Failure of anchor from bone should not be considered a critical error because loss of fixation could be related to bone quality	
VI. Suture management—first anchor (steps 22-23)	Unanimous affirmative
Should we include retrieval of broken suturing device (generally thought this would be quite rare, so not a useful metric)?	
The definition for adequacy of capsulolabral tissue capture seems adequate	
VII. Knot tying—first anchor (steps 24-26)	Unanimous affirmative
Consider deleting the term "sliding" from the knot description	
Add error of a knot completed and left on the articular surface	
Need to drop "alternating posts" metric	
VIII. Insertion of second anchor (steps 27-30)	Unanimous affirmative
Failure of anchor from bone should not be considered a critical error because loss of fixation could be related to bone quality	
IX. Suture management—second anchor (steps 31-32)	Unanimous affirmative
X. Knot tying—second anchor (steps 33-35)	Unanimous affirmative
Consider deleting the term "sliding" from the knot description	
Add error of a knot completed and left on the articular surface	
Probably need to drop "alternating posts" metric	
XI. Insertion of third anchor (steps 36-39)	Unanimous affirmative
Failure of anchor from bone should not be considered a critical error because loss of fixation could be related to bone quality	
XII. Suture management—third anchor (steps 40-41)	Unanimous affirmative
XIII. Knot tying—third anchor (steps 42-44)	Unanimous affirmative
Consider deleting the term "sliding" from the knot description	
Add error of a knot completed and left on the articular surface	
Need to drop "alternating posts" metric	
XIV. Final assessment (step 45)	Unanimous affirmative

into the essential steps necessary for the effective completion of the repair, (2) the potential errors related to the procedure are able to be identified and characterized, and (3) face and content validity for the resulting step and error metrics can be obtained through use of the modified Delphi Panel technique. Traditionally, surgeons have been trained using the "apprenticeship" model, which is related to "process" or time, based (a certain variety of rotations, exposure to numbers of specific cases, etc.). A paradigm shift toward proficiency-based progression training, which is outcomes based, is occurring and mandates that the trainee be able to demonstrate the ability to meet specific skills benchmarks to progress in training. These benchmarks must have specific, clear, objective, and fair standards of performance. Validated metrics will be essential in defining these standards. In addition, as the move toward including surgical skills credentialing and procedural competency occurs for licensing, the same validated standards will be needed. The methodology used in this study provides a framework for the development of those metrics and standards.

An arthroscopic Bankart repair (index procedure) was selected as the reference surgical procedure to study for several reasons. For the patient with unidirectional anterior instability due primarily to a Bankart lesion without significant bone loss, a suture anchor repair with 3 implants is a commonly accepted method used to obtain a successful patient outcome.[13-16] In addition, the essential components of the procedure are well outlined regardless of whether the patient is placed in the lateral decubitus or beach-chair orientation.[17-21]

The task analysis stage of metric development is crucial because metrics are the fundamental building blocks of a good training program. Metrics, thus, not only define how the training should be characterized and the procedure performed by the trainee but also must afford the opportunity for meaningful assessment of the trainee's performance and progress. The entire process of metric development should be as transparent, objective, and unambiguous as possible. Metric definitions should be characterized in such a way that they are sufficiently complete and detailed for an individual—not associated with the initial development process—to use them to score performance reliably. Metric definitions should include behavioral markers that indicate the beginning points and endpoints of the performance characteristics (steps) to be assessed. These parameters will become particularly important in the future as the procedural metrics are used with higher-fidelity simulators. The details of the metric definitions will be necessary for the simulator to be appropriately programmed to provide the trainee with performance assessments and accurate feedback.

Other approaches to the measurement of surgical performance use qualitative descriptions of performance and require the user to rate items on a graduated Likert-type scale (Table 1), which ascribes a quantitative value to qualitative data to make them amenable to statistical analysis. Likert scales (often with a range from 1 to 5 or from 1 to 7) are typically constructed with responses (opinion) around a neutral option (e.g., "suture delivery was 1, awkward, . . . 3, effective, . . . 5, highly efficient") and were originally designed to assess a range of attitudes. Because of the component of subjectivity, this method of attempting to rate objective performance can render it difficult to obtain acceptable levels of inter-rater reliability (>80%) in the scoring of events. In contrast, the approach to the assessment of performance in our study uses precise definitions of performance and simply requires the reviewer to report whether the specific event occurred or not. This binary approach to the measurement of individual events has been shown to facilitate the reliable scoring of metric-based performance units across a variety of functions during skills training[22-25] of individuals with different experience levels.[26,27] This approach has also been shown to be more reliable than Likert-scale scoring.[28]

Behaviors that deviate from optimal performance (errors) can be characterized, including those of a "more serious nature." The issue of whether those more serious errors should be termed "critical errors" or some alternative label was raised at the outset during the metric definition process. It was agreed on by the Metrics Group that use of the term "critical error" could imply that the event was life-threatening or might have serious medicolegal implications. It was elected, instead, to use the term "sentinel" (Table 1) to connote an error that should be carefully "watched for and to avoid." Sentinel errors involve a serious deviation from optimal performance during a procedure because they can either jeopardize the success/desired result of the procedure or create iatrogenic insult to the patient's tissues. A single specific sentinel error may not always lead to a poor outcome but should stringently be avoided.[6] The underlying philosophy of this approach to errors is that suboptimal outcomes do not happen by accident but usually result from the coalescence of deviations from optimal procedure performance.

The face and content validity of the metric-based procedure characterization by subject specialists can be verified using the modified Delphi Panel methodology reported in this study. The metrics developed were informed by research studies, professional guidelines, clinical experience, and manufacturers' guidelines.[6,7] Although the surgeons in the Metrics Group are very experienced in the performance of the Bankart procedure, the Delphi process provided an excellent method to ensure that the procedure characterization is appropriate, represents best practice, and is acceptable to a larger group of experienced master and associate master Bankart faculty. As anticipated, many surgeons pointed out that they might perform a specific step in a different manner but that the approach outlined by the Metrics Group was "not incorrect" or inadvisable. The members of the panel made very helpful suggestions for improving the definitions.

Assuming that the Bankart metric identifications and definitions represent a real-world surgical procedure, these performance characteristics should be able to distinguish between experienced (skilled) surgeons and novices, that is, provide construct validity. Future studies regarding construct validity will seek to provide information about which metrics best distinguish between experienced and novice surgeon performance. This information will facilitate the establishment of a benchmark to define the "proficiency level" that trainees should acquire before progressing to in vivo practice.[6,29]

Limitations

A limitation of this study resides in the fact that every potential error, regardless of how rare the occurrence, might not have been included. The Delphi Panel, however, confirmed that the errors listed were those most likely to occur and that should be avoided in the safe performance of an arthroscopic Bankart repair. Although common errors may be relatively easy to agree on, it is somewhat more challenging to decide which errors should be designated as "sentinel," without a specific weighting methodology. Although the issue of using this designation for events that cause iatrogenic damage is more straightforward, the concept of also using the term for events that might "potentially lead to a suboptimal outcome" is more subject to the opinion of the Metrics Group and the Delphi Panel.

Furthermore, data are not available to confirm that the specific steps identified by the Metrics Group and the Delphi Panel directly correlate with a successful surgical outcome for patients with unidirectional shoulder instability. Therefore the metrics created remain predominantly based on the opinion of experienced surgeons and instructors. An outcomes study will be needed to fully establish the predictive validity for the Bankart metrics as authored. In addition, the surgeons comprising the Metrics Group and Delphi Panel were all North American surgeons. International arthroscopists may have created somewhat different metrics for an arthroscopic Bankart repair.

Conclusions

This study rejects the null hypothesis and confirms that a core group of experienced arthroscopic surgeons is able to perform task deconstruction of an arthroscopic Bankart repair and create unambiguous step and error definitions (metrics) that accurately characterize the essential components of the procedure. Analysis and revision by a larger panel of experienced arthroscopists were able to validate the Bankart metrics.

References

1. Adams JA. A closed-loop theory of motor learning. *J Mot Behav* 1971;3:111-149.
2. Adams JA. Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychol Bull* 1987;101:41-74.
3. Adams JA. The changing face of motor learning. *Hum Mov Sci* 1990;9:209-220.
4. Dreyfus HL, Dreyfus SE. *Mind over machine: The power of human intuition and expertise in the era of the computer*. New York: The Free Press, 1986;50.
5. Andersen DK. How can educators use simulation applications to teach and assess surgical judgment? *Acad Med* 2012;87:934-941.
6. Gallagher AG, O'Sullivan GC. *Fundamentals of surgical simulation; principles & practices*. London: Springer Verlag, 2011.
7. Gallagher AG. Metric-based simulation training to proficiency in medical education: What it is and how to do it. *Ulster Med J* 2012;81:107-113.
8. Burks RT, Presson AP, Weng H-Y. An analysis of technical aspects of the arthroscopic Bankart procedure as performed in the United States. *Arthroscopy* 2014;30:1246-1253.
9. Morgan CD, Bodenstab AB. Arthroscopic Bankart suture repair: Technique and early results. *Arthroscopy* 2010;26:819-820.
10. Streubel PN, Krych AJ, Simone JP, et al. Anterior glenohumeral instability: A pathology-based surgical treatment strategy. *J Am Acad Orthop Surg* 2014;22:283-294.
11. Seymour NE, Gallagher AG, Roman SA, et al. Analysis of errors in laparoscopic surgical procedures. *Surg Endosc* 2004;18:592-595.
12. Cuschieri A, Francis N, Crosby J, Hanna GB. What do master surgeons think of surgical competence and revalidation? *Am J Surg* 2001;182:110-116.
13. Ee GW, Mohamed S, Tan AH. Long term results of arthroscopic Bankart repair for traumatic anterior shoulder instability. *J Orthop Surg Res* 2011;6:28.
14. van der Linde JA, van Kampen DA, Willems JW. How should we define failure after surgical shoulder stabilization? *Arthroscopy* 2013;29:1602-1603.
15. Waterman BR, Burns TC, McCriskin B, et al. Outcomes after Bankart repair in a military population: Predictors for surgical revision and long-term disability. *Arthroscopy* 2014;30:172-177.
16. Shibata H, Gotoh M, Mitsui Y, et al. Risk factors for shoulder re-dislocation after arthroscopic Bankart repair. *J Orthop Surg Res* 2014;9:53.
17. Angelo R. Arthroscopic Bankart repair for unidirectional shoulder instability. *Instr Course Lect* 2009;58:305-313.
18. Angelo RL. Controversies in arthroscopic shoulder surgery: Arthroscopic versus open Bankart repair, thermal treatment of capsular tissue, acromioplasties—Are they necessary? *Arthroscopy* 2003;19:224-228.
19. Ryu RK. Arthroscopic approach to traumatic anterior shoulder instability. *Arthroscopy* 2003;19:94-101.
20. Kim S-H, Ha K-I, Cho Y-B, et al. Arthroscopic anterior stabilization of the shoulder: Two to six-year follow-up. *J Bone Joint Surg Am* 2003;85:1511-1518.
21. Cole BJ, L'Insalata J, Irrgang J, Warner JJ. Comparison of arthroscopic and open anterior shoulder stabilization. A two to six-year follow-up study. *J Bone Joint Surg Am* 2000;82:1108-1114.
22. McClusky DA, Gallagher A, Ritter EM, et al. Virtual reality training improves junior residents' operating room performance: Results of a prospective, randomized, double-blinded study of the complete laparoscopic cholecystectomy. *J Am Coll Surg* 2004;199:73 (abstract).
23. Van Sickle K, Smith B, McClusky DA III, et al. Evaluation of a tensiometer to provide objective feedback in knot-tying performance. *Am Surg* 2005;71:1018-1023.
24. Van Sickle K, Gallagher AG, Smith CD. The effect of escalating feedback on the acquisition of psychomotor skills for laparoscopy. *Surg Endosc* 2007;21:220-224.

25. Van Sickle K, Ritter EM, Baghai M, et al. Prospective, randomized, double-blind trial of curriculum-based training for intracorporeal suturing and knot tying. *J Am Coll Surg* 2008;207:560-568.
26. Neary PC, Boyle E, Delaney CP, et al. Construct validation of a novel hybrid virtual-reality simulator for training and assessing laparoscopic colectomy; results from the first course for experienced senior laparoscopic surgeons. *Surg Endosc* 2008;22:2301-2309.
27. Nicholson WJ, Cates CU, Patel AD, et al. Face and content validation of virtual reality simulation for carotid angiography: Results from the first 100 physicians attending the Emory NeuroAnatomy Carotid Training (ENACT) program. *Simul Healthc* 2006;1:147-150.
28. Gallagher AG, O'Sullivan GC, Leonard G, Bunting BP, McGlade KJ. Objective structured assessment of technical skills and checklist scales reliability compared for high stakes assessments. *ANZ J Surg* 2012;84:568-573.
29. Gallagher AG, Ritter EM, Champion H, et al. Virtual reality simulation for the operating room: Proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg* 2005;241:364-372.

Appendix 6, (Chapter 4)

The Bankart Performance Metrics Combined With a Shoulder Model Simulator Create a Precise and Accurate Training Tool for Measuring Surgeon Skill

Richard L. Angelo, M.D., Robert A. Pedowitz, M.D., Ph.D., Richard K. N. Ryu, M.D., and Anthony G. Gallagher, Ph.D., D.Sc.

Purpose: To determine if a dry shoulder model simulator coupled with previously validated performance metrics for an arthroscopic Bankart repair (ABR) would be a valid tool with the ability to discriminate between the performance of experienced and novice surgeons, and to establish a proficiency benchmark for an ABR using a model simulator. **Methods:** We compared an experienced group of arthroscopic shoulder surgeons (Arthroscopy Association of North America faculty) ($n = 12$) with a novice group ($n = 7$) (postgraduate year 4 or 5 orthopaedic residents). All surgeons were instructed to perform a diagnostic arthroscopy and a 3 suture anchor Bankart repair on a dry shoulder model. Each procedure was videotaped in its entirety and scored in blinded fashion independently by 2 trained reviewers. Scoring used previously validated metrics for an ABR and included steps, errors, and "sentinel" (more serious) errors. **Results:** The inter-rater reliability among pairs of raters averaged 0.93. The experienced group made 63% fewer errors, committed 79% fewer sentinel errors, and performed the procedure in 42% less time than the novice group (all significant differences). The greatest difference in errors between the groups involved anchor preparation and insertion, suture delivery and management, and knot tying. **Conclusions:** The tool comprised by validated ABR metrics coupled with a dry shoulder model simulator is able to accurately distinguish between the performance of experienced and novice orthopaedic surgeons. A performance benchmark based on the mean performance of the experienced group includes completion of a 3 anchor Bankart repair, enacting no more than 4 total errors and 1 sentinel error. **Clinical Relevance:** The combination of performance metrics and an arthroscopic shoulder model simulator can be used to improve the effectiveness of surgical skills training for an ABR. The methodology used may serve as a template for outcomes-based procedural skills training in general.

Some authors from professional bodies and health care training organizations around the world argue that the surgical trainee should acquire basic procedural skills outside of the surgical theater before operating on real patients.[1-3] Furthermore, evidence now clearly indicates that when performed to a quantitatively defined level, skills practiced and acquired outside the operating room are superior to skills acquired in a traditional apprenticeship manner primarily in the operating room.[4,5] Satava[6] first introduced the concept of simulation-based training in the early 1990s, with quantitative evidence from prospective, randomized, double-blinded clinical studies showing that simulation-based training is a powerful tool for the acquisition of surgical skills.[7-10] The simulator can be either a physical model or computer-generated video images[7,10] because both are equally effective if used as part of a "metric-based training curriculum."[11] Table 1 shows a glossary of terms used throughout this article.

An implicit assumption in a simulator-based training process is the use of validated metrics that appropriately characterize the procedure to be trained. Previously, Angelo et al.[12] reported on the development of a tool defining "performance metrics" ("steps" and "errors" as defined in Table 1) for a standard "reference approach"

From ProOrtho Clinic (R.L.A.), Kirkland, Washington; Professor Emeritus, University of California (R.A.P.), Los Angeles, California; The Ryu Hurvitz Orthopedic Clinic (R.K.N.R.), Santa Barbara, California, U.S.A.; and ASSERT, University College Cork (A.G.G.), Cork, Ireland.

The authors report the following potential conflict of interest or source of funding: R.L.A. receives support from DePuy Mitek. R.K.N.R. receives support from MedBridge and Mitek.

Received January 24, 2014; accepted April 15, 2015.

Address correspondence to Richard L. Angelo, M.D., ProOrtho Clinic, 12911 120th Ave NE, Ste H-210, Kirkland, WA 98072, U.S.A. E-mail: rlamdortho@comcast.net

0749-8063/1454/$36.00

http://dx.doi.org/10.1016/j.arthro.2015.04.092

Arthroscopy: The Journal of Arthroscopic and Related Surgery, Vol 31, No 9 (September), 2015: pp 1639-1654 1639

Table 1. Glossary

	Definition
Construct validity	A type of evidence that supports that specific test items identify the quality, ability, or trait they were designed to measure
Content validity	An estimate (opinion) by experts of the validity of a testing instrument based on a detailed examination of the contents of the test items
Definition	A definite, distinct, and clear objective characterization providing an accurate and reliable identification of whether an event was or was not observed to have occurred
Damage to non-target tissue	Iatrogenic damage to tissues not intended to be addressed in the specific step (i.e., articular cartilage damage)
Delphi Panel (modified)	A structured communication technique originally developed as a systematic, interactive forecasting method that relies on the opinion of a panel of experts; in the modified form, experts answer queries/vote in 2 or more rounds (cycles) on the appropriateness of the metric-based operational definitions of detailed aspects of procedure performance with the goal of achieving consensus—voting is not anonymous
Error	A deviation from optimal performance
Face validity	An estimate (opinion) by experts who review the content of an assessment or tool to see if it seems appropriate and relevant to the concept it purports to measure
Inter-rater reliability	The extent of agreement between 2 raters on the occurrence of a series of observed events; it ranges between 0, no agreement, and 1.0, complete agreement
Metric	A standard of measurement of quantitative assessments used for objective evaluations to make comparisons or to track performance
Metric-based training curriculum	A procedural skills training program based on clear, specific, detailed definitions of the steps to be accomplished and errors avoided
Performance metric	The features determining the accomplishment of a given task measured against preset known standards of accuracy and completeness
Procedure phase	A group or series of integrally related events or actions that, when combined with other phases, make up or constitute a complete operative procedure
Proficiency/proficient	A specific level of performance defined by a quantitative score (benchmark) or scores on a standardized test or other form of assessment
Reference procedure	A straightforward operative procedure; an agreed on/accepted approach to the performance of an uncomplicated procedure under ideal circumstances
Sentinel error	An event or occurrence involving a serious deviation from optimal performance during a procedure that either (1) jeopardizes the success/desired result of the procedure or (2) creates iatrogenic insult to the patient's tissues
Step	A component task, the series aggregate of which constitutes the completion of a specific procedure
Task deconstruction	To break down a procedure into constituent tasks, steps, or components

to performing an arthroscopic Bankart repair.[13-17] This tool was derived from a careful "task deconstruction" (Table 1) using videos of complete Bankart procedures performed with patients in either the lateral decubitus or beach-chair orientation. The metrics were constructed so that they could be scored in an identical manner with the patient in either orientation. "Face validity" and "content validity" of the metrics were verified using a "modified Delphi Panel" methodology (Table 1).

The purpose of this study was to determine if a dry shoulder model simulator coupled with previously validated performance metrics for an arthroscopic Bankart repair would be a valid tool with the ability to discriminate between the performance of experienced and novice surgeons. We also sought to establish a "proficiency" (Table 1) benchmark for the arthroscopic Bankart procedure using the model simulator. The null hypothesis was that when using a shoulder model simulator, the Bankart metrics would fail to discriminate between experienced and novice surgeon performance.

Methods

No institutional review board (IRB) approval was obtained for this study investigating the validity of the Bankart metrics coupled with the model simulator. IRB approval was sought for the final Copernicus Study proper, which will compare 3 different training protocols. The Western IRB (No. 1-776362-1) opined that, as an educational curriculum study, this investigation was exempt from the need for full IRB approval [based on the criteria of 45 CFR 46.101(b)(1)]. The final study comparing the 3 training protocols was registered with the National Institutes of Health (ClinicalTrials.gov No. NCT01921621).

Study Groups

Two groups were compared in their performance of an arthroscopic Bankart procedure on a shoulder model simulator. The experienced group consisted of all faculty members who served as either a master or associate master instructor for a standard 3-day Arthroscopy Association of North America (AANA) Resident Course conducted at the Orthopedic

Learning Center (Rosemont, IL). The novice group was limited to postgraduate year (PGY) 4 and PGY 5 orthopaedic residents who had registered for a Resident Course and who volunteered to participate in the investigation.

Arthroscopic Bankart Repair Metrics

Metrics have been previously defined for a standard reference arthroscopic Bankart repair.[12] Forty-five essential steps in 13 "phases" (Table 1) (Roman numerals) were defined with beginning points and endpoints (Table 2). Twenty-nine potential unique errors were specified (Table 3), 8 of which were designated as "sentinel" (Table 1). The more serious (sentinel) errors were defined as those expected to either (1) substantially compromise the outcome of the shoulder stabilization (e.g., "capsular penetration of the suture passing instrument is superior to the anchor hole," resulting in failure to achieve retentioning of the capsule and inferior glenohumeral ligaments) or (2) potentially lead to iatrogenic damage to the shoulder (e.g., "laceration of the intact labrum"). Some of the same errors could be enacted more than once during different phases of the procedure. Thus a total of 77 potential errors, 20 of which were sentinel errors, were specified for the complete procedure. In addition, events that led to less consequential "damage to non-target tissue" (DNTT) (Table 1) were recorded as a standard error (e.g., scuffing of the articular cartilage). A perfect score would indicate that all 45 steps were completed satisfactorily without committing any errors.

Dry Shoulder Model Simulator

The shoulder simulator used is a physical model composed of a dense foam plastic endoskeleton including a humerus, scapula, glenoid, coracoid, acromial spine, and acromion with proportions appropriate to the human skeleton (Sawbones; Pacific Research Laboratories, Vashon, WA) (Fig 1). The articulating surfaces of the humerus and glenoid are laminated with a softer, white layer designed to mimic articular cartilage. A Hill-Sachs lesion measuring 1 cm by 3.5 cm is oriented vertically on the posterior aspect of the humeral head and is represented by a red impaction trough. A rim of off-white, rubber-like material encircles and lightly adheres to the glenoid neck, simulating the labrum. Red staining in the region where the labrum is joined to the anteroinferior glenoid represents the Bankart lesion. The adhesive attachment of the labrum requires the operator to intentionally "liberate" the labrum from the glenoid to demonstrate mobilization of the capsulolabral tissues. A more medial and superficial pink layer of soft foam represents the subscapularis muscle. A tubular strand of rubber simulates the long head of the biceps tendon and courses from its anatomic attachment to the superior labrum, out of the shoulder joint, into the bicipital groove of the humerus. The capsule is replicated by a pliable, rubberized material containing the glenohumeral joint and has a molded imprint of the inferior glenohumeral ligaments on the articular surface. A separate band represents the superior border of the subscapularis tendon. Holes measuring 8 mm in diameter are created in the capsule during molding and enable cannulas for the posterior, mid-anterior, and anterosuperior portals to pass through the relatively tough capsular material. Beige-colored, soft, moderately dense foam represents the skin and soft tissues exterior to the glenohumeral joint and possesses a contour and bulk that mimic the shape of the human shoulder. The acromion, acromial spine, and coracoid landmarks are readily palpable through the "soft tissues" and assist in locating proper portal placement.

Arthroscopic Bankart Repair

During a single weekend AANA resident arthroscopy course, the surgeons from both groups were instructed to establish portals (posterior, anterosuperior, and mid-anterior), complete a thorough diagnostic arthroscopy, and perform a 3-anchor arthroscopic Bankart repair on the simulator model. Furthermore, they were instructed to demonstrate/complete all of the steps for the Bankart repair that they would normally perform in clinical practice on a real patient. The model was secured in either the lateral decubitus or beach-chair orientation according to surgeon preference. Equipment representatives from multiple different vendors served as surgical assistants and were randomly assigned to participant surgeons. The assistants were instructed to act only at the specific direction of the operating surgeon. Prompting and coaching (of technique) were prohibited (the procedures were proctored by staff from the Orthopedic Learning Center). A standard equipment tower with a 30° arthroscope, along with all instruments necessary to complete an arthroscopic Bankart repair, was provided (Table 4).

The surgeon created the required portals based on the palpable "bony" landmarks of the shoulder and then progressed to complete the diagnostic arthroscopy and Bankart repair. A continuous video recording was made beginning with the first arthroscopic view of the joint from the posterior portal and ending with the withdrawal of the arthroscope after the surgeon's examination of the completed repair with a hook probe. No time limit was imposed on the performance of the procedure on the simulator model.

Video Reviewer Training

Once the construction of the metrics for an arthroscopic Bankart repair was completed and face and

Table 2. Thirteen Phases of Bankart Procedure (in Roman Numerals) and Brief Summary of 45 Steps of Procedure

I. Portals
  1. Posterior portal established
  2. View posterior humeral head and extent of the Hill-Sachs when present
  3. Introduce mid-anterior spinal needle immediately superior to the subscapularis and direct it toward the anteroinferior glenoid and labrum
  4. Establish a cannula that abuts the superior border of the subscapularis near the lateral subscapularis insertion
  5. Demonstrate instrument access to the anteroinferior glenoid/labrum
  6. Introduce anterosuperior spinal needle at the superolateral aspect of the rotator interval and direct it toward the anterior glenoid
  7. Establish an anterosuperior cannula, arthroscopic sheath, or switching stick

II. Arthroscopic instability assessment
  View from posterior portal
    8. View or probe the superior labral attachment onto the glenoid
    9. View or probe articular surface of the rotator cuff
    10. Probe anteroinferior glenoid/Bankart pathology including rim fracture, articular defect
  View from anterosuperior portal
    11. View or probe the mid-substance of the anterior-inferior glenohumeral ligaments
    12. View or probe the insertion of the anterior glenohumeral ligaments onto the anterior humeral neck

III. Capsulolabral mobilization/glenoid preparation
  13. Elevate the capsulolabral tissue from the glenoid neck and articular margin
  14. View the subscapularis muscle superficial to the mobilized capsule
  15. With an instrument, grasp and perform an inferior to superior shift of the capsulolabral tissue (to show tension is restored)
  16. Obtain a view of the anterior glenoid neck
  17. Mechanically abrade the glenoid neck

IV. Inferior anchor preparation/insertion
  18. Seat the guide for the most inferior anchor hole at the inferior region of the anteroinferior quadrant
  19. Drill anchor hole oblique to the glenoid articular face
  20. Insert anchor
  21. Test the anchor security by pulling on the suture tails

V. Suture delivery/management
  22. Pass a cannulated suture hook or suture retriever through the capsular tissue—inferior to the anchor
  23. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula

VI. Knot tying
  24. Deliver an arthroscopic sliding knot
  25. Back up with 3 or 4 half-hitches
  26. Cut suture tails

VII. Second anchor preparation/insertion
  27. Seat the drill guide for the second anchor superior to the first anchor and inferior to the equator of the glenoid
  28. Drill anchor hole oblique to the glenoid articular face
  29. Insert suture anchor
  30. Test anchor security by pulling on the suture tails

VIII. Suture delivery/management
  31. Pass a cannulated suture hook or suture retriever through the capsular tissue inferior to the suture anchor
  32. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula

IX. Knot tying
  33. Deliver an arthroscopic sliding knot
  34. Back up with 3 or 4 half-hitches
  35. Cut suture tails

X. Third anchor preparation/insertion
  36. Seat the drill guide for the third anchor at or superior to the equator
  37. Drill anchor hole oblique to the glenoid articular face
  38. Insert suture anchor
  39. Test anchor security by pulling on suture tails

XI. Suture delivery/management
  40. Pass a cannulated suture hook or suture retriever through the capsular at or inferior to the suture anchor
  41. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula

XII. Knot tying
  42. Deliver an arthroscopic sliding knot
  43. Back up with 3 or 4 half-hitches
  44. Cut suture tails

XIII. Procedure review
  45. View and/or probe final completed repair

Table 3. Summary of 29 Different Bankart Procedure Metric Errors

Failure to maintain intra-articular position of the posterior cannula
Failure to maintain intra-articular position of the mid-anterior cannula
Failure to maintain intra-articular position of the anterosuperior cannula
Damage to the superior border of the subscapularis during creation of the mid-anterior portal
Damage to the anterior border of the supraspinatus during creation of the anterosuperior portal
Loss of intra-articular position of arthroscope/sheath or operating cannula (loss of each portal is scored only once for each Roman numeral, i.e., up to a total of 3 for arthroscope + 2 portals)
Laceration of intact capsulolabral tissue (sentinel error)
Failure to maintain control of a working instrument (sentinel error)
Guide is not located in the inferior region of the anteroinferior quadrant of the glenoid for the most inferior anchor
Entry of the completed tunnel lies outside safe zone of 0 to 3 mm from the bony glenoid rim (sentinel error)
Shallow undermining and deformation of articular cartilage (sentinel error)
Failure to maintain secure seating of the drill guide during anchor insertion
Breakage of the implant
Implant remains visibly proud (sentinel error)
Failure to insert the anchor with the inserter laser line (when present) to or beyond the laser line on the drill guide
Anchor fails to remain securely fixed within bone at the appropriate depth
Capsular penetration is at or superior to anchor hole (sentinel error)
Capsular penetration is not at or peripheral to the capsulolabral junction
Instrument breakage
Tearing of capsulolabral tissue
Uncorrected entanglement of shuttling device or suture
Off-loading of suture anchor
Breakage of suturing device
Failure to create and maintain indentation of the capsule or labral tissue on knot completion (sentinel error)
Visible void is present between throws of the completed primary knot (sentinel error)
Completed knot abuts articular cartilage
Visible void is present between throws of the completed half-hitches
Suture breakage
Guide is inferior to the equator of the glenoid (for the third and final anchor)

NOTE. Metric errors can be associated with multiple phases and steps of the procedure (77 total errors).

content validity were verified,[12] a final version of a score sheet was formatted. Ten AANA master/associate master faculty surgeons (none belonging to the experienced group from this study) formed the panel of reviewers designated to score the videos. This group included the 3 members (R.L.A., R.K.N.R., R.A.P.) who developed the arthroscopic Bankart metric definitions (Table 1) in conjunction with a consultant experimental psychologist (A.G.G.). The 10 reviewers were assigned by the AANA research coordinator to 1 of 5 fixed pairs, which remained constant throughout the scoring of all videos. Assignments were made based on similar time zones of the reviewers' residence/practice. Reviewer training was initiated with an 8-hour in-person meeting during which time each metric was studied in detail. Multiple video examples of live patient cases were shown to illustrate each particular metric. Videos of the patients in both the lateral decubitus and beach-chair orientations were represented. Discussion helped to clarify how each step and error were to be scored, including the nuances and conventions to be used. Several weeks later, full-length practice videos 1 and 2 (one each in the lateral decubitus and beach-chair orientation) were sent to and independently scored by each of the 10 reviewers, and the scores were then tabulated. During 2 subsequent 2-hour group phone conferences, the differences and discrepancies among all reviewers were compared and discussed, seeking conformity in scoring. Each of the designated pairs of reviewers also participated in 1 to 3 additional phone conferences to analyze the specific instances in which the 2 of them scored particular events differently. Subsequently, all reviewers scored practice videos 3 and 4, and the results were tabulated (each patient orientation was again represented). The scores for each of the 5 designated pairs of reviewers were compared for the second set of practice videos. In only 1 of 10 comparisons (2 videos × 5 reviewer pairs) did the inter-rater reliability (IRR) (Table 1) calculation (as described later) fall below an acceptable level of 0.8,[18] with an IRR value of 0.76.

Video Scoring

The AANA research coordinator randomly assigned each of the 19 full-length study videos of experienced and novice surgeons performing an arthroscopic Bankart procedure on the shoulder simulator model to a single pair of reviewers. Other than the research coordinator and the study consultant, all reviewers remained blinded to the source of all videos. Each of the 19 videos were independently reviewed and scored by the 2 members of an assigned pair of reviewers. All scores were tabulated for each of 13 phases of the procedure

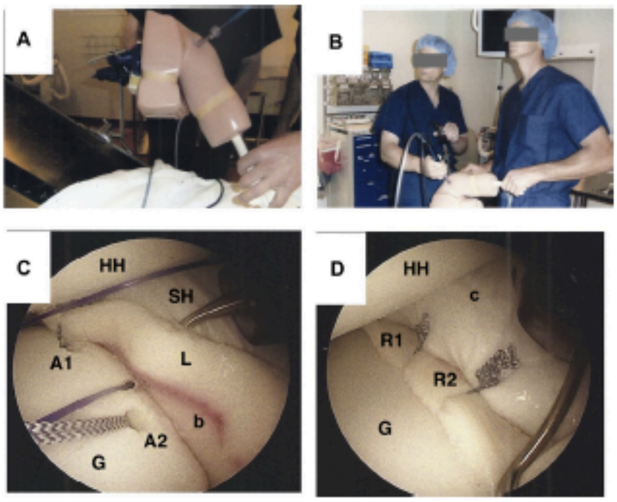


Fig 1. Dry shoulder simulator model used in study (left shoulder). (A) Anterior view of shoulder simulator oriented in beach-chair position. (B) An operator and assistant performing arthroscopic surgery on the simulator model oriented in the lateral decubitus position. (C, D) Arthroscopic view from anterosuperior portal. (C) Placement of the inferior-most anchor and sutures has been completed. The second anchor has been inserted, and the sutures have been retrieved out of the posterior cannula. A cannulated suture hook (SH) enters through the midanterior portal and is passed through the capsule and labrum (L) inferior to the exit of the suture anchor hole; a monofilament shuttle suture is then delivered. (A1, anchor position 1; A2, anchor position 2; b, Bankart lesion; G, glenoid; HH, humeral head.) (D) A hook probe is used to examine the completed repair (with the third anchor just out of view beneath the hook probe); the capsule (c) has been retensioned and the labrum secured to the glenoid rim. (G, glenoid; HH, humeral head; R1, repair position 1; R2, repair position 2.)

(Tables 5 and 6). Each step and error metric was scored as either yes or no, designating whether the specific event was or was not observed to have occurred by the reviewer. In addition to scoring steps and errors, each event characterized as DNTT was scored. There was no limit to the number of individual instances DNTT could be scored, with each occurrence simply tallied as a single error event. The score sheet also contained a box for specific reviewer comments for each metric. The 2 individual scores from a pair of reviewers were averaged to obtain the overall score for each step, error, or DNTT event. In addition, the score agreement or disagreement between the specific pair of reviewers was tabulated for each individual event (step, errors, and DNTT events) and used to calculate IRR correlations (as described in the "Statistical Methods" section). The total time in minutes was documented for each video, beginning with the first view of the arthroscope from the posterior portal and ending with withdrawal of the arthroscope after examination of the completed repair.

Table 4. Arthroscopic Instruments Used to Perform Bankart Procedure on Cadaveric Shoulder

5.5- and 8.5-mm obturator cannulas
Switching sticks
Hook probe
Regular and looped graspers
Liberator/elevator
Shaver
Drill guide/drill
Push-in anchor loaded with single suture
Mallet
Cannulated suture hook
Penetrator
Monofilament suture
Knot pusher
Arthroscopic scissors

Performance Benchmark

Prior research has used the metric-based mean performance of a group of experienced or expert operators to objectively define "proficiency."[7,9-11,19] Before initiating this study, the 4 primary investigators specified

Table 5. Copernicus Resident/Master Model Metric Validation: Novices

	64A	64B	64 Ave	94A	94B	94 Ave	24A	24B	24 Ave	34A	34B	34 Ave	14A	14B	14 Ave	44A	44B	44 Ave	84A	84B	84 Ave
Video																					
I. Portals																					
Steps uncompleted	0	0	0	0	0	0	0	0	0	3	2	2.5	2	3	2.5	1	2	1.5	1	2	1.5
Errors made	0	2	1	3	3	3	0	1	0.5	3	0	1.5	1	1	1	1	1	1	2	0	1
II. Instabl assess																					
Steps uncompleted	2	1	1.5	1	0	0.5	1	1	1	1	1	1	3	3	3	0	1	0.5	2	1	1.5
Errors made	0	0	0	1	1	1	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0
III. Caps/gen prep																					
Steps uncompleted	2	2	2	3	3	3	1	1	1	3	4	3.5	5	5	5	3	3	3	3	4	3.5
Errors made	1	0	0.5	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IV. First inf anch prep																					
Steps uncompleted	0	0	0	0	1	0.5	0	1	0.5	0	0	0	1	2	1.5	1	3	2	1	1	1
Errors made	1	0	0.5	2	1	1.5	3	3	3	1	4	2.5	1	2	1.5	0	1	0.5	0	0	0
V. First sut del/mgmt																					
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0	1	1	1	3	3	3	0	0	0	0	0	0	1	0	0.5
VI. First knot tying																					
Steps uncompleted	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	1	0.5	0	0	0	0	0	0	2	2	2	1	1	1	1	0	0.5
VII. Second anch prep																					
Steps uncompleted	0	0	0	0	1	0.5	1	0	0.5	0	0	0	1	1	1	1	1	1	1	0	0.5
Errors made	0	0	0	2	1	1.5	0	1	0.5	2	2	2	1	0	0.5	0	0	0	0	0	0
VIII. Second sut del/mgmt																					
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	2	1	1.5	0	0	0	1	1	1	3	3	3	0	0	0	0	0	0	0	0	0
IX. Second knot tying																					
Steps uncompleted	1	0	0.5	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
Errors made	2	1	1.5	0	1	0.5	0	0	0	1	1	1	0	0	0	0	0	0	1	0	0.5
X. Third anch prep																					
Steps uncompleted	0	0	0	0	1	0.5	1	1	1	0	0	0	1	1	1	1	1	1	0	0	0
Errors made	0	0	0	0	1	0.5	0	0	0	2	2	2	0	0	0	1	0	0.5	3	3	3
XI. Third sut del/mgmt																					
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	1	2	1.5	0	0	0	1	1	1	2	2	2	1	1	1	1	1	1	2	0	1
XII. Third knot tying																					
Steps uncompleted	0	1	0.5	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
Errors made	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	3	0	1.5
XIII. Eval repair																					
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total time (Port, Dx, and Rx), min			105			93			67			76			62			67			39
Rating pairs																					
Steps completed (45)	40	41	40.5	40	39	39.5	41	41	41	36	36	36	32	31	31.5	38	34	36	37	37	37
Errors made (77)	7	6	6.5	10	11	10.5	8	6	7	17	17	17	7	6	6.5	5	5	5	13	3	8

(continued)

Table 5. Continued

	64A	64B	64 Ave	94A	94B	94 Ave	24A	24B	24 Ave	34A	34B	34 Ave	14A	14B	14 Ave	44A	44B	44 Ave	84A	84B	84 Ave
Sentinel errors	3	2	2.5	1	4	2.5	3	3	3	3	4	3.5	1	3	2	0	1	0.5	4	1	2.5
Steps completed																					
Agreement			42			40			43			43			41			41			39
Disagreement			3			5			2			2			4			4			6
Step IRR			0.93			0.89			0.95			0.95			0.91			0.91			0.87
Errors made																					
Agreement			73			70			72			69			70			73			65
Disagreement			4			7			5			8			7			4			12
Error IRR			0.95			0.91			0.93			0.9			0.91			0.95			0.84
Total score (S + E)																					
Agreement			115			110			115			112			111			113			104
Disagreement			7			12			7			10			11			9			18
Total IRR			0.94			0.9			0.94			0.92			0.91			0.93			0.85

anch, anchor; caps, capsule; Dx, diagnostic; E, errors; eval, evaluate; glen, glenoid; inf, inferior; instab, instability; IRR, inter-rater reliability; mgmt, management; min, minutes; port, portals; Rx, treatment; S, steps; sut, suture; sut del, suture delivery.

that any participant from the experienced group who was performing more than 2 SDs better or worse than the group as a whole would be deemed an "outlier" and not representative of the experienced group. If any such performance by a participant in this group occurred, the participant's scores would be removed from the analysis so as not to skew the establishment of the reference benchmark.

Statistical Methods

For each of the 13 separate phases of the procedure, the numbers of "uncompleted steps" (steps that were not performed) and "errors made" were tabulated and the scores for the 2 reviewers averaged (Tables 5 and 6). These data were used to determine which of the procedural phases showed the greatest differences in performance when comparing the experienced and novice surgeons (1-factor analysis of variance) (IBM SPSS statistical software program; IBM, Armonk, NY). Furthermore, for the entire procedure, the total numbers of steps "completed," errors made, and sentinel errors enacted were also averaged and tabulated for the pair of reviewers.

The 2 raw score sheets were compared for each of the individual steps (45 steps in total) and the number of "agreements" tabulated (either both reviewers documented that a step was performed or both scored the step as not being completed). In addition, the number of "disagreements" in scoring steps was tabulated (one of the reviewers indicated that the step had been completed and the other scored that the step had not been completed). The IRR for the steps was calculated according to the following formula: Number of agreements/Number of agreements + Number of disagreements.

In a similar manner, there was either agreement or disagreement in the 2 scores for each of the potential errors (77 errors in total). The IRR for error scoring was calculated in the same manner as that for the steps. Finally, the IRR for scoring the metrics for the complete procedure was calculated using both the step and error agreements/disagreements for the entire procedure (122 in total). An acceptable IRR is equal to or greater than 0.80.[18]

Results

Participants

The entire group of 12 master and associate master instructors serving as faculty for an AANA Resident Course chose to participate and comprised the experienced group. The faculty, all fellowship trained in arthroscopy or sports medicine, averaged over 17 years in clinical practice, with each having routine experience in arthroscopic shoulder techniques. All faculty members have been recognized nationally by

Table 6. Copernicus Resident/Master Model Metric Validation: Experienced

	Excluded																																			
	15A	15B	15 Ave	25A	25B	25 Ave	35A	35B	35 Ave	65A	65B	65 Ave	75A	75B	75 Ave	45A	45B	45 Ave	55A	55B	55 Ave	85A	85B	85 Ave	95A	95B	95 Ave	105A	105B	105 Ave	115A	115B	115 Ave	125A	125 B	125 Ave
Video																																				
I. Portals																																				
Steps uncompleted	3	2	2.5	1	0	0.5	2	1	1.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	1.5	2	1	1.5
Errors made	0	1	1.5	0	0	0	1	0	0.5	1	0	0.5	0	1	0.5	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	3	2	2.5	0	0	0
II. Install anest																																				
Steps uncompleted	3	2	2.5	1	1	1	1	2	1.5	1	0	0.5	2	1	1.5	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.5	2	0	1	2	1	1.5
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
III. Caps/gen prep																																				
Steps uncompleted	2	3	2.5	2	1	1.5	4	4	4	1	3	2	2	3	2.5	2	2	2	1	3	2	1	2	1.5	1	3	2	1	3	2	4	3	3.5	3	3	3
Errors made	0	0	0	0	0	0	0	1	0.5	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
IV. First inf anch prep																																				
Steps uncompleted	1	1	1	0	0	0	0	1	0.5	0	1	0.5	0	0	0	1	0	0.5	0	0	0	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1
Errors made	2	2	2	0	1	0.5	2	2	2	1	0	0.5	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	2	2	2	0	0	0	0	0	0
V. First sut del/mgmt																																				
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	2	0	1	1	0	0.5	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	1	0.5	0	1	0.5	1	0	0.5	0	0	0
VI. First knot tying																																				
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	0	1	0.5	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0
VII. Second anch prep																																				
Steps uncompleted	1	1	1	0	0	0	1	1	1	0	1	0.5	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	0	2	1	1	1	1
Errors made	0	0	0	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	3	1	0	0.5
VIII. Second sut del/mgmt																																				
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0
Errors made	1	0	0.5	0	0	0	1	1	1	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
IX. Second knot tying																																				
Steps uncompleted	1	0	0.5	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.5	0	0	0	0	0	0
Errors made	0	0	0	0	0	0	1	1	1	0	0	0	0	1	0.5	1	2	1.5	0	0	0	0	0	0	0	0	0	0	0	0	2	1	1.5	0	1	0.5
X. Third anch prep																																				
Steps uncompleted	1	1	1	2	0	1	1	1	1	1	1	1	0	0	0	1	3	2	1	1	1	1	1	1	0	0	0	0	0	0	0	1	0.5	1	1	1
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
XI. Third sut del/mgmt																																				
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0
Errors made	1	0	0.5	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
XII. Third knot tying																																				
Steps uncompleted	1	1	1	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	1	0	0.5	0	0	0	1	1	1	0	0	0	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
XIII. Eval repair																																				

(continued)

Table 6. Continued

	Excluded																	
	15A	15B	15 Ave	25A	25B	25 Ave	35A	35B	35 Ave	65A	65B	65 Ave	75A	75B	75 Ave	45A	45B	45 Ave
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total time (Port, Dx, and Rx), min			38			47			37			37			29			43
Rating pairs																		
Steps completed (45)	33	34	33.5	41	43	42	36	33	34.5	42	39	40.5	42	41	41.5	40	42	41
Errors made (77)	5	3	4	3	1	2	7	6	6.5	3	2	2.5	4	6	5	3	4	3.5
Sentinel errors	1	0	0.5	0	0	0	3	3	3	0	0	0	1	1	1	0	3	1.5
Steps completed																		
Agreement			42			39			38			40			43			42
Disagreement			3			6			7			5			2			3
Step IRR			0.93			0.87			0.84			0.89			0.95			0.93
Errors made																		
Agreement			73			73			73			75			75			73
Disagreement			4			4			4			2			2			4
Error IRR			0.95			0.95			0.95			0.97			0.97			0.95
Total score (S + E)																		
Agreement			115			112			111			115			118			115
Disagreement			7			10			11			7			4			7
Total IRR			0.94			0.92			0.91			0.94			0.97			0.94

	Excluded																	
	55A	55B	55 Ave	85A	85B	85 Ave	95A	95B	95 Ave	105A	105B	105 Ave	115A	115B	115 Ave	125A	125 B	125 Ave
Steps uncompleted	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Total time (Port, Dx, and Rx), min			43			52			42			31			35			27
Rating pairs																		
Steps completed (45)	39	42	40.5	41	40	40.5	42	42	42	43	39	41	36	35	35.5	37	37	37
Errors made (77)	1	1	1	3	1	2	3	4	3.5	3	4	3.5	8	7	7.5	1	1	1
Sentinel errors	0	0	0	0	0	0	0	0	0	0	0	0	2	2	2	1	0	0.5
Steps completed																		
Agreement			43			44			41			41			38			41
Disagreement			2			1			4			4			7			4
Step IRR			0.95			0.98			0.91			0.91			0.84			0.91
Errors made																		
Agreement			75			75			76			76			72			75
Disagreement			2			2			1			1			5			2
Error IRR			0.97			0.97			0.99			0.99			0.93			0.97
Total score (S + E)																		
Agreement			118			119			117			117			111			116
Disagreement			4			3			5			5			11			6
Total IRR			0.97			0.97			0.96			0.96			0.91			0.95

anch, anchor; caps, capsule; Dx, diagnostic; E, errors; eval, evaluate; glen, glenoid; inf, inferior; instab, instability; IRR, inter-rater reliability; min, minutes; mgmt, management; port, portals; Rx, treatment; S, steps; sut, suture; sut del, suture delivery.

AANA for their talent and ability to teach and communicate shoulder arthroscopy skills to trainees. The novice group was made up of 7 volunteers (from a total of 46 orthopaedic residents registered for a weekend course) who elected to participate in the study and perform an arthroscopic Bankart repair on a shoulder model simulator.

IRR Assessments

The IRR calculations across each of the assessments were strong. The mean IRR for the paired scoring of all 19 videos was 0.92 (range, 0.84 to 0.98) for procedural steps and 0.94 (range, 0.84 to 0.99) for errors (includes DNTT events). The mean IRR for the total of steps and errors was 0.93 (range, 0.85 to 0.97). In no instance did any of the 3 IRR calculations for the 19 scored videos fall below 0.80.

Outlier Experienced Surgeon Performance

Before analysis of the data, score profiles of participants in the experienced group were examined for atypical performance. One subject in the experienced group was found to have enacted 3 sentinel errors in comparison with a mean experienced-group sentinel error rate equal to 0.71 (SD, 0.98). The mean and standard deviation were used to convert the data to Z scores, with a mean of 0 and an SD of 1. The outlier subject's Z score for sentinel errors was 2.31 (>2 SDs from the mean of the experienced group) ($P = .01$). When this subject's score is removed from the sentinel error data, the new mean equals 0.5 (SD, 0.71). Figure 2 shows that the exclusion of this 1 experienced surgeon had little impact on the overall experienced-group scores. In subsequent statistical analysis, all of this outlier's data were excluded.

Overall Performance Comparisons

On average, the experienced surgeons completed more steps than participants in the novice group (39.54 *v* 37.36), but this difference was not statistically significant (Fig 2A). The experienced group took significantly less time to perform the procedure on the shoulder model in comparison with the novice surgeons (39 minutes *v* 66 minutes, $P < .001$) (Fig 2B). The experienced surgeons also made significantly fewer errors than the novice surgeons (3.23 *v* 8.64, $P = .001$) (Fig 2C) and significantly fewer sentinel errors (0.5 *v* 2.36, $P < .001$) (Fig 2D). Overall, experienced arthroscopists made 63% fewer errors, made 79% fewer sentinel errors, and performed the procedure in 42% less time than novice surgeons. The procedural phases that exhibited the greatest differences in performance between the groups included anchor preparation and insertion, suture delivery and management, and knot tying. The experienced group also performed better than the novices on the phases of portal placement, arthroscopic instability

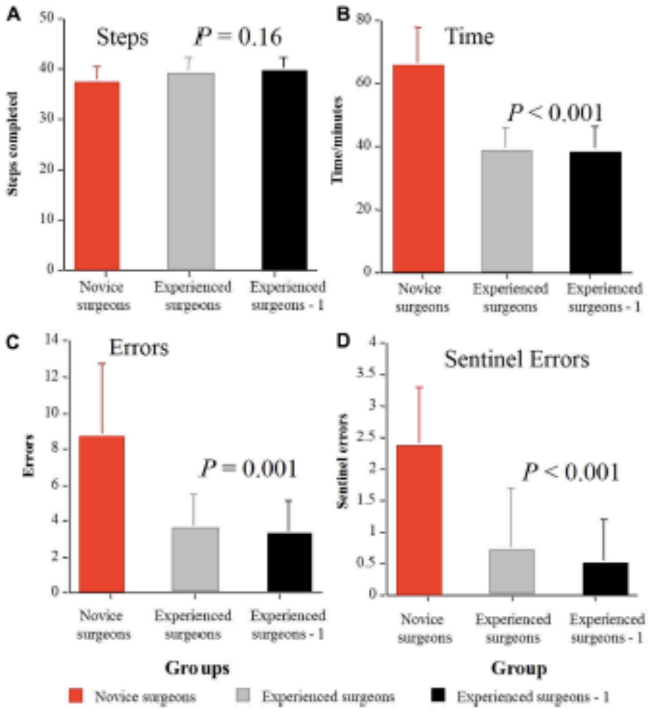


Fig 2. Summary performance data for the shoulder model simulator showing the mean and standard deviation scores for (A) steps completed, (B) time taken, (C) errors enacted, and (D) sentinel errors made by novice and experienced operators. Also shown are the mean scores of the experienced group with the 1 outlier's dataset excluded.

assessment, and capsulolabral mobilization/glenoid preparation, but the differences were not significant.

Anchor Preparation and Insertion Steps

Figure 3A shows the mean number of steps not completed by both groups during the anchor preparation/insertion phase of the procedure. Few steps were not completed or were omitted by either group. Experienced surgeons performed better on anchor 1, but surgeons in the novice group performed marginally better on anchors 2 and 3. These differences were not statistically significant.

Anchor Preparation and Insertion Errors

Figure 3B shows the mean (and standard deviation) number of errors made by the experienced group and the novice group during the anchor preparation and insertion phase of the procedure. Across all 3 anchors, the experienced group made fewer errors than the novices and showed more consistent performance as indicated by the smaller standard deviation scores. Experienced arthroscopists made significantly fewer errors than novices on the preparation and insertion of anchor 1 (0.32 *v* 1.36, $P = .02$). Although experienced surgeons also made considerably fewer errors than novices on the preparation and insertion of anchor 3, this did not reach statistical significance (0.14 *v* 0.86, $P = .07$).

Suture Delivery and Management Steps

The number of uncompleted steps during suture delivery and management for anchors 1, 2, and 3 was few, and there were almost no differences between the groups (Fig 3C).

Suture Delivery and Management Errors

Figure 3D shows the mean (and standard deviation) number of errors made by experienced and novice surgeons on the suturing steps of the procedure. The number of errors made by experienced arthroscopists was small

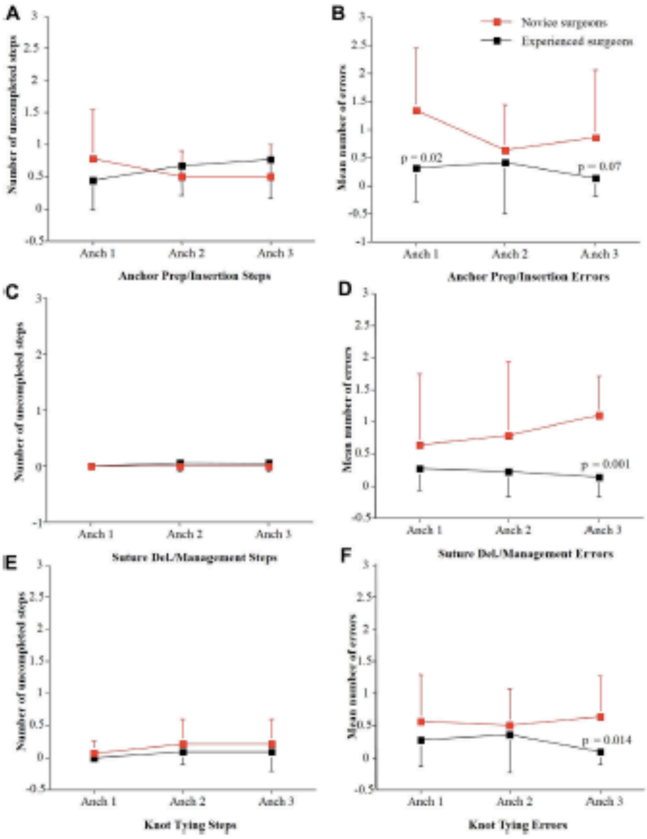


Fig 3. Mean and standard deviation data for novice and experienced surgeons for (A) anchor preparation (Prep) and insertion steps and (B) errors (phases IV, VII, and X); (C) suture delivery (Del) and management steps and (D) errors (phases V, VIII, and XI); and (E) knot-tying steps and (F) errors (phases VI, IX, and XII) during Bankart procedure. (Anch, anchor.)

across all 3 anchors and showed substantial consistency as indicated by the small standard deviation scores. In addition, their performance showed slight improvement across the anchors. In contrast, the novices showed considerable performance variability and performance deterioration across the 3 anchors. Only the differences in suture management and delivery on anchor 3, however, were found to be statistically significant (0.14 *v* 1.1, $P = .001$).

Knot-Tying Steps

Similar to the results for the suturing steps, only a small number of uncompleted or omitted steps during knot tying for anchors 1, 2, and 3 was observed for both groups (Fig 3E) and the small differences that did exist were not statistically significant.

Knot-Tying Errors

The experienced group consistently made fewer errors during the knot-tying phase of the model procedure (Fig 3F). They performed best on anchor 3 as indicated by their mean score and very small standard deviation. The novice group made, on average, less than 1 error per anchor. A significant difference between the groups was, however,

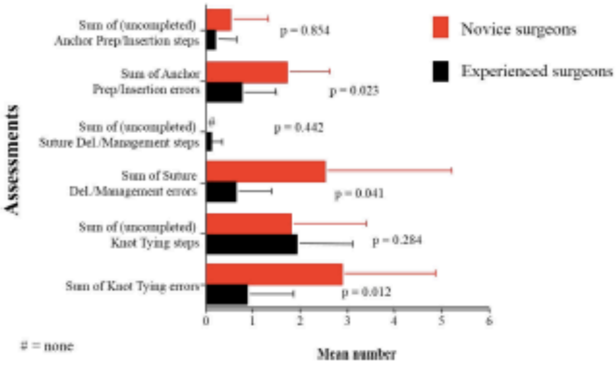


Fig 4. Mean and standard deviation data for novices and experienced surgeons across 6 groups of metrics: anchor preparation (Prep) and insertion steps and errors (phases IV, VII, and X); suture delivery (Del) and management steps and errors (phases V, VIII, and XI); and knot-tying steps and errors (phases VI, IX, and XII).

observed for knot-tying errors on anchor 3 (0.09 *v* 0.64, $P = .014$).

Performance Summary Assessments

The performances of both groups across the 6 measures presented in Figure 3 were summed to give an indicator of each group's overall performance on the 3-anchor repair. These data are presented in Figure 4. Although the experienced group completed more of the procedure steps than the novice group for the anchor preparation and insertion, suture management, and knot-tying phases, none of these differences in steps completed were statistically significant. In contrast, all of the error variables did show large and statistically significant differences between the groups. The experienced group made significantly fewer (*i.e.*, 70%) anchor preparation and insertion and suture management errors (0.86 *v* 2.9, $P = .012$). The experienced group also made 74% fewer suture delivery and management errors in comparison with the novice group; this difference was found to be statistically significant (0.6 *v* 2.5, $P = .041$). The smallest difference between the error performances of the 2 groups was in the knot-tying phase. The experienced group still made 57% fewer knot-tying errors than the novice group; this difference was also found to be statistically significant (0.73 *v* 1.7, $P = .023$).

Procedure Review

Both groups completed the final examination of the repair, and there were no errors committed during this phase of the procedure by either group.

Proficiency Benchmark

The experienced group had a mean total error rate of 3.23. A surgeon could not make a portion of an error, so for practical purposes, the error benchmark was rounded to the next greater whole number, 4. The experienced group also created a mean number of 0.5 sentinel errors (SD, 0.71), rounded to 1. Thus the overall benchmark is set at completing a 3-anchor Bankart repair with no more than 4 total errors and no more than 1 sentinel error.

Discussion

The most important findings of this study include verification that the arthroscopic Bankart metrics coupled with a shoulder model simulator are able to accurately discriminate between the performance of experienced and novice orthopaedic surgeons and show "construct validity" (Table 1). In addition, a performance benchmark was able to be established based on the mean performance of the experienced group and included completion of a 3-anchor Bankart repair, enacting no more than 4 total errors and 1 sentinel error.

Bankart Metrics

The primary intent of the study was to determine whether construct validity could be demonstrated for the previously established arthroscopic Bankart metrics coupled with the use of a shoulder model simulator. For construct validity to be demonstrated, the combination of ABR metrics and simulator tools must be able to discriminate between the performance of experienced and novice surgeons. The differences between the 2 groups were significant, and those that best distinguished between experienced and novice surgeons in the performance of an arthroscopic Bankart procedure included (1) the errors enacted in the performance of the procedure, (2) the number of sentinel errors made, and (3) the time it took to perform the procedure.

In this study, error scores were a very powerful and accurate[20] discriminator between the groups, with the novices making more than twice as many errors as the experienced group, with a difference in the standard deviation scores of a similar magnitude. The goal of surgical education should be to help trainees perform well with as few errors as possible. The trainee, however, should be afforded the opportunity to create errors in an inconsequential manner (e.g., on a simulator without associated patient morbidity) and learn from them. Effective progression in training should be demonstrable with a concomitant reduction in errors. For the individual trainee, the identification of specific errors facilitates a focused correction of deficiencies. Performance errors can also be used as a powerful metric tool to shape and configure the related educational curriculums and to establish benchmarks that trainees must meet and demonstrate before progressing.[11,19,21] Lastly, defined errors can serve to guide the development of simulators, that is, not only what they should emulate but also what they should measure.

Within surgical and procedural disciplines, there is unanimous agreement that certain types of technique errors are so egregious and pose such a threat to either the success of the procedure or patient safety that they should constitute their own performance category. We have elected to term these more serious deviations from optimal performance "sentinel errors." Similar designations have not been formally made in other studies that have objectively assessed surgical performance and evaluated the construct validity of specific metrics.[7-10,22,23] The use of such a special metric classification could have profound implications for "high-stakes" assessments (i.e., determining whether a trainee is allowed to progress in the specific educational/residency program) and proficiency-based progression approaches to training.

Although the number of steps completed did not distinguish between the groups in this study, it is nonetheless important to include all of the essential steps in a training program. The procedure cannot be completed without knowing and performing all of the correct steps in the proper order.[11,19,21] The fact that this performance unit is being assessed also increases the probability that the steps will be learned.[24] However, the proper steps and sequences should be communicated and learned outside of the skills training proper, such as in an online educational module, because its inclusion is sensible rather than essential. A thorough diagnostic evaluation of pathology potentially related to shoulder instability is necessary for the comprehensive and appropriate treatment of the unstable shoulder and, thus, its inclusion in the steps of the procedure.

Simulator Model

Gallagher and O'Sullivan[11] have proposed that to be effective, a simulation model should provide the learner with the span of appropriate sensory responses to physical actions that are behaviorally consistent with what would be experienced in the real situation, including the opportunity to enact both appropriate actions (steps) and inappropriate actions (errors). The simulator should also afford the opportunity to execute the procedure in the same order and with the same tools and devices with which the procedure would normally be performed.[21] The simulator model used in this study was sufficiently realistic to provide the learner the opportunity to perform each of the 45 steps in a realistic fashion using the same tools, implants, and techniques used for an anterior stabilization in a real patient.

Benchmarking

The second purpose of this study was to establish a performance benchmark for the arthroscopic Bankart metrics coupled with a shoulder model simulator. The definition of "proficiency," in distinction to qualitatively described "competency," is based on objectively defined performance metrics. Proficiency-based progression training requires the establishment of a benchmark that trainees must reach to be able to progress. We sought to establish an objective, reliable, transparent, and fair performance benchmark for an anterior Bankart repair on the shoulder model simulator. Because the benchmark was to be established based on the mean performance of the group of experienced surgeons, it was important that the performance of the members be representative of that group. On the basis of a pre-study stipulation, the scores of 1 member of the experienced group were removed from the analysis because the performance was 2.3 SDs worse than that of this participant's peers. This policy was established so as not to skew the creation of the reference benchmark and was shown to have little impact on the overall scores of the experienced group.

An IRR equal to or greater than 0.80 is considered acceptable.[18,25] The very high IRR for the scores from reviewer pairs for the entire group of metrics (0.93) is reflective of the clarity and precision of the arthroscopic Bankart metrics drafted, as well as the thorough training of the 10 reviewers. The ability to score the steps and errors consistently is essential in obtaining a reliable measure of the surgeon's performance and skill level for a particular procedure.

Limitations

A limitation of this study is that there was no confirmation that the participants identified as master/associate master surgeons and representative of the experienced group possessed a specified level of expert skill in performing an arthroscopic Bankart procedure. Nevertheless, the individual surgeons so identified have been recognized by the AANA as skilled and effective educators either from lecture presentations with video exhibiting skilled shoulder arthroscopic techniques or from their performance in an arthroscopic laboratory setting in which they demonstrated the ability to teach each of the key components of a Bankart procedure. Therefore this group was defined as "experienced" rather than "expert." Similarly, other than identification of the year in training, no information was obtained to determine the extent of the novice group's (resident) experience with arthroscopic shoulder surgery, that is, the number of arthroscopy/sports medicine rotations previously completed or the number of shoulder arthroscopic procedures in which the participant served as an assistant surgeon. Even with those pieces of information, it would not be possible to gain any reliable measure of an individual resident's level of first-hand "experience" or skill with arthroscopic shoulder surgery. Furthermore, the structure of residency rotations and level of independence permitted vary a great deal among training programs. As a result, the residents' knowledge and skill sets are unlikely to be uniform but the PGY 4 and 5 levels provide a general measure of their training experience consistent with the designation "novice."

The numbers of surgeons in both groups were small, but one of the strengths of the detailed metric-based procedure characterization method that we used is the sensitivity to detect differences when in fact they exist. We obtained 123 data points (metrics) including the duration of the procedure in minutes for each scored video. Thus small numbers of subjects can still produce statistically powerful differences, assuming that performance has been reliably measured. An a priori power analysis was not performed because we found no previous studies of arthroscopic shoulder repair that used a similar detailed assessment methodology. This is the first study of its kind in this field, and our results will afford other researchers the opportunity to develop their sample sizes based on the reported mean and standard deviation scores. The only published scientific reports we could draw on were similar types of studies published in the laparoscopic surgical literature. Those reports could only give an indication of the possible sample sizes required.

Although not specifically a limitation of the study or design, the option to use either the lateral decubitus or beach-chair position could potentially introduce some variability. Both patient positions are in common use among practicing surgeons. The metrics were carefully constructed to facilitate unbiased scoring for the model simulator/patient in either orientation with no penalty. Several metrics require that the arthroscope be placed in the anterosuperior portal to adequately complete the step; however, this is true for both orientations (e.g., step 12, "view or probe the insertion of the anterior glenohumeral ligaments onto the anterior humeral neck," and step 16, "obtain a view of the anterior glenoid neck"). The challenge in completing these steps relates more to the position of the arthroscope (posterior *v* anterosuperior portal) than to the patient orientation. Although we believe that the lateral decubitus orientation makes some steps easier to perform (e.g., appropriate seating of the drill guide in the anterior/posterior dimension relative to the bony rim, as well as accurate passage of the suturing device through the capsulolabral tissue inferior to the anchor site), no inherent bias is introduced in scoring the metrics for procedures performed with either patient orientation.

On the basis of the data from this study, the null hypothesis is rejected. The shoulder model used, coupled with previously validated arthroscopic Bankart metrics, is able to accurately distinguish between experienced and novice operators. Construct validity is demonstrated for the simulator model coupled with Bankart performance metrics.

Conclusions

The tool comprising validated arthroscopic Bankart repair metrics coupled with a dry shoulder model simulator is able to accurately distinguish between the performance of experienced and novice orthopaedic surgeons. A performance benchmark based on the experienced group includes completion of a 3-anchor Bankart repair, enacting no more than 4 total errors and 1 sentinel error.

References

1. Healy GB. The college should be instrumental in adapting simulators to education. *Bull Am Coll Surg* 2002;87:10-11.
2. Pellegrini CA. Surgical education in the United States: Navigating the white waters. *Ann Surg* 2006;244:335-342.
3. Sachdeva AK, Pellegrini CA, Johnson KA. Support for simulation-based surgical education through American College of Surgeons—accredited education institutes. *World J Surg* 2008;32:196-207.
4. Selzer DJ, Dunnington GL. Surgical skills simulation: A shift in the conversation. *Ann Surg* 2013;257:594-595.
5. Zendejas B, Brydges R, Hamstra SJ, et al. State of the evidence on simulation-based training for laparoscopic surgery: A systematic review. *Ann Surg* 2013;257:586-593.
6. Satava RM. Virtual reality surgical simulator. The first steps. *Surg Endosc* 1993;7:203-205.
7. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: Results of a randomized, double-blinded study. *Ann Surg* 2002;236:458-464.
8. Grantcharov TP, Kristiansen VB, Bendix J, et al. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg* 2004;91:146-150.
9. Ahlberg G, Enochsson L, Gallagher AG, et al. Proficiency-based virtual reality training significantly reduces the

error rate for residents during their first 10 laparoscopic cholecystectomies. *Am J Surg* 2007;193:797-804.

10. Van Sickle K, Ritter EM, Baghai M, et al. Prospective, randomized, double-blind trial of curriculum-based training for intracorporeal suturing and knot tying. *J Am Coll Surg* 2008;207:560-568.
11. Gallagher AG, O'Sullivan GC. *Fundamentals of surgical simulation; principles & practices.* London: Springer Verlag, 2011.
12. Angelo RL, Ryu RKN, Pedowitz RA, Gallagher AG. Metric development for an arthroscopic Bankart procedure: Assessment of face and content validity. *Arthroscopy* 2015;31:1430-1440.
13. Morgan CD, Bodenstab AB. Arthroscopic Bankart suture repair: Technique and early results. *Arthroscopy* 2010;26:819-820.
14. Streubel PN, Krych AJ, Simone JP, et al. Anterior glenohumeral instability: A pathology-based surgical treatment strategy. *J Am Acad Orthop Surg* 2014;22:283-294.
15. Waterman BR, Burns TC, McCriskin B, et al. Outcomes after Bankart repair in a military population: Predictors for surgical revision and long-term disability. *Arthroscopy* 2014;30:172-177.
16. Shibata H, Gotoh M, Mitsui Y, et al. Risk factors for shoulder re-dislocation after arthroscopic Bankart repair. *J Orthop Surg Res* 2014;9:53.
17. Ryu RK. Arthroscopic approach to traumatic anterior shoulder instability. *Arthroscopy* 2003;19:94-101.
18. American Educational Research Association. Standards for educational and psychological testing. 2014 ed. Available at www.apa.org/science/programs/testing/standards.aspx. Accessed January 12, 2014.
19. Gallagher AG, Ritter EM, Champion H, et al. Virtual reality simulation for the operating room: Proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg* 2005;241:364-372.
20. Rossi MJ, Lubowitz JH, Provencher MT, Poehling GG. Precision versus accuracy: A case for common sense. *Arthroscopy* 2012;28:1043-1044.
21. Gallagher AG. Metric-based simulation training to proficiency in medical education—What it is and how to do it. *Ulster Med J* 2012;81:107-113.
22. Seymour NE, Gallagher AG, Roman SA, et al. Analysis of errors in laparoscopic surgical procedures. *Surg Endosc* 2004;18:592-595.
23. Larsen CR, Soerensen JL, Grantcharov TP, et al. Effect of virtual reality training on laparoscopic surgery: Randomized controlled trial. *BMJ* 2009;338:b1802.
24. Reinhardt-Rutland AH, Gallagher AG. Visual depth perception in minimally invasive surgery. In: Robertson SA, ed. *Contemporary ergonomics.* London: Taylor and Francis, 1995;531-536.
25. Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: Rigorous science for the assessment of surgical education and training. *Surg Endosc* 2003;17:1525-1529.

Appendix 7 (Chapter 5)

The Bankart Performance Metrics Combined With a Cadaveric Shoulder Create a Precise and Accurate Assessment Tool for Measuring Surgeon Skill

Richard L. Angelo, M.D., Richard K. N. Ryu, M.D., Robert A. Pedowitz, M.D., Ph.D., and Anthony G. Gallagher, Ph.D., D.Sc.

Purpose: To determine if previously validated performance metrics for an arthroscopic Bankart repair (ABR) coupled with a cadaveric shoulder are a valid assessment tool with the ability to discriminate between the performances of experienced and novice surgeons and to establish a proficiency benchmark for an ABR using a cadaveric shoulder. **Methods:** Ten master/associate master faculty from an Arthroscopy Association of North America Resident Course (experienced group) were compared with 12 postgraduate year 4 and postgraduate year 5 orthopaedic residents (novice group). Each group was instructed to perform a diagnostic arthroscopy and a 3 suture anchor Bankart repair on a cadaveric shoulder. The procedure was videotaped in its entirety and independently scored in blinded fashion by a pair of trained reviewers. Scoring was based on defined and previously validated metrics for an ABR and included steps, errors, "sentinel" (more serious) errors, and time. **Results:** The inter-rater reliability was 0.92. Novice surgeons made 50% more errors (5.86 *v* 2.95, $P = .013$), showed more performance variability (SD, 1.86 *v* 0.55), and took longer to perform the procedure (45.5 minutes *v* 25.9 minutes, $P < .001$). The greatest difference in errors related to suture delivery and management (exclusive of knot tying) (1.95 *v* 0.45, $P = .024$). **Conclusions:** The assessment tool composed of validated arthroscopic Bankart metrics coupled with a cadaveric shoulder accurately distinguishes the performance of experienced from novice orthopaedic surgeons. A benchmark based on the mean performance of the experienced group includes completion of a 3-anchor Bankart repair, and enacting no more than 3 total errors and 1 sentinel error. **Clinical Relevance:** Validated procedural metrics combined with the use of a cadaveric shoulder can be used to assess the performance of an ABR. The methodology used may serve as a template for outcomes-based procedural skills training in general.

The traditional manner in which surgical trainees have acquired their operative skills is under considerable pressure. Concerns about patient safety,[1,2] pressures on operating room efficiency,[3] and the reduced availability of work hours[4,5] have resulted in fewer opportunities for in vivo operative experience. As a consequence, trainees are graduating from residency programs with considerably less operative experience and almost certainly less technical skill than residents graduating in the past who were exposed to greater surgical volumes. For example, Bell et al.[6] found that of the 121 surgical procedures that general surgery residency program directors believed residents should be competent in by the time of graduation, only 18 of them had been performed with sufficient frequency by residents for them to acquire competence during their training. They also found that the mode frequency with which the 121 procedures were performed was 0. The implications of these findings for surgical training are considerable and concerning. At a more practical level, it means that surgical skills training must be optimized and preparation for a surgical practice maximized.

Traditionally, surgical residents have been trained using the "apprenticeship" model, dependent in part on exposure to surgical cases, variable graduated participation in surgery, and time spent on specific clinical

From ProOrtho Clinic (R.L.A.), Kirkland, Washington; The Ryu Hurvitz Orthopedic Clinic (R.K.N.R.), Santa Barbara, California; Professor Emeritus, University of California (R.A.P.), Los Angeles, California, U.S.A.; and ASSERT, University College Cork (A.G.G.), Cork, Ireland.

The authors report the following potential conflict of interest or source of funding: R.L.A. receives support from DePuy Mitek. R.K.N.R. receives support from MedBridge, Mitek, and Rotation Medical.

Received June 25, 2014; accepted May 13, 2015.

Address correspondence to Richard L. Angelo, M.D., ProOrtho Clinic, 12911 120th Ave NE, Ste H-210, Kirkland, WA 98072, U.S.A. E-mail: rlangelo@comcast.net

0749-8063/14538/$36.00

http://dx.doi.org/10.1016/j.arthro.2015.05.006

Table 1. Glossary

	Definition
Construct validity	A type of evidence that supports that specific test items identify the quality, ability, or trait they were designed to measure
Content validity	An estimate (opinion) by experts of the validity of a testing instrument based on a detailed examination of the contents of the test items
Damage to non-target tissue	Iatrogenic damage to tissues not intended to be addressed in the specific step (e.g., articular cartilage damage)
Definition	A definite, distinct, and clear objective characterization providing an accurate and reliable identification of whether an event was or was not observed to have occurred
Delphi Panel (modified)	A structured communication technique originally developed as a systematic, interactive forecasting method that relies on the opinion of a panel of experts; in modified form, experts answer queries or vote in 2 or more rounds (cycles) on the appropriateness of the metric-based operational definitions of detailed aspects of procedure performance with the goal of achieving consensus—voting is not anonymous
Error	A deviation from optimal performance
Face validity	An estimate (opinion) by experts who review the content of an assessment or tool to see if it seems appropriate and relevant to the concept it purports to measure
Inter-rater reliability	The extent of agreement between 2 raters on the occurrence of a series of observed events; it ranges between 0, no agreement, and 1.0, complete agreement
Metric	A standard of measurement of quantitative assessments used for objective evaluations to make comparisons or to track performance
Operational definition	Terms used to define a variable or event in terms of a process (or set of validation tests) needed to determine its existence, quantity, and duration
Performance metric	The features determining the accomplishment of a given task measured against preset known standards of accuracy and completeness
Procedure phase	A group or series of integrally related events or actions that, when combined with other phases, make up or constitute a complete operative procedure
Proficiency/proficient	A specific level of performance defined by a quantitative score (benchmark) or scores on a standardized test or other form of assessment
Proficiency-based progression	A training program that dictates that skill performance be demonstrated, to a predetermined benchmark level, by the trainee before advancement to more complex techniques
Sentinel error	An event or occurrence involving a serious deviation from optimal performance during a procedure that either (1) jeopardizes the success/desired result of the procedure or (2) creates iatrogenic insult to the patient's tissues
Step	A component task, the series aggregate of which constitutes the completion of a specific procedure
Task analysis	An assessment of how a procedure is accomplished, including a detailed (functional) description of the manual activities or tasks along with their duration, frequency, and complexity and any other unique and distinguishing factors
Task deconstruction	To break down a procedure into constituent tasks, steps, or components

rotations. At the outset, we sought to determine if a "proficiency-based progression" (PBP) method was potentially a more effective manner in which to train surgical skills than the apprenticeship model (Table 1 includes a glossary of terms used throughout the article). A PBP training program dictates that skill performance, to a predetermined benchmark level, be demonstrated by the trainee before advancing to more complex techniques. This method relies on a comprehensive and quantitative characterization of the skills to be learned. These performance characteristics, or "metrics," and their "operational definitions" (rather than descriptions) (Table 1) offer very specific goals and guidelines as part of the training curriculum. Previously, we reported on the development of "performance metrics" ("steps" and "errors")[7] (Table 1) for a standard reference approach to performing an arthroscopic Bankart repair (ABR).[8-12] Those metrics were derived from a careful "task analysis" and "deconstruction" (Table 1) using videos of complete Bankart procedures performed with patients in either the lateral decubitus or beach-chair orientation. The metrics were constructed in such a manner that they could be scored in an identical manner with the patient in either orientation. "Face validity" and "content validity" of the metrics were verified using a modified Delphi Panel methodology (Table 1). The Delphi Panel was composed of 27 experienced shoulder arthroscopists who have all served as master or associate master faculty for Arthroscopy Association of North America (AANA) shoulder courses at the Orthopedic Learning Center (Rosemont, IL). The Delphi Panel obtained excellent consensus on the metric-based characterization of the Bankart procedure.

In a subsequent report, we verified "construct validity" (the ability to discriminate between the performance of experienced and novice groups of surgeons) (Table 1) for the use of the ABR metrics with a shoulder model simulator as a training tool.[13] In the present study we evaluate the construct validity of these exact same metrics on a much higher-fidelity platform, the human cadaveric shoulder. At present, an ABR in a cadaveric shoulder provides the closest approximation to a similar surgical repair in a live patient. Full-physics, high-fidelity virtual reality simulators with haptic feedback are likely to play a greater role in the future but are expensive to develop and not currently available. Even with such simulators, validated metrics will be needed to substantiate their effectiveness.

The purpose of this study was to determine if previously validated performance metrics for an ABR coupled with a cadaveric shoulder are a valid assessment tool with the ability to discriminate between the performances of experienced and novice surgeons. We also sought to establish a "proficiency" (Table 1) benchmark for this procedure using the cadaveric shoulder. The null hypothesis was that when using a cadaveric shoulder, the Bankart metrics would fail to discriminate between experienced and novice surgeon performance.

Methods

No institutional review board (IRB) approval was obtained for this study investigating the validity of the Bankart metrics coupled with the cadaveric shoulder. IRB approval was sought for the final Copernicus Study proper, which will compare 3 different training protocols evaluating surgical simulation and proficiency-based training methods. The Western IRB (No. 1-776362-1) opined that, as an educational curriculum study, this study was exempt from the need for full IRB approval [based on the criteria of 45 CFR 46.101(b)(1)]. The final study comparing the 3 training protocols was registered with the National Institutes of Health (ClinicalTrials.gov No. NCT01921621).

Study Groups

Two groups were compared in their performance of an arthroscopic Bankart procedure on a cadaveric shoulder. The experienced group consisted of all faculty members who served as master or associate master instructors for a standard 3-day AANA Resident Course conducted at the Orthopedic Learning Center. "Experienced" meant that they performed the procedure consistently in practice and taught the principles at the Orthopedic Learning Center during shoulder courses. An "expert" would not be possible to define without the surgeon meeting objective performance criteria (metrics) that achieved a specific benchmark (that some group or body determined meant "expert" performance). The novice group was limited to postgraduate year (PGY) 4 and PGY 5 orthopaedic residents who had registered for a Resident Course and who volunteered to participate in the investigation.

ABR Metrics

Metrics have been previously defined for a standard reference ABR.[7] Forty-five essential steps in 13 "procedural phases" (Table 1) (in Roman numerals) were defined with beginning points and endpoints (Table 2). Twenty-nine potential unique errors were specified (Table 3), 8 of which were designated as "sentinel" (Table 1). The more serious (sentinel) errors were defined as those expected to either (1) substantially compromise the outcome of the shoulder stabilization (e.g., "capsular penetration of the suture passing instrument is superior to the anchor hole" resulting in failure to achieve retention of the capsule and inferior glenohumeral ligaments) or (2) potentially lead to iatrogenic damage to the shoulder (e.g., "laceration of the intact labrum"). Some of the same errors could be enacted more than once during separate but similar phases of the procedure (e.g., suture delivery and management for each of 3 anchors). Thus a total of 77 potential errors, 20 of which were sentinel errors, were specified for the complete procedure. In addition, events that led to less consequential "damage to non-target tissues" (DNTT) (Table 1) were simply recorded as standard errors (e.g., scuffing of the articular cartilage). A perfect score would indicate that all 45 steps were completed satisfactorily without committing any errors.

There are many ways to perform a Bankart procedure, but the task deconstruction was designed for a "reference" procedure (a routine Bankart repair), breaking it down into essential components. Each of the metrics was specifically crafted to accommodate different methods that can be used to accomplish the steps—for example, suture passage could be performed with a number of different instruments and techniques, but to accomplish re-tensioning of the capsulolabral tissue, the capsule must be purchased inferior to the anchor site (one of the metrics). The modified Delphi Panel procedure used to obtain face and content validity asks the following question of each panel member: "Is this metric (step or error) acceptable as written?," that is, "It is not incorrect" (although a particular panel member might perform the step in a different manner). The 45 steps and 77 errors were drafted, revised, and stress tested by the core group of primary investigators (R.L.A., R.K.N.R., R.A.P., A.G.G.) and then submitted to the Delphi Panel for comment, modification, and revision. The panel then obtained consensus for all of the 122 metrics, which were found to be acceptable in their final form.

Cadaveric Shoulder Study Specimens

Fresh-frozen cadaveric specimens with a complete shoulder girdle from the scapula and associated soft tissues to the mid humerus were used. After appropriate thawing, the scapula was mounted with a clamp

Table 2. Thirteen Phases of Bankart Procedure (in Roman Numerals) and Brief Summary of 45 Steps of Procedure

I. Portals
1. Posterior portal established
2. View posterior humeral head and extent of the Hill-Sachs when present
3. Introduce mid-anterior spinal needle immediately superior to the subscapularis and direct it toward the anteroinferior glenoid and labrum
4. Establish a cannula that abuts the superior border of the subscapularis near the lateral subscapularis insertion
5. Demonstrate instrument access to the anteroinferior glenoid/labrum
6. Introduce anterosuperior spinal needle at the superolateral aspect of the rotator interval and direct it toward the anterior glenoid
7. Establish an anterosuperior cannula, arthroscopic sheath, or switching stick
II. Arthroscopic instability assessment
View from posterior portal
8. View or probe the superior labral attachment onto the glenoid
9. View or probe articular surface of the rotator cuff
10. Probe anteroinferior glenoid/Bankart pathology including rim fracture, articular defect
View from anterosuperior portal
11. View or probe the midsubstance of the anterior-inferior glenohumeral ligaments
12. View or probe the insertion of the anterior glenohumeral ligaments onto the anterior humeral neck
III. Capsulolabral mobilization/glenoid preparation
13. Elevate the capsulolabral tissue from the glenoid neck and articular margin
14. View the subscapularis muscle superficial to the mobilized capsule
15. With an instrument, grasp and perform an inferior to superior shift of the capsulolabral tissue (to show tension is restored)
16. Obtain a view of the anterior glenoid neck
17. Mechanically abrade the glenoid neck
IV. Inferior anchor preparation/insertion
18. Seat the guide for the most inferior anchor hole at the inferior region of the anteroinferior quadrant
19. Drill anchor hole oblique to the glenoid articular face
20. Insert anchor
21. Test the anchor security by pulling on the suture tails
V. Suture delivery/management
22. Pass a cannulated suture hook or suture retriever through the capsular tissue—inferior to the anchor
23. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula
VI. Knot tying
24. Deliver an arthroscopic sliding knot
25. Back up with 3 or 4 half-hitches
26. Cut suture tails
VII. Second anchor preparation/insertion
27. Seat the drill guide for the second anchor superior to the first anchor and inferior to the equator
28. Drill anchor hole oblique to the glenoid articular face
29. Insert suture anchor
30. Test anchor security by pulling on the suture tails
VIII. Suture delivery/management
31. Pass a cannulated suture hook or suture retriever through the capsular tissue inferior to the suture anchor
32. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula
IX. Knot tying
33. Deliver an arthroscopic sliding knot
34. Back up with 3 or 4 half-hitches
35. Cut suture tails
X. Third anchor preparation/insertion
36. Seat the drill guide for the third anchor at or superior to the equator
37. Drill anchor hole oblique to the glenoid articular face
38. Insert suture anchor
39. Test anchor security by pulling on suture tails
XI. Suture delivery/management
40. Pass a cannulated suture hook or suture retriever through the capsular tissue at or inferior to the suture anchor
41. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula
XII. Knot tying
42. Deliver an arthroscopic sliding knot
43. Back up with 3 or 4 half-hitches
44. Cut suture tails
XIII. Procedure review
45. View and/or probe final completed repair

Table 3. Summary of 29 Different Bankart Procedure Metric Errors

Failure to maintain intra-articular position of the posterior cannula
Failure to maintain intra-articular position of the mid-anterior cannula
Failure to maintain intra-articular position of the anterosuperior cannula
Damage to the superior border of the subscapularis during placement of the mid-anterior portal
Damage to the anterior border of the supraspinatus during placement of the anterosuperior portal
Loss of intra-articular position of arthroscope/sheath or operating cannula (loss of each portal is scored only once for each Roman numeral, i.e., up to a total of 3 for arthroscope + 2 portals)
Lacerate intact capsulolabral tissue (sentinel error)
Failure to maintain control of a working instrument (sentinel error)
Guide is not located in the inferior region of the anteroinferior quadrant of the glenoid for the most inferior anchor
Entry of the completed tunnel lies outside safe zone of 0 to 3 mm from the bony glenoid rim (sentinel error)
Shallow undermining and deformation of articular cartilage (sentinel error)
Failure to maintain secure seating of the drill guide during anchor insertion
Breakage of the implant
Implant remains visibly proud (sentinel error)
Failure to insert the anchor with the inserter laser line (when present) to or beyond the laser line on the drill guide
Anchor fails to remain securely fixed within bone at the appropriate depth
Capsular penetration is at or superior to anchor hole (sentinel error)
Capsular penetration is not at or peripheral to the capsulolabral junction
Instrument breakage
Tearing of capsulolabral tissue
Uncorrected entanglement of shuttling device or suture
Offloading of suture anchor
Breakage of suturing device
Failure to create and maintain indentation of the capsule or labral tissue on knot completion (sentinel error)
Visible void is present between throws of the completed primary knot (sentinel error)
Completed knot abuts articular cartilage
Visible void is present between throws of the complete half-hitches
Suture breakage
Guide is inferior to the equator of the glenoid (for the third and final anchor)

NOTE. Metric errors can be associated with multiple phases and steps of the procedure (77 total errors).

in the subject surgeon's orientation of preference (lateral decubitus *v* beach chair). The cadaveric specimens were considered acceptable if (1) arthroscopic visibility of the target tissues was obtainable; (2) the specimen (flexibility) permitted adequate access to the target tissues; and (3) the integrity of the capsulolabral tissues was sufficient to permit mobilization, suture delivery, and knot tying. One of 3 designated AANA shoulder arthroscopy master instructors (the surgeon members of the group who created the arthroscopic Bankart metrics; R.L.A., R.K.N.R., R.A.P.) determined the acceptability of the cadaveric specimens.

Arthroscopic Bankart Repair

During a single weekend AANA resident arthroscopy course, the surgeons from both groups were instructed to establish portals (posterior, anterosuperior, and mid-anterior), complete a thorough diagnostic arthroscopy, and perform a 3-anchor ABR on the cadaveric shoulder. Furthermore, they were instructed to demonstrate/complete all of the steps for the Bankart repair that they would normally perform in clinical practice on a real patient. Equipment representatives from multiple different vendors served as surgical assistants and were randomly assigned to participating surgeons. The assistants were instructed to act only at the specific direction of the operating surgeon. Prompting and coaching (of technique) were prohibited (the procedures were proctored by staff from the Orthopedic Learning Center). A standard equipment tower with a 30° arthroscope was provided along with all instruments necessary to complete an ABR (Table 4).

The subject surgeon identified bony landmarks with a marking pen, established their desired portals, and performed a diagnostic examination. The arthroscope was then withdrawn from the shoulder joint. One of the 3 Master surgeons who evaluated the cadaver acceptability according to the criteria noted earlier, then

Table 4. Arthroscopic Instruments Used to Perform Bankart Procedure on Cadaveric Shoulder

5.5- and 8.5-mm obturator cannulas
Switching sticks
Hook probe
Regular and looped graspers
Liberator/elevator
Shaver
Drill guide/drill
Push-in anchor loaded with single suture
Mallet
Cannulated suture hook
Penetrator
Monofilament suture
Knot pusher
Arthroscopic scissors

Table 5. Copernicus Cadaver - Novice

			33			73			123			53			83			113			13			93			43			63			23	103	103	103
Video #	33A	33B	ave	73A	73B	ave	123A	123B	ave	53A	53B	ave	83A	83B	ave	113A	113B	ave	13A	13B	ave	93A	93B	ave	43A	43B	ave	63A	63B	ave	23A	23B	ave	A	B	ave
I - Portals																																				
Steps uncompl.	1	1	1	1	1	1	3	3	3	0	0	0	1	1	1	0	3	1.5	1	1	1	3	1	2	1	0	0.5	0	0	0	0	0	0	0	0	0
Errors made	2	2	2	0	0	0	1	1	1	0	3	1.5	0	0	0	2	1	1.5	1	0	0.5	1	0	0.5	0	1	0.5	0	0	0	0	2	1	0	0	0
II - Instabl Anes.																																				
Steps uncompl.	3	5	4	3	5	4	3	3	3	0	1	0.5	2	3	2.5	1	2	1.5	1	1	1	3	3	3	1	0	0.5	1	0	0.5	2	0	1	1	1	1
Errors made	0	0	0	0	0	0	1	0	0.5	1	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
III - Caps/Gen Prep																																				
Steps uncompl.	1	2	1.5	3	3	3	4	4	4	1	1	1	5	5	5	4	4	4	1	2	1.5	3	2	2.5	3	1	2	5	3	4	3	3	3	1	3	2
Errors made	0	1	0.5	0	0	0	0	0	0	1	1	1	0	1	0.5	0	0	0	0	3	1.5	1	1	1	1	0	0.5	2	1	1.5	0	0	0	0	0	0
IV - 1st Inf Anch Prep																																				
Steps uncompl.	0	0	0	1	1	1	1	0	0.5	0	0	0	1	0	0.5	1	1	1	1	0	0.5	0	0	0	0	0	0	0	1	0.5	1	0	0.5	0	0	0
Errors made	0	0	0	0	0	0	1	2	1.5	0	0	0	1	0	0.5	1	2	1.5	0	0	0	3	2	2.5	1	1	1	0	0	0	0	1	0.5	0	0	0
V - 1st Sut Del / Mgmt																																				
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	1	1	1	0	0	0	2	1	1.5	1	1	1	2	1	1.5	1	1	1	0	0	0	2	3	2.5	0	0	0	4	1	2.5	0	0	0	2	2	2
VI - 1st Knot Tying																																				
Steps uncompl.	0	0	0	0	0	0		Procedure terminated by surgeon		0	0	0	1	0	0.5	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	1	0	0.5				0	1	0.5	2	1	1.5	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5
VII - 2nd Anch Prep																																				
Steps uncompl.	1	1	1	1	0	0.5				1	1	1	2	1	1.5	0	0	0	1	0	0.5	0	1	0.5	0	1	0.5	1	1	1	1	1	1	2	1	1.5
Errors made	0	0	0	0	0	0				1	2	1.5	2	2	2	0	0	0	0	2	1	1	2	1.5	0	0	0	0	0	0	0	0	0	1	2	1.5
VIII - 2nd Sut Del/ Mgmt																																				
Steps uncompl.	0	0	0	0	0	0				0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0				0	0	0	1	3	2	0	1	0.5	0	0	0	3	2	2.5	0	0	0	2	0	1	0	1	0.5	0	0	0
IX - 2nd Knot Tying																																				
Steps uncompl.	0	0	0	0	0	0				0	0	0	3	2	2.5	3	3	3	1	1	1	1	0	0.5	0	0	0	0	1	0.5	0	0	0	0	0	0
Errors made	0	0	0	0	0	0				0	0	0	0	5	2.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
X - 3rd Anch Prep																																				
Steps uncompl.	0	0	0	1	0	0.5				1	1	1	1	1	1	4	4	4	1	0	0.5	4	4	4	0	1	0.5	1	1	1	1	1	1	1	0	0.5
Errors made	0	0	0	0	0	0				0	0	0	1	2	1.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
XI - 3rd Sut Del/ Mgmt																																				
Steps uncompl.	1	0	0.5	0	0	0				0	0	0	0	0	0	2	2	2	0	0	0	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0				1	0	0.5	1	0	0.5	0	0	0	0	0	0	0	0	0	0	1	0.5	2	1	1.5	0	1	0.5	0	0	0
XII - 3rd Knot Tying																																				
Steps uncompl.	0	0	0	0	0	0				0	0	0	0	0	0	3	3	3	0	0	0	3	3	3	0	0	0	1	1	1	0	0	0	0	0	0
Errors made	0	1	0.5	0	0	0				0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	1.5	1	1	1	1	1	1	0	1	0.5
XIII - Eval Repair																																				
Steps uncompl.	0	0	0	0	0	0				0	0	0	0	0	0	2	2	2	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0				0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Portal / Dx Time (B4 Bank)			6			13			26			14			17			36			5			19			9			34			23			7
Bankart Repair Time			50			59			15			45			75			35			31			68			47			39			32			33
Total Time (Dx+Rx)			56			72			41			59			92			71			36			87			56			73			55			40
									DNC																											

(continued)

Table 5. Continued

Video #	33A	33B	33 ave	73A	73B	73 ave	123A	123B	123 ave	53A	53B	53 ave	83A	83B	83 ave	113A	113B	113 ave
Rating Pairs																		
Steps completed (45)	38	37	37.5	35	35	35	*12*	*14*	*13*	41	42	41.5	30	32	31	26	22	24
Errors made (77)	3	5	4	1	0	0.5	*5*	*4*	*4.5*	4	8	6	10	16	13	5	5	5
Sentinel errors	1	3	2	0	0	0	*1*	*1*	*1*	1	2	1.5	3	4	3.5	1	1	1
Steps completed																		
Agreement		39			41						44			40			41	
Disagreement		6			4						1			5			4	
Step IRR		0.87			0.91						0.98			0.89			0.91	
Errors made																		
Agreement		75			76						68			59			69	
Disagreement		2			1						9			18			8	
Error IRR		0.97			0.99						0.88			0.77			0.9	
Total score (S+E)																		
Agreement		114			117						112			99			110	
Disagreement		8			5						10			23			12	
Total IRR		0.93			0.96						0.92			0.81			0.9	

Video #	13A	13B	13 ave	93A	93B	93 ave	43A	43B	43 ave	63A	63B	63 ave	23A	23B	23 ave	103 A	103 B	103 ave
Rating Pairs																		
Steps completed (45)	38	39	38.5	19	25	22	40	42	41	36	37	36.5	37	40	38.5	40	40	40
Errors made (77)	1	5	3	11	10	10.5	5	3	4	11	4	7.5	2	7	4.5	4	5	4.5
Sentinel errors	0	1	0.5	2	1	1.5	2	0	1	5	3	4	0	1	0.5	1	1	1
Steps completed																		
Agreement		40			37			39			40			42			40	
Disagreement		5			8			6			5			3			5	
Step IRR		0.89			0.82			0.87			0.89			0.93			0.89	
Errors made																		
Agreement		71			68			73			70			72			73	
Disagreement		6			9			4			7			5			4	
Error IRR		0.92			0.88			0.95			0.91			0.93			0.95	
Total score (S+E)																		
Agreement		113			105			112			110			114			113	
Disagreement		9			17			10			12			8			9	
Total IRR		0.93			0.86			0.92			0.90			0.93			0.93	

NOTE. Data in italics indicate incomplete procedure.

anch, anchor; caps, capsule; DNC, did not complete; Dx, diagnostic; E, errors; eval, evaluate; inf, inferior; instabl, instability; IRR, inter-rater reliability; mgmt, management; Rx, treatment; S, steps; sut del, suture delivery.

reintroduced the arthroscope and created a standardized Bankart lesion from the 2- to 6-o'clock position and 6 to 9 mm deep (medial). While the lesion was clearly delineated, the capsulolabral tissues were not mobilized to any additional extent. Before the study, we attempted to use different techniques to create an anteroinferior capsular detachment from the glenoid (Bankart lesion) in a cadaveric shoulder specimen (i.e., a long-handled scalpel blade, a hook-tip cautery, and manual dissection with an elevator). It was determined that the most consistent lesion could be created using a liberator/elevator along with a mallet to provide a gentle, controlled impact force to the elevator to effectively "sculpt" the Bankart pathology. This method optimized preservation of the integrity of the capsulolabral tissues for subsequent repair.

Once the Bankart lesion was completed, the arthroscope was withdrawn and reintroduced by the subject surgeon, who operated for the duration of the procedure. A continuous video recording was made, beginning with the first arthroscopic view of the joint from the posterior portal and ending with the final examination of the completed procedure by the surgeon. In calculating the total operating time of the Bankart repair procedure for the subject surgeon, the segment of time required by the master faculty surgeon to create the Bankart pathology was subtracted from the total absolute running time. No time limit was imposed on the performance of the procedure in the cadaveric specimen.

Video Reviewer Training

Once the construction of the metrics for an ABR was completed and face and content validity verified,[7] a final version of a score sheet was formatted. Ten AANA master/associate master faculty surgeons (none belonging to the experienced group from this study) formed the panel of reviewers designated to score the videos. This group included the 3 members (R.L.A., R.K.N.R., R.A.P.) of the group who, in conjunction with a consultant experimental psychologist (A.G.G.), created the arthroscopic Bankart metric "definitions" (Table 1). The 10 reviewers were randomly assigned to form 5 fixed pairs, which remained constant throughout the scoring of all videos. Reviewer training was initiated with an 8-hour in-person meeting, during which time each metric was studied in detail. Multiple video examples of live patient cases were shown to illustrate each particular metric. Videos of the patients in both the lateral decubitus and beach-chair orientations were represented. Discussion helped to clarify how each step and error were to be scored, including the nuances and conventions to be used. Several weeks later, full-length practice videos 1 and 2 (one each in the lateral decubitus and beach-chair orientation) were sent to and independently

Table 6. Copernicus Cadaver - Experienced

			12			22			112			122			32			42			52			62			82			92
Video #	12A	12B	ave	22A	22B	ave	112A	112B	ave	122A	122B	ave	32A	32B	ave	42A	42B	ave	52A	52B	ave	62A	62B	ave	82A	82b	ave	92A	92B	ave
I - Portals																														
Steps uncompl.	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	1	0.5	2	2	2	0	0	0	2	0	1
Errors made	1	1	1	1	2	1.5	1	1	1	0	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
II - Instabl Asses.																														
Steps uncompl.	1	3	2	0	2	1	1	1	1	0	1	0.5	1	2	1.5	3	3	3	1	1	1	2	0	1	1	1	1	2	0	1
Errors made	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0
III - Caps/Gen Prep																														
Steps uncompl.	2	3	2.5	4	3	3.5	2	1	1.5	0	1	0.5	1	1	1	2	2	2	0	1	0.5	1	1	1	3	0	1.5	4	3	3.5
Errors made	0	0	0	0	1	0.5	2	3	2.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IV - 1st Inf Anch Prep																														
Steps uncompl.	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	1	0.5
Errors made	0	0	0	0	0	0	0	0	0	2	1	1.5	0	0	0	1	1	1	0	0	0	0	0	0	0	1	0.5	0	1	0.5
V - 1st Sut Del/Mgmt																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0	1	0	0.5	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5
VI - 1st Knot Tying																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1
Errors made	0	1	0.5	0	0	0	0	1	0.5	0	0	0	1	2	1.5	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0
VII - 2nd Anch Prep																														
Steps uncompl.	1	1	1	0	0	0	1	1	1	1	1	1	0	0	0	1	1	1	1	0	0.5	1	1	1	0	1	0.5	1	1	1
Errors made	0	0	0	2	1	1.5	0	0	0	0	0	0	0	0	0	0	1	0.5	0	2	1	1	1	1	0	0	0	0	0	0
VIII - 2nd Sut Del/Mgmt																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	1	0.5	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0.5
IX - 2nd Knot Tying																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
X - 3rd Anch Prep																														
Steps uncompl.	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	0	0.5	1	1	1	1	1	1
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	0	2	1	0	1	0.5	0	0	0	0	0	0	1	1	1	0	0	0
XI - 3rd Sut Del/Mgmt																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	1	1	1	0	0	0	0	0	0	0	1	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
XII - 3rd Knot Tying																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	1	1	1	2	2	2
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
XIII - Eval Repair																														
Steps uncompl.	0	0	0	0	0	0	0	0	0	1	0	0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Errors made	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Portal/Dx Time (B4 Bank)			8			9			5			6			7			2			18			1			6			3
Bankart Repair Time			36			30			26			28			63			17			24			28			27			23
Total Time (Dx+Rx)			44			39			31			34			70			19			42			29			33			26
Rating Pairs																														

(continued)

Table 6. Continued

			12			22			112			122
Video #	12A	12B	ave	22A	22B	ave	112A	112B	ave	122A	122B	ave
Steps completed (45)	37	35	36	39	38	38.5	40	41	40.5	41	40	40.5
Errors made (77)	1	4	2.5	4	5	4.5	5	5	5	2	1	1.5
Sentinel errors	0	1	0.5	1	1	1	3	2	2.5	1	0	0.5
Steps completed												
Agreement		40			42			44			42	
Disagreement		5			3			1			3	
Step IRR		0.89			0.93			0.98			0.93	
Errors made												
Agreement		73			74			71			76	
Disagreement		4			3			6			1	
Error IRR		0.95			0.96			0.92			0.99	
Total score (S+E)												
Agreement		113			116			115			118	
Disagreement		9			6			7			4	
Total IRR		0.93			0.95			0.94			0.97	

			32			42			52			62			82			92
Video #	32A	32B	ave	42A	42B	ave	52A	52B	ave	62A	62B	ave	82A	82b	ave	92A	92B	ave
Steps completed (45)	36	37	36.5	36	36	36	43	42	42.5	38	41	39.5	38	38	38.5	30	36	33
Errors made (77)	3	9	6	3	6	4.5	0	3	1.5	1	1	1	1	2	1.5	1	2	1.5
Sentinel errors	0	2	1	1	3	2	0	0	0	1	1	1	0	0	0	1	1	1
Steps completed																		
Agreement		41			45			42			43			38			39	
Disagreement		4			0			3			2			7			6	
Step IRR		0.91			1			0.93			0.95			0.84			0.87	
Errors made																		
Agreement		69			74			74			77			76			74	
Disagreement		8			3			3			0			1			3	
Error IRR		0.9			0.96			0.96			1			0.99			0.96	
Total score (S+E)																		
Agreement		110			119			116			120			114			113	
Disagreement		12			3			6			2			8			9	
Total IRR		0.9			0.97			0.95			0.98			0.93			0.93	

anch, anchor; caps, capsule; Dx, diagnostic; E, errors; eval, evaluate; inf, inferior; instabl, instability; IRR, inter-rater reliability; mgmt, management; Rx, treatment; S, steps; sut del, suture delivery.

scored by each of the 10 reviewers, and the scores were tabulated. During 2 subsequent 2-hour group phone conferences, the differences and discrepancies among all reviewers were compared and discussed, seeking conformity in scoring. In addition, each designated pair of reviewers conducted 1 to 3 additional phone conferences to analyze the specific instances in which the 2 of them scored particular events differently. Subsequently, all reviewers scored practice videos 3 and 4, and the results were again tabulated (each patient orientation was again represented). The scores for each of the 5 designated pairs of reviewers were compared for the second set of practice videos. In only 1 of 10 comparisons (2 videos × 5 reviewer pairs) did the inter-rater reliability (IRR) (Table 1) calculation (as described later) fall below an acceptable level of 0.8,[14] with an IRR value of 0.76.

Video Scoring

The AANA research coordinator randomly assigned each of the 22 full-length study videos of the experienced and novice surgeons performing an ABR on a cadaveric shoulder to a single pair of reviewers. Other than the research coordinator and the study consultant, the primary investigators and all video reviewers remained blinded to the source of the video being reviewed. Each of the 22 videos was independently reviewed and scored by the 2 members of an assigned pair of reviewers. Each step and error metric was scored as either yes or no, designating whether the specific event was or was not observed to have occurred by the reviewer. In addition to scoring steps and errors, each event characterized as DNTT was scored. There was no limit to the number of individual instances DNTT could be scored, with each occurrence simply tallied as a single error event. The score sheet also contained a box for specific reviewer comments for each metric. The 2 individual scores from a pair of reviewers were averaged to obtain the overall score for each step, error, or DNTT event. Reviewer scores are tabulated in Table 5 (novice) and Table 6 (experienced). The number of uncompleted steps are listed for each phase rather than completed steps because they provided a more meaningful assessment of performance. Each phase was tabulated separately to determine which phases best discriminated between novice and experienced surgeon performance. The total time in minutes taken by the participant to perform the diagnostic and procedural components was documented for each video, beginning with the first view of the arthroscope from the posterior portal and ending with the withdrawal of the arthroscope after examination of the completed review. The time required by the designated faculty member to create the Bankart lesion was subtracted from the total running time of the video. The total number of steps, errors and sentinel errors are listed for the entire procedure. In addition, the score agreement or disagreement between the specific pair of reviewers was tabulated seperately for the steps, errors, and total of steps and errors, and were used to calculate IRR correlations (as described in the "Statistical Methods" section).

Performance Benchmark

Prior research has used the mean performance (metric based) of a group of experienced or expert operators to objectively define proficiency.[15-19] To assess and ensure performance homogeneity among the group of experienced surgeons for establishment of an accurate benchmark, their performances were converted to *Z* scores. The standard score (more commonly referred to as a "Z score") is a very useful statistic because it (1) creates the ability to calculate the probability of a score occurring within a normal distribution and (2) enables one to compare more precisely the scores from 2 individuals on a standardized scale. It is then possible to objectively and transparently determine whether and precisely how much a given subject's score is above or below the mean of their peers. In the pre-study design phase of the project, a stipulation was made by the primary investigators to remove the data from analysis for an experienced surgeon performing more than 2 SDs from the mean of the experienced group for any of the 4 assessments: steps, errors, sentinel errors, and time. Any such performance by a participant in the experienced group would be deemed an "outlier," and the scores would be removed from the analysis so as not to skew the establishment of the reference benchmark. This stipulation applied to both the previously reported shoulder model construct validity study[13] and the cadaveric construct validity study presented in this report.

Incomplete Repairs

It was prospectively determined that if a surgeon did not substantially complete the 3-anchor Bankart repair, his or her partial scores would be removed from the analysis. This policy was established because if only a portion of a procedure was performed, it would not have been possible to accurately estimate or extrapolate how many errors, sentinel errors, or instances of DNTT the surgeon may have enacted had he or she completed the entire procedure. Furthermore, no estimate of total time for the procedure would be sensible.

Statistical Methods

For each of the 13 separate phases of the procedure, the numbers of "uncompleted steps" and "errors made" were tabulated and the scores for the 2 reviewers averaged (Tables 5 and 6). These data were used to determine which of the procedural phases showed the greatest differences in performance when comparing the experienced and novice surgeons (1-factor analysis of variance) (IBM SPSS statistical software program; IBM, Armonk, NY). Furthermore, for the entire procedure, the total numbers of steps completed, errors made, and sentinel errors enacted were also averaged for the pair of reviewers.

The 2 raw score sheets were compared for each of the individual steps (45 steps in total) and the number of "agreements" tabulated (either both reviewers documented that a step was performed or both scored the step as not being completed). In addition, the number of "disagreements" in scoring steps was tabulated (one of the reviewers indicated that the step had been completed and other scored that the step had not been completed). The IRR for the steps was calculated according to the following formula: Number of agreements/Number of agreements + Number of disagreements.

In a similar manner, there was either agreement or disagreement in the 2 scores for each of the potential errors (77 errors in total). The IRR for error scoring was calculated in the same manner as that for the steps. Finally, the IRR for scoring the entire procedure was calculated using both the step and error agreements/disagreements for the entire procedure (122 in total). An acceptable IRR is equal to or greater than 0.80.[14]

Results

Participants

Two groups were compared in their performance of an ABR on a cadaveric shoulder. The entire group of master and associate master instructors, serving as faculty for an AANA Resident Course, chose to participate and comprised the experienced group (n = 10). The faculty, all fellowship trained in arthroscopy or sports medicine, averaged over 16 years in clinical practice, with each having routine experience in performing arthroscopic shoulder techniques. All faculty members have been recognized nationally by the AANA for their talent and ability to teach and communicate shoulder arthroscopy skills to trainees. The novice group (n = 12) comprised 11 PGY 5 and 1 PGY 4 orthopaedic resident volunteers (from a total of 44 orthopaedic residents registered for the weekend course) who elected to participate in the study and perform an ABR on a cadaveric shoulder. These volunteers had previously registered for an AANA Resident Course with no prior knowledge of the Bankart repair assessment protocol. Other than their year in training, no information regarding their arthroscopic experience or surgical skill was obtained.

Cadaveric Specimens

One specimen for each of the study groups was rejected and replaced. One specimen for the novice group was large with very noncompliant tissues, substantially restricting the anterior working space and making the use of instruments difficult. One specimen from the experienced group was arthritic, which limited the ability to distract the humeral head from the glenoid; thus visualization and the use of instruments from the posterior portal were unacceptably restricted. Each of these 2 specimens was replaced by an acceptable fresh cadaveric shoulder.

IRR Assessments

The IRR calculations across each of the assessments were strong. Twenty-one videos could be scored completely. One novice completed only a single-anchor repair during the entire duration of the procedure, which provided incomplete data. The mean IRR for the paired scoring of the 21 videos for procedural steps was 0.91 (range, 0.82 to 1.00), and the mean IRR for errors including DNTT events was 0.93 (range, 0.77 to 1.00 [the value 0.77 for the error IRR calculation for 1 video was the single instance that fell below 0.80 among the 63 IRR calculations]). The mean IRR for the total of steps and errors was 0.93 (range, 0.81 to 0.98).

Outlier Performance

One novice subject completed only 13 of 45 steps and a 1-anchor repair before failing to progress and electively terminating the procedure. During that time, 4.5 errors and 1 sentinel error were created (the average of the pair of designated reviewers). Inclusion of the data for the relatively small number of errors enacted during the partial repair would bias and understate the average number of errors for the novice group. As a result, all scores for this outlier were removed. The number of steps could theoretically have been used because this number accurately reflected the number of steps that were actually completed, but we elected to use none of the data from this subject's limited repair.

Before the data for complete repairs were analyzed, score profiles were examined for significantly atypical performance in the experienced group. One subject in the experienced group took dramatically longer than the subject's colleagues to perform the procedure, primarily because of substantial difficulties with suture delivery and management. This subject required 63 minutes to complete the Bankart repair in comparison with the colleagues' mean of 29.5 minutes (SD, 12.6 minutes). For the experienced surgeon outlier, Z equaled 2.61 with an associated probability value of .005, which indicated that the difference in performance from this surgeon's peers was highly significant. Consistent with the prospectively established policy of removing an experienced subject's scores if his or her performance was more than 2 SDs from the group mean, this subject's data were removed from further statistical comparisons between the groups. Thus, the data for one subject from each of the two groups was removed from the comparative statistical analysis.

Experienced Group and Novice Group Comparisons

Comparisons were made separately for steps and errors for each of the 13 phases of the Bankart procedure, as well as the summary data for steps, errors, and sentinel errors (Table 5 shows data for the novice group and Table 6 shows data for the experienced group). The phases of anchor preparation/insertion, suture delivery and management, and knot tying were repeated for each of the 3 anchors. The 3 sets of data for the 3 similar phases (1 for each of 3 anchors) were combined. The novice surgeons made significantly more objectively scored overall procedure errors (Fig 1) than the experienced surgeons (5.68 errors for novice surgeons *v* 2.95 errors for experienced surgeons, $P = .013$). Not only did the novice surgeons make more errors but they also showed greater performance variability, as shown by the considerably larger standard deviation score (3.51 *v* 1.85). The greatest difference in the mean number of errors made occurred during the suture delivery and management phases of the procedure, which was statistically significant (1.95 errors for novice surgeons *v* 0.45 errors for experienced surgeons, $P = .024$) (Fig 1). The novice surgeons also made more sentinel errors than the experienced surgeons (1.5 *v* 0.95), but this difference was not statistically significant.

The most common errors and sentinel errors are shown in Table 7, with those errors common to all of the 3 anchors being summed. With respect to regular errors, failure to maintain intra-articular position of the cannulas was frequently observed for the novice group. Both groups experienced occasional instances of anchor pullout, the experienced group somewhat more often than the novice group. By far the most common sentinel error enacted by the novice group was improper introduction of the suture-delivery device into the capsule at or above the anchor hole, resulting in failure to achieve retentioning of the capsuloligamentous tissues superiorly. Damage (laceration) of the intact labrum during attempts to mobilize the capsulolabral tissues was also notably more common among the novice group compared with the experienced group. Overall, the novice surgeons also completed fewer steps than the experienced surgeons (35.04 *v* 38.15), but this difference was not statistically significant ($P = .187$).

Figure 2 shows the mean amount of time both groups of subjects took to perform the procedure. The novice

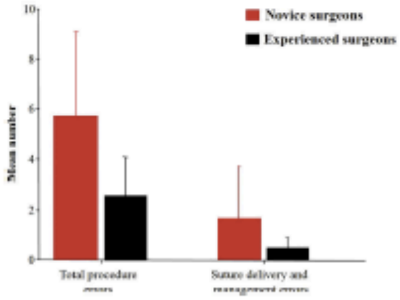


Fig 1. Mean total errors enacted by novice and experienced surgeon groups ($P = .013$) and mean suture delivery and management errors enacted by novice and experienced surgeon groups ($P = .024$).

surgeons took significantly more time to perform the repair than the experienced group (45.5 minutes [SD, 14.95 minutes] for novice surgeons *v* 25.9 minutes [SD, 5.33 minutes] for experienced surgeons; $P < .001$).

Table 7. Common Errors and Sentinel Errors for Novice and Experienced Subjects

	Frequency of Errors by Group	
	Novice	Experienced
Error		
Failure to maintain intra-articular position of the posterior cannula	5	1
Failure to maintain intra-articular position of the mid-anterior cannula	4	0
Failure to maintain intra-articular position of the anterosuperior cannula	1	2
Elevate capsulolabral tissue from the glenoid neck	5	0
Anchor fails to remain securely fixed at appropriate depth*	4	6
Offloading of suture anchor	1	0
Sentinel error		
Lacerate intact capsulolabral tissue	4	1
Capsular penetration is at or superior to anchor hole*	13	3
Entry of completed tunnel lies outside safe zone of 0-3 mm from bony glenoid rim	1	1
Failure to create and maintain indentation of capsule or labral tissue	2	1

*Errors that were common to the 3 anchors were summed.

Discussion

Novice Versus Experienced Surgeon Performance

This study shows robust construct validity for the use of the arthroscopic Bankart procedure metrics with a cadaveric shoulder. The Bankart metrics are both precise (high IRR) and accurate (able to distinguish between novice and experienced surgeon performance).[20] Overall, experienced surgeons performed better than novice orthopaedic surgeons when evaluated using an objectively assessed and blinded review of video-recorded operative performance. Although the objectively assessed performance of the experienced surgeons was better than that of the novice surgeons across all of the measures, the metrics that best distinguished the 2 groups were procedure errors, particularly suture management errors. Operative time was also significantly different. We are unaware of previous similar studies using a detailed metric-based assessment of a complete surgical procedure with which to compare and contrast our results.

Tool Development

At the outset of the series of investigations termed the AANA Copernicus Initiative (a paradigm shift from the apprenticeship model to one of PBP training), we sought to study the effectiveness of PBP training for surgical skills. This investigation required the development and validation of 3 separate, specific tools to conduct the analysis. The first component to be created was a "metric tool" (steps, errors, and sentinel errors) for a specific procedure (an ABR was selected). This metric tool was shown to have face and content validity.[7] Second, a "training tool" (a shoulder model simulator coupled with the ABR metrics) was shown to have construct validity for an ABR with the ability to distinguish between experienced and novice surgeon performance.[13] Lastly, the current study shows construct validity for the cadaveric shoulder (coupled with the ABR metrics) as a valid "assessment tool" for comparing the performance of different surgeons.

Inter-rater Reliability

The very high IRR for the scores from the reviewer pairs for the entire group of metrics (0.92) is reflective of the clarity and precision of the arthroscopic Bankart metrics drafted, as well as the thorough training of the 10 reviewers. The ability to score the steps and errors consistently is essential to obtain a reliable measure of the surgeon's performance and skill level for a particular procedure.

Shoulder Simulator Model Versus Cadaveric Shoulder

For the prior study undertaken to assess construct validity for the shoulder model simulator and Bankart metrics, surgeons in the experienced group made 63% fewer errors, committed 79% fewer sentinel errors, and performed the procedure in 42% less time than those in

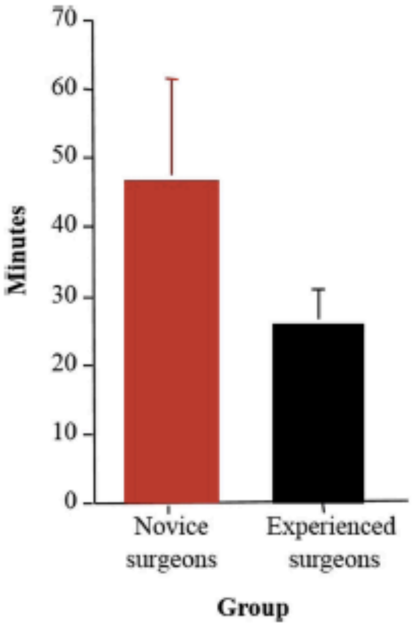


Fig 2. Mean time (in minutes) taken by novice and experienced surgeon groups to complete 3 suture anchor Bankart repair ($P < .001$).

the novice group (all differences being significant). The greatest difference in errors between the groups involved anchor preparation and insertion, suture delivery and management, and knot tying. In the current study using a cadaveric shoulder, experienced surgeons made 50% fewer errors and performed the procedure in 44% less time (both differences being significant). The number of sentinel errors was significantly less for the experienced group in the model validation study, and sentinel errors were also committed less frequently by the experienced surgeons in this cadaveric evaluation, although the difference was not statistically significant. With respect to specific phases of the procedure, the greatest discriminator in both investigations was for the phases of suture delivery and management. This finding is not surprising because the steps involved in those phases are among the most challenging for the Bankart repair. The number of steps performed did not differentiate between the novice and experienced groups in either the simulator model or cadaver studies. This result is not unexpected because the intent and effort to perform each of the steps predominantly reflect a familiarity and knowledge of the steps necessary to perform the procedure. Overall, in both studies, the experienced group showed less performance variability than the novice surgeons as demonstrated by their smaller standard deviation scores.

The benchmark was established based on the mean performance of the group of experienced surgeons. For this cadaveric study, it included completion of a 3-anchor Bankart repair with no more than 3 total errors and no more than 1 sentinel error. For the similar previous study on the simulator model using the identical metrics, the one difference in the benchmark was that no more than 4 (instead of 3) total errors were permitted.

Novice and Experienced Outliers

The partial data from 1 novice surgeon were removed from the analysis because the surgeon was not able to complete the suturing and knot tying for the first anchor before electively terminating the procedure because of the inability to make progress. A relatively large number of errors were enacted (mean, 4.5) (Table 5) for the portion of the procedure performed, but it was not possible to accurately estimate or extrapolate to the total number of errors that might have been created had the entire procedure been completed. The total number of errors enacted would likely have been substantial had the procedure been completed. Thus the average total numbers of errors and sentinel errors are likely to have been substantially understated for the novice group. A relatively small number of steps were accomplished (mean, 13) (Table 5), and had the data for this novice been included in that analysis, they would have affected the average total number of steps that the novices completed. Given that the diagnostic portion of the procedure took over 25 minutes for this subject, the overall time for completion of the 3-anchor repair would also likely have been much longer. Because the data for the majority of the analysis were incomplete, it was not possible to include any of this surgeon's performance data in the analysis.

One of the issues that emerged during this study, and indeed in one of our previous studies,[13] was atypical performance of 1 experienced subject. Atypical performance is an important issue as it relates to establishing benchmarks, which may have considerable implications for trainee progression. Since PBP training was first introduced and validated in 2002,[15] the average, or mean, performance of experienced operators has been used as the performance benchmark that trainees must meet and demonstrate before being allowed to progress in their training.[15-19] If an experienced individual's performance score was dramatically worse than his or her peers' scores and if this score were to be included in the establishment of the benchmark, the reference level would clearly be lowered. The lowering of the performance threshold (benchmark) could have important patient safety implications. For example, in a study on bariatric surgery, it was found that surgeons performing at the lower end of the performance range had significantly poorer outcomes than surgeons performing at the upper end.[21] This was one of the first studies to quantitatively link objectively assessed surgeon skill performance with patient outcomes.

The criteria for removing outlier data from the group being used to create a performance benchmark must be established before the data are collected and should be objective, transparent, and fair. At the outset, before conducting this study, we identified and discussed the possibility of encountering atypical experienced surgeon performance. The core group of 4 primary investigators agreed that the resolution to this potential issue would be to remove all of a subject's scores from subsequent analysis if it could be unambiguously established that the subject's performance was statistically atypical (>2.0 SDs from the group mean). The experienced individual participating in this study performed considerably worse than that for operative time (2.61 SDs from the mean).

Performance Errors

The enactment of errors is emerging as one of the most important indicators of skill for operative performance.[13] While an individual may be able to perform all of the correct steps in an acceptable order with the appropriate instruments and score very well on those parameters, he or she may still perform the steps poorly. Procedural errors are operative behaviors that deviate from optimal performance. These metrics are a reliable measure of performance quality and are likely to be the most sensitive assessment tool in the evaluation of operative performance and safety.[22] Although simulation-based education and the resulting transfer of training will influence other performance parameters as well, such as procedural time, the greatest impact of such a training strategy appears to be on limiting performance errors.[22] It is therefore a necessity, at the outset, that the performance characterization of "deviations from optimal performance" (errors) must be particularly robust and well validated. It also implies that error performance assessed using a less rigid Likert scale (Table 1) (global rating scale) may result in a less focused approach to minimizing errors because the deviations from optimal performance have been less clearly defined.[23,24] A Likert-type scale is a method of ascribing a quantitative value to qualitative data to make the data amenable to statistical analysis. Likert scales (often with a range from 1 to 5 or from 1 to 7) are typically constructed with responses (opinion) around a neutral option (e.g., "suture delivery was 1, awkward, . . . 3, effective, . . . 5, highly efficient") and were originally designed to assess a range of respondent attitudes.[25] Given the inherent subjectivity in this method of attempting to rate objective performance, it may be difficult to obtain acceptable levels of IRR (≥80%) in the scoring of events.[24] In contrast, the approach to the assessment of performance used in this study uses precise definitions of performance and simply requires the video reviewer to determine whether the specific event did or did not occur. This binary approach to the measurement of performance has been shown to facilitate the reliable scoring of metric-based performance units across a variety of functions from skills training[18,26-28] at different experience levels.[29,30] It has also been shown to considerably enhance assessment reliability levels in comparison with Likert-scale scoring.[23]

The effectiveness of a deliberate practice, proficiency-based training curriculum using simulation relies on a clear and specific identification not only of the proper steps that the trainee should perform but also of what the trainee is doing wrong and how to prevent or correct his or her error. Other advantages of creating comprehensive procedure characterizations and explicit operational step and error definitions exist. Detailed metrics provide very clear guidance for the construction of simulation training platforms, specifying exactly what the simulator should be capable of emulating and, more importantly, measuring.[31] Comprehensive procedure characterization is challenging and time-consuming the first time it is undertaken and requires robust validation of all of the performance metrics. With experience, however, this methodology is considerably easier to apply to subsequent characterizations of different procedures by the same group.

Limitations

A limitation of this study relates to the use of cadaveric specimens for the arthroscopic Bankart lesion creation and repair. The specimens lacked some uniformity in the integrity of the capsule and labrum, soft-tissue compliance, shoulder mobility/distractibility, and bulk of the extra-articular tissues. In addition, although specific parameters were used for the creation of the Bankart lesions (i.e., 2- to 6-o'clock position on the glenoid rim and 6 to 9 mm deep/medial), the lesions could not be made absolutely uniform. Further variability existed in the presence of coexisting pathology (arthritis, synovial proliferation, rotator cuff partial tears, and so on). The "acceptability criteria" for the specimens (listed earlier) were used to minimize the impact of this potential problem.

An additional limitation of this study is that there was no confirmation that those serving as master/associate master surgeons and being representative of the "experienced" group possessed a specified level of expert skill in performing an arthroscopic Bankart procedure. Nevertheless, the individual surgeons so identified have been recognized by the AANA as valuable educators either from lecture presentations with videos exhibiting skilled shoulder arthroscopy techniques or from repeated experience teaching in an arthroscopic laboratory setting with the ability to demonstrate and teach each of the key components of a Bankart procedure. Thus, "experienced" rather than "expert" is a reasonable description of the group. Similarly, other than identifying the year in training, no additional information was obtained to determine the extent of the residents' experience (novice group) with arthroscopic shoulder surgery (i.e., the number of arthroscopy/sports medicine rotations previously completed, the number of shoulder arthroscopic surgical procedures in which they served as assistant surgeons, and so on). Even with these data, accurate knowledge of the level of skill possessed by an individual resident would not be possible. Thus, although the arthroscopic skill sets of the subjects are representative of their respective groups and experience, those skills are highly likely to be somewhat heterogeneous.

We acknowledge that only a single operative procedure was analyzed for each of the subject surgeons. It is possible that data averaged over several procedures would be somewhat different from those obtained in this study. Cost and time considerations made the performance of a single ABR on a cadaveric shoulder most feasible. Finally, it should be noted that the participants in each group had no prior specific knowledge of the metrics to be scored in the review of their procedure, and we suspect that the experienced surgeons, in particular, might have performed and scored differently (better) for certain non-crucial parts of the procedure (e.g., the diagnostic steps at the beginning) had they been familiar with the metrics to be evaluated.

Conclusions

The assessment tool composed of validated arthroscopic Bankart metrics coupled with a cadaveric shoulder accurately distinguishes the performance of experienced from novice orthopaedic surgeons. A benchmark based on the mean performance of the experienced group includes completion of a 3-anchor Bankart repair, while enacting no more than 3 total errors and 1 sentinel error.

References

1. Committee on Quality of Health Care in America. *To err is human: Building a safer health system.* Washington, DC: National Academies Press, 2000;196-197.
2. Smith R. All changed, changed utterly. *BMJ* 1998;316:1917-1918.
3. Bridges M, Diamond DL. The financial impact of teaching surgical residents in the operating room. *Am J Surg* 1999;177:28-32.
4. Barden CB, Specht MC, McCarter MD, et al. Effects of limited work hours on surgical training. *J Am Coll Surg* 2002;195:531-538.
5. Temple J. Time for training: A review of the impact of the European Working Time Directive on the quality of training. Available at http://hee.nhs.uk/healtheducationengland/files/2012/08/Time-for-training-report.pdf. Accessed March 2, 2014.
6. Bell RH Jr, Biester TW, Tabuenca A, et al. Operative experience of residents in US general surgery programs: A gap between expectation and experience. *Ann Surg* 2009;249:719-724.
7. Angelo RL, Ryu RKN, Pedowitz RA, Gallagher AG. Metric development for an arthroscopic Bankart procedure: Assessment of face and content validity. *Arthroscopy* 2015;31:1430-1440.
8. Morgan CD, Bodenstab AB. Arthroscopic Bankart suture repair: Technique and early results. *Arthroscopy* 2010;26:819-820.
9. Streubel PN, Krych AJ, Simone JP, et al. Anterior glenohumeral instability: A pathology based surgical treatment strategy. *J Am Acad Orthop Surg* 2014;22:283-294.
10. Waterman BR, Burns TC, McCriskin B, et al. Outcomes after Bankart repair in a military population: Predictors for surgical revision and long-term disability. *Arthroscopy* 2014;30:172-177.
11. Shibata H, Gotoh M, Mitsui Y, Kai Y, Nakamura H, et al. Risk factors for shoulder re-dislocation after arthroscopic Bankart repair. *J Orthop Surg Res* 2014;9:53.
12. Ryu RK. Arthroscopic approach to traumatic anterior shoulder instability. *Arthroscopy* 2003;19:94-101.
13. Angelo RL, Pedowitz RA, Ryu RKN, Gallagher AG. The Bankart performance metrics combined with a shoulder model simulator create a precise and accurate training tool for measuring surgeon skill. *Arthroscopy* 2015;31:1639-1654.
14. American Educational Research Association. Standards for educational and psychological testing. Available at www.apa.org/science/programs/testing/standards.aspx. Accessed January 12, 2014.
15. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: Results of a randomized, double-blinded study. *Ann Surg* 2002;236:458-464.
16. Gallagher AG, Ritter EM, Champion H, et al. Virtual reality simulation for the operating room: Proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg* 2005;241:364-372.
17. Ahlberg G, Enochsson L, Gallagher AG, et al. Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *Am J Surg* 2007;193:797-804.
18. Van Sickle K, Ritter EM, Baghai M, et al. Prospective, randomized, double-blind trial of curriculum-based training for intracorporeal suturing and knot tying. *J Am Coll Surg* 2008;207:560-568.

19. Gallagher AG, O'Sullivan GC. *Fundamentals of surgical simulation; principles & practices*. London: Springer Verlag, 2011.
20. Rossi MJ, Lubowitz JH, Provencher MT, Poehling GG. Precision versus accuracy: A case for common sense. *Arthroscopy* 2012;28:1043-1044.
21. Birkmeyer JD, Finks JF, O'Reilly A, et al. Surgical skill and complication rates after bariatric surgery. *N Engl J Med* 2013;369:1434-1442.
22. Gallagher AG, Seymour NE, Jordan-Black JA, et al. Prospective, randomized assessment of transfer of training (ToT) and transfer effectiveness ratio (TER) of virtual reality simulation training for laparoscopic skill acquisition. *Ann Surg* 2013;257:1025-1031.
23. Gallagher AG, O'Sullivan GC, Leonard G, Bunting BP, McGlade KJ. Objective structured assessment of technical skills and checklist scales reliability compared for high stakes assessments. *ANZ J Surg* 2014;84:568-573.
24. Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: Rigorous science for the assessment of surgical education and training. *Surg Endosc* 2003;17:1525-1529.
25. Likert R. A technique for the measurement of attitudes. *Arch Psychol* 1932;140:44-53.
26. McClusky DA, Gallagher A, Ritter EM, et al. Virtual reality training improves junior residents' operating room performance: Results of a prospective, randomized, double-blinded study of the complete laparoscopic cholecystectomy. *J Am Coll Surg* 2004;199:73 (abstract, suppl).
27. Van Sickle K, Smith B, McClusky DA, et al. Evaluation of a tensiometer to provide objective feedback in knot-tying performance. *Am Surg* 2005;71:1018-1023.
28. Van Sickle K, Gallagher AG, Smith CD. The effect of escalating feedback on the acquisition of psychomotor skills for laparoscopy. *Surg Endosc* 2007;21:220-224.
29. Neary PC, Boyle E, Delaney CP, et al. Construct validation of a novel hybrid virtual-reality simulator for training and assessing laparoscopic colectomy; results from the first course for experienced senior laparoscopic surgeons. *Surg Endosc* 2008;22:2301-2309.
30. Nicholson WJ, Cates CU, Patel AD, et al. Face and content validation of virtual reality simulation for carotid angiography: Results from the first 100 physicians attending the Emory NeuroAnatomy Carotid Training (ENACT) program. *Simul Healthc* 2006;1:147-150.
31. Gallagher AG. Metric-based simulation training to proficiency in medical education:- What it is and how to do it. *Ulster Med J* 2012;81:107-113.

Appendix 8 (Chapter 6)

Objective Assessment of Knot-Tying Proficiency With the Fundamentals of Arthroscopic Surgery Training Program Workstation and Knot Tester

Robert A. Pedowitz, M.D., Ph.D., Gregg T. Nicandri, M.D., Richard L. Angelo, M.D., Richard K. N. Ryu, M.D., and Anthony G. Gallagher, Ph.D., D.Sc.

Purpose: To assess a new method for biomechanical assessment of arthroscopic knots and to establish proficiency benchmarks using the Fundamentals of Arthroscopic Surgery Training (FAST) Program workstation and knot tester. **Methods:** The first study group included 20 faculty at an Arthroscopy Association of North America resident arthroscopy course (19.9 ± 8.25 years in practice). The second group comprised 30 experienced surgeons attending an Arthroscopy Association of North America fall course (17.1 ± 19.3 years in practice). The training group included 44 postgraduate year 4 or 5 orthopaedic residents in a randomized, prospective study of proficiency-based training, with 3 subgroups: group A, standard training (n = 14); group B, workstation practice (n = 14); and group C, proficiency-based progression using the knot tester (n = 16). Each subject tied 5 arthroscopic knots backed up by 3 reversed hitches on alternating posts. Knots were tied under video control around a metal mandrel through a cannula within an opaque dome (FAST workstation). Each suture loop was stressed statically at 15 lb for 15 seconds. A calibrated sizer measured loop expansion. Knot failure was defined as 3 mm of loop expansion or greater. **Results:** In the faculty group, 24% of knots "failed" under load. Performance was inconsistent: 12 faculty had all knots pass, whereas 2 had all knots fail. In the second group of practicing surgeons, 21% of the knots failed under load. Overall, 56 of 250 knots (22%) tied by experienced surgeons failed. For the postgraduate year 4 or 5 residents, the aggregate knot failure rate was 26% for the 220 knots tied. Group C residents had an 11% knot failure rate (half the overall faculty rate, $P = .013$). **Conclusions:** The FAST workstation and knot tester offer a simple and reproducible educational approach for enhancement of arthroscopic knot-tying skills. Our data suggest that there is significant room for improvement in the quality and consistency of these important arthroscopic skills, even for experienced arthroscopic surgeons. **Level of Evidence:** Level II, prospective comparative study.

Knot tying is an essential skill for proficiency in arthroscopic surgery.[1,2] Arthroscopic knot tying is difficult to teach and to assess objectively. At this time, most trainees are assessed by visual inspection of arthroscopic knots, either by direct view or by an arthroscopic image. Hanypsiak et al.[3] recently showed that even experienced practicing surgeons are relatively inconsistent when it comes to arthroscopic knot tying. Technical inconsistency could have a negative impact on surgical outcomes.

In the laboratory setting, arthroscopic knots (more accurately, the suture loops created after knot tying) are usually tested with expensive material testing devices that allow sophisticated variation of load magnitude, cyclic versus single pull, loop preload, and load application rate.[2,4-10] However, these devices are not practical for day-to-day education of residents and fellows or for continuing medical education of practicing surgeons. It would be advantageous to have a cost-effective and relatively simple-to-use tester for objective assessment of knot performance, as opposed to knot appearance.

In the teaching laboratory, knot-tying skills are generally developed using knot-tying boards under direct visualization with the trainee's eyes looking directly at the hands, suture, and associated surgical

From the University of California (R.A.P.), Los Angeles, California; University of Rochester School of Medicine and Dentistry (G.T.N.), Rochester, New York; Evergreen Orthopaedic Clinic (R.L.A.), Seattle, Washington; The Ryu Hurvitz Orthopedic Clinic (R.K.N.R.), Santa Barbara, California, U.S.A.; and University College Cork (A.G.G.), Cork, Ireland.

The authors report the following potential conflict of interest or source of funding: R.A.P. receives support from Virtamed. R.L.A. receives support from DePuy Mitek. R.K.N.R. receives support from MedBridge, Mitek, and Rotation Medical.

Received December 2, 2014; accepted June 17, 2015.

Address correspondence to Robert A. Pedowitz, M.D., Ph.D., 4043 Irving Pl, Ste 304, Culver City, CA 90232, U.S.A. E-mail: robertpedowitz@gmail.com

0749-8063/141016/$36.00

http://dx.doi.org/10.1016/j.arthro.2015.06.021

instruments. However, in the clinical setting, arthroscopic knots and backup hitches are created outside of the body, delivered through an arthroscopic cannula, and then tensioned within the joint, with visualization provided by a 2-dimensional video screen. This combination requires an integrated chain of complex psychomotor skills that are performed in 3-dimensional space with mostly binocular cues.[11] Such skills are best acquired and rehearsed in a gradual and systematic fashion.[12,13]

The Fundamentals of Arthroscopic Surgery Training (FAST) Program is a collaborative initiative of the Arthroscopy Association of North America (AANA), the American Academy of Orthopaedic Surgeons, and the American Board of Orthopaedic Surgery. The FAST Program offers a basic arthroscopic motor skills curriculum with associated teaching modules to facilitate core training in orthopaedic surgery. It is logical to achieve a baseline level of technical proficiency, if possible, before operating on patients.[14]

The FAST Program curriculum was developed after task deconstruction of basic arthroscopic skills (available, with open access, at http://www.aana.org/FASTProgram/FASTProgramSurgicalSkillsContent.aspx). The FAST workstation (Sawbones, Vashon Island, WA) was custom designed for training of these skills. The system allows for initial practice under direct visualization, then advances to triangulation through simulated portals under direct visual control, and finally moves to skill rehearsal through simulated portals using a video camera, with the direct surgeon view eliminated. The purposes of this study were to assess the FAST knot tester and to establish benchmarks for knot-tying proficiency using this system. Our hypothesis was that the FAST knot tester would facilitate objective, accurate, and immediate mechanical assessment of knot performance.

Methods

For all groups in this study, 5 consecutive knots were created by each subject on the FAST workstation using No. 2 FiberWire (Arthrex, Naples, FL) under dry, room-temperature conditions through a 7-mm plastic cannula. Each subject created an arthroscopic knot of his or her choice, backed up by 3 reversed half-hitches on alternating posts. Each suture was labeled, well away from the knot and suture loop, for later identification. The 5 knots were gently placed within a labeled plastic bag for each subject and set aside for subsequent analysis using the FAST knot tester.

Faculty Reference Group ($n = 20$)

The first group (faculty) was composed of 20 experienced surgeons teaching at a dedicated AANA resident arthroscopy skills course at the Orthopedic Learning Center (Rosemont, IL). This expert group reported clinical practice experience of 19.9 ± 8.25 years and performed 381 ± 150 arthroscopies per year.

Resident Comparison Groups ($n = 44$)

Orthopaedic surgery residents (postgraduate years 4 and 5) participated in a randomized, prospective study of proficiency-based training at the Orthopedic Learning Center (the AANA Copernicus Study, described in detail in a separate publication[15] that did not include specific information about knot-tying performance, benchmarks, and associated methodology). Residents were divided into 3 subgroups. Group A included 14 residents who were instructed on knot-tying skills using standard educational methodology during a regular AANA resident arthroscopy course. Standard educational methodology included didactic instruction, faculty demonstration, practice with rope, and progression to knot tying using suture around a metal hook and then through an arthroscopic cannula, all under direct visualization. When group A participants felt ready for testing, each resident used the FAST workstation and USB camera system to create 5 arthroscopic knots in sequence without interval feedback.

Group B included 14 residents who received similar didactic instruction to group A, but they were also allowed to practice knot-tying skills using the FAST workstation/USB camera system until they were ready to create 5 knots for later testing. Group C comprised 16 residents who received the same didactic instruction and practiced with the workstation/USB camera setup. However, group C residents were allowed to use the FAST knot tester after each knot was tied, providing immediate performance feedback during the practice phase, until they were ready to create 5 test knots. Practice time before knot testing was not a controlled variable for the resident study groups.

Surgeon Reference Group ($n = 30$)

Thirty surgeons volunteered to create 5 knots using the FAST workstation and USB camera setup at the 2013 AANA fall course. For the purposes of setting the benchmark for resident proficiency (as described later), we only used the faculty from the Copernicus course as the reference group. We were surprised at the high knot failure rate among the Copernicus faculty, so we pursued an additional cohort of practicing surgeons (whether faculty or non-faculty surgeons) from the AANA fall course. We thought that this would, at a minimum, represent arthroscopic surgeons in practice, and we believed that observations of 50 practicing surgeons would enhance overall confidence in the observations. The knots were tested later with no feedback provided to the surgeon during knot creation. All participants were in clinical practice for at least 1 year (maximum, 40 years). The group had an average

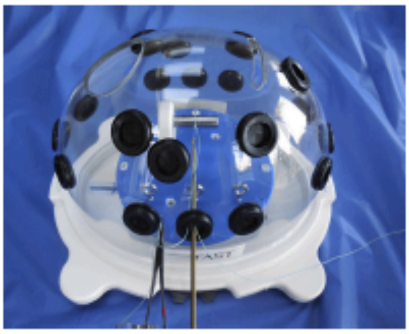


Fig 1. Fundamentals of Arthroscopic Surgery Training (FAST) base station with knot-tying mandrel and lucent dome for skills rehearsal under direct visualization.

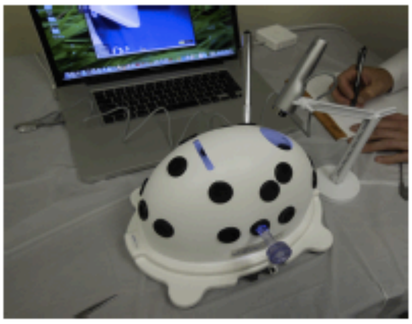


Fig 2. Fundamentals of Arthroscopic Surgery Training (FAST) workstation with opaque dome and light-emitting diode penlight. An inexpensive USB camera is directed at the knot-tying mandrel, and the image is displayed on a laptop computer. This arrangement simulates arthroscopic visualization.

practice experience of 17.1 ± 19.3 years (mean ± SD). Ten surgeons self-reported as course faculty, 12 self-reported as attendees, and 8 surgeons did not indicate whether they were faculty or attendees.

Study Participants

Each subject was verbally informed about the purpose of the study, and all volunteers were assigned a unique identification number. All test knots were labeled using subject identification numbers (subject names excluded). The study protocol was reviewed by the Western Institutional Review Board (Puyallup, WA) and deemed exempt.

FAST Workstation

The FAST workstation is composed of a base unit, which accommodates various snap-in teaching modules that complement the FAST Program curriculum. The base station can be used for basic skills practice under direct visualization without the need for triangulation. Two snap-in dome units allow for skills rehearsal either under direct visualization with the lucent dome (Fig 1) or with video imaging using the opaque dome (Fig 2). Both domes have multiple, identically positioned access holes that mimic portal positions and geometries of knee and shoulder arthroscopy.

The FAST workstation has a horizontally positioned, smooth, stainless steel knot-tying mandrel for practice with suture (Fig 1). The circumference of the knot-tying mandrel matches the first marked position on the conical loop sizer of the FAST knot tester (Fig 3). The loop sizer is calibrated in 1-mm increments to measure up to 5 mm of expansion relative to the knot-tying mandrel. On the basis of prior literature,[3,16] 3 mm of loop expansion or more was deemed to indicate knot failure. This is considered an amount of suture loop elongation that might be associated with biological healing failure at the tendon-bone interface after rotator cuff repair.[14,17]

Visualization Protocol

To mimic clinical conditions, the FAST workstation was designed for use with an inexpensive USB camera mounted on a stand (Fig 2). For this study, a

Fig 3. Conical loop sizer of Fundamentals of Arthroscopic Surgery Training (FAST) knot tester. The first mark indicates zero loop expansion, and each subsequent mark reflects 1 mm of additional loop expansion compared with the knot-tying mandrel.

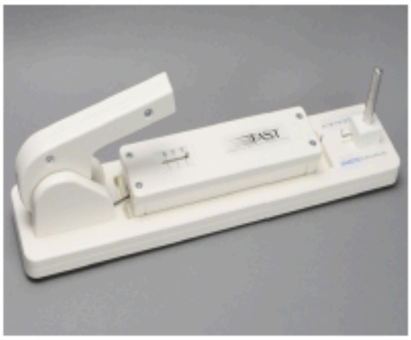


Fig 4. Fundamentals of Arthroscopic Surgery Training (FAST) knot tester. The suture loop is positioned on the 2 tines, and the handle allows for controlled application of 15 lb of longitudinal load.

high-resolution Point 2 View camera (IPEVO, Sunnyvale, CA) was mounted on its base station and connected to a laptop computer, which provided the image on the video monitor (Fig 2). The camera was directed at the knot-tying mandrel through a hole in the opaque workstation dome, which forced subjects to look at the image on the laptop screen during knot tying. The camera was set at 2× screen magnification using the IPEVO software. Illumination within the dome was augmented using a disposable light-emitting diode penlight (PLED23A; Energizer, St Louis, MO), although we found that ambient room light was generally sufficient because of the high sensitivity of the USB camera.

FAST Knot Tester

The FAST knot tester is composed of a rigid base with an integrated spring for application of linear tension to a suture loop (Fig 4). The tester was designed to apply 15 lb of tension (60 N) based on published theoretical modeling of relevant clinical forces because direct in vivo measurements of postoperative suture tension are not available. Burkhart et al.[18] estimated that 60 N would be the maximal force per suture that might be created by muscle contraction after a balanced suture anchor repair of a medium-sized rotator cuff tear. Peak loads would potentially be greater with an unbalanced repair or with an abrupt event, such as a postoperative fall. High-strength No. 2 sutures are relatively stiff, and the material can withstand loads greater than 300 N before rupture.[4,8] Therefore the 60-N load of the FAST knot tester was selected to assess knot performance as opposed to suture performance.

After a knot is created on the workstation, the loop is gently slipped off the knot-tying mandrel and transferred to the conical loop sizer, which gives a baseline measurement of the suture loop. The loop is then transferred to the 2 tines of the knot tester. One of the tines is solidly attached to the rigid base. Tension is applied by an actuator handle connected by a calibrated spring to the other tine. A force gauge allows the user to apply 15 lb of axial tension to the suture loop. We chose 15 seconds of steady force application because pilot studies indicated that significant additional loop expansion did not occur beyond this time point (in fact, most of the loop expansion was observed within a few seconds of force application). After 15 seconds of static load, the actuator handle is released, and the suture loop is removed from the tines and transferred back to the conical loop sizer. The calibrated markings are used to assess final loop size compared with initial loop size. A loop created with a perfect knot would therefore have zero baseline difference from the knot-tying mandrel and zero expansion after load application.

Reproducibility of Load Application

A digital force scale with a maximum load capacity of 30 lb (HS-30; CCI Scale Company, Clovis, CA) was used to measure the consistency of load application created by the FAST knot tester. Two sets of 10 measures each were acquired by independent observers. One observer applied 15 lb of load with the actuator handle for 15 seconds while looking at the force scale of the FAST knot tester. The other observer recorded axial load using the digital force scale after 15 seconds. The force scale was connected by a rigid metal link to the suture tine of the knot tester. The digital force scale was re-zeroed after each pull of the actuator handle. The roles of the 2 observers were reversed during the second set of measurements. By use of this protocol, mean force application measured by the digital force scale was 15.03 lb, with an SD of 0.05 lb (expressed as the SD for 2 independent sets of 10 measures).

Statistical Analysis

Logistic regression analysis was used to compare relative differences between knot-tying performance of the 3 resident trainee groups and faculty knot-tying performance. Statistical significance was considered at $P < .05$.

Results

Performance data from the AANA Copernicus Study participants and from the AANA fall course subjects are presented in Table 1 and Table 2, respectively. This information is stratified according to the number of knots that failed (defined as $\geq$3 mm of loop expansion) after application of 15 lb of static load for 15 seconds. Of the 20 Copernicus course faculty, 12 had 0 knots fail. Four faculty had 2 knot failures, 2 faculty had 3 knot failures, and 2 faculty had all 5 of their knots fail on the FAST knot tester. Overall, 24% of the faculty knots

Table 1. Knot Tying at Arthroscopy Association of North America Residents' Copernicus Course

	Faculty (n = 20)	Group A Residents (n = 14)	Group B Residents (n = 14)	Group C Residents (n = 16)
Years in practice, mean ± SD (range)	19.9 ± 8.3 (4-32)			
Knot performance, n				
0 of 5 failed	12	3	3	11
1 of 5 failed	0	3	5	2
2 of 5 failed	4	7	2	2
3 of 5 failed	2	0	1	1
4 of 5 failed	0	1	1	0
5 of 5 failed	2	0	2	0
No. of knots that failed	24 of 100 (24%)	21 of 70 (30%)	26 of 70 (35%)	9 of 80 (11%)

NOTE. Knot failure was defined as 3 mm of loop expansion or greater with application of 15 lb of load for 15 seconds.

were considered failures. Only 1 faculty surgeon tied 5 consecutive "perfect" knots (zero loop expansion compared with the knot-tying mandrel at baseline and zero loop expansion after 15 lb of load application).

A similarly high rate of knot failure was noted in the second group of experienced surgeons (Table 2). Overall, for these practicing surgeons, 21% of the knots were noted to be failures. Five of the 30 participants had 3 of 5 knot failures. Taking in aggregate all knots tied by faculty and practicing surgeons at the 2 courses, this study found that 56 of 250 knots (22%) were deemed failures by mechanical testing.

Overall, the knot failure rate was 26% for the 220 knots that were tied by the orthopaedic surgery residents (Table 1). However, the group C residents, who were allowed to use the knot tester for feedback during the training experience, had an overall 11% knot failure rate, which was exactly one-half the knot rate of the Copernicus course faculty.

Data from the Copernicus course faculty were used to create a proficiency benchmark for "passing" resident performance. This passing benchmark was applied to the group C Copernicus course residents (proficiency-based progression group; full details were provided by Angelo et al.[15]). On the basis of the Copernicus faculty data, we defined a proficiency benchmark of less than or equal to 2 knot failures out of 5 knot attempts as a passing grade. When we apply these same thresholds to the surgeons in practice at the fall course (Table 2), with the bar set at less than or equal to 2 knot failures out of 5 attempts, 5 of the 30 surgeons (17%) missed the passing mark.

Relative to the Copernicus course faculty, groups A and B were more likely to have their knots fail, but these differences were not statistically significant (P = .384 for group A *v* faculty and P = .07 for group B *v* faculty by logistic regression analysis). In contrast, residents in group C were more than twice as likely to have their knots pass in comparison with the faculty reference group (odds ratio, 2.84), and this difference was statistically significant (P = .013). Logistic regression analysis was also used to compare the relative differences between the 3 trainee subgroups, using standard training (group A) as the reference. There was no statistical difference between groups A and B (odds ratio, 0.725; P = .372). In contrast, group C residents were almost 4 times as likely to have their knots pass as group A (odds ratio, 3.857; P = .002).

If we apply a proficiency threshold of no more than 2 knot failures out of 5 trials (as described earlier), 6 of the 44 orthopaedic residents (14%) fell below the passing bar. Of note, 15 of 16 group C residents (94%) exceeded the passing threshold; this was the best performance for any subgroup in this study.

Table 2. Knot Tying at 32nd Arthroscopy Association of North America Fall Course

	Faculty (n = 10)	Surgeon Attendees (n = 12)	Faculty or Attendee Not Defined (n = 8)	Total for All Participants (n = 30)
Year in practice, mean ± SD (range)	20.5 ± 7.6 (3-30)	14.6 ± 12.4 (1-40)	19.0 ± 9.3 (9-32)	17.1 ± 19.3 (1-40)
Knot performance, n				
0 of 5 failed	5	3	4	12
1 of 5 failed	3	3	3	9
2 of 5 failed	1	3	0	4
3 of 5 failed	1	3	1	5
4 of 5 failed	0	0	0	0
5 of 5 failed	0	0	0	0
No. of knots that failed	8 of 50 (16%)	18 of 60 (30%)	6 of 40 (15%)	32 of 150 (21%)

NOTE. Knot failure was defined as 3 mm of loop expansion or greater with application of 15 lb of load for 15 seconds.

Discussion

The FAST workstation and FAST knot tester facilitated direct, objective measurements of arthroscopic knot-tying performance. Overall, 22% of knots tied by practicing surgeons "failed" using this testing protocol. A proficiency-based progression training protocol resulted in improved resident knot-tying skills (11% knot failure rate) compared with standard training methodology.

The FAST Program provides core education for orthopaedic surgery residents, fellows, and practicing surgeons who wish to develop and enhance their arthroscopic motor skills. The FAST Program is intended to create and enhance robust psychomotor skills, right from the outset of training. It is relatively difficult to correct bad surgical habits once they are firmly established.[12,13] Training of many surgical skills can be performed outside of the operating room in an efficient and cost-effective manner that maximizes educational quality and eliminates patient morbidity. There has been a significant shift toward structured simulation training, including recent simulation mandates by the American Board of Orthopaedic Surgery and the Orthopaedic Residency Review Committee of the Accreditation Council for Graduate Medical Education.[19-21] The FAST Program was designed to satisfy these educational mandates for arthroscopic surgery with a cost-effective, practical, modular system.

In the traditional approach to training, operative skills were acquired in an apprentice model of training that meant that learning was serendipitous. Resident experience was affected by when residents were on duty, which patients and procedures they encountered, and who was supervising and mentoring them. It also relied on learning by repeated practice.[22] The proficiency-based progression approach to training, afforded by technologies such as FAST, encourages a "deliberate practice" approach.[13,23] This means that the trainee receives objective metric-based feedback on his or her performance proximate to the measured task, thus augmenting the learning experience for him or her. Seeing knots slip when pulled is a very impactful learning experience, even for very experienced surgeons.

For decades, orthopaedic educators have been using knot-tying boards with rope and suture to train arthroscopic knot tying. In most cases, proficiency assessment has been based on subjective, visual observation of the knot-tying process and the visual appearance of the surgical knot.[24-34] However, what is most important, in terms of surgical outcome, is knot performance as opposed to knot appearance. A "pretty" knot has no clinical value if it does not hold under physiological loads.

Before this study, biomechanical assessment of sutures and knots has generally been restricted to analyses with sophisticated material testing devices (from MTS Systems [Eden Prairie, MN], Instron [Norwood, MA], and so on). These devices are quite expensive, and they are impractical for day-to-day training applications. However, they do have advantages for complex load-application paradigms, including cyclic load protocols. Nonetheless, we thought that it would be advantageous to create a very simple and inexpensive knot tester that could be used on the educational front lines. The FAST knot tester was not designed to be a sophisticated bioengineering research tool.

It should be emphasized that the level of load application (15 lb) for the FAST knot tester was specifically selected for testing of high-strength No. 2 sutures used commonly during many arthroscopic procedures. The knot tester could be adapted with different spring loads to assess the performance of other suture materials. Objective proficiency benchmarks could be established for various sutures under specific performance conditions. This strategy is relevant to training and objective assessment of surgical knot tying across the medical spectrum because knot tying is a pervasive technical skill requirement for most procedural specialties. The FAST approach to measurement and quality assurance, by achieving performance benchmarks before training progression, fits well with the recommendations of the Institute of Medicine report on graduate medical education.[35] The Institute of Medicine proposed that medical education should move away from training that is process driven (i.e., time in training, number of procedures completed, duration of rotation) to an "outcome"-driven enterprise.[36] This means that trainees would be required to demonstrate a benchmark performance level.[12]

We were quite surprised by the high incidence of knot failures in this study for course faculty and surgeons in practice. After our data collection, Hanypsiak et al.[3] published similar observations in their study of 73 expert orthopaedic arthroscopists who tied 365 individual knots with No. 2 FiberWire suture. In their study, surgeons created knots under direct visualization, without magnification or video control, and the knots were tested using a sophisticated electromechanical dynamic testing system. The authors observed significant variations between surgeons and between knot configurations. Perhaps even more important, they concluded that "considerable variation and inconsistencies in knot strength exist between arthroscopic knots of the same type tied by the same surgeon."[3] Individual subject performance inconsistency was also noted in our study for knots created by experts under video control through arthroscopic cannulas. The observations of Hanypsiak et al. and the findings of our study are extremely important because technical consistency is a hallmark of surgical proficiency and patient safety.

We used the performance data for our expert faculty surgeons to create objective proficiency benchmarks that could be applied to resident and fellow training. Proficiency benchmarks must be reasonable and achievable. It would not make sense to set benchmarks that are unachievable for a high percentage of competent, experienced surgeons. Given the relatively high incidence of knot failure for our experienced surgeons, we defined a proficiency benchmark of no more than 2 of 5 knot failures to achieve a passing score for residents. Of course, surgeons should strive for technical perfection, with 0 knot failures, and we observed that level of high performance for some of our expert subjects. However, our data suggest that many arthroscopic surgeons (even experienced and expert surgeons) have substantial opportunities for improvement. Such opportunities are facilitated by direct, objective, and immediate performance feedback.

On the basis of the Copernicus faculty data, we defined a proficiency benchmark (a passing grade) of less than or equal to 2 knot failures out of 5 knot attempts. If the threshold had been set to no more than 1 failure in 5 attempts, 8 of our own Copernicus faculty (40%) would have fallen below the passing bar (Table 1). We thought it was important to avoid unrealistic or unachievable proficiency benchmarks for the residents, so we selected the more lenient proficiency standard of no more than 2 knot failures out of 5 knot attempts. For the surgeons in practice at the fall course (Table 2), if the threshold was set at less than or equal to 1 knot failure in 5 attempts, 9 of 30 surgeons in practice (30%) would not have passed. These data further support the use of the more lenient proficiency standard for training purposes.

Overall, the knot failure rate was 26% for the 220 knots that were tied by the orthopaedic surgery residents (Table 1). Surprisingly, this failure rate was not dramatically different than the overall failure rate for our faculty and surgeons in practice. However, the group C residents (who were allowed to use the knot tester for feedback during the training experience) had an overall 11% knot failure rate, which was significantly better than the Copernicus course faculty (P = .02). We were impressed by the strong performance of group C, the proficiency-based progression subgroup (Table 1). These residents could assess their performance based on direct proximate feedback, make adjustments in knot-tying technique, and then see for themselves whether their performance had improved. This approach appears to have resulted in substantial enhancement of this group's performance.

On the basis of our observations and the recent findings of Hanypsiak et al.,[3] we believe that it would be very challenging to objectively assess knot performance using video review of arthroscopic procedures. Some overt suture failures are easily observed. For example, it is visually obvious when sutures break or become entangled or when there is an overly loose suture loop that does not indent soft tissue. However, our findings suggest that some "visually acceptable" knots may fail under relevant mechanical loads even in the hands of very experienced surgeons.

Limitations

One of the limitations of this study was that we did not afford opportunities for self-directed performance feedback to our faculty surgeons or to the practicing surgeons at the AANA fall course. This study did not involve a homogeneous population of faculty and practicing surgeons, with prior clinical experience ranging from 1 to 40 years of practice.

We did not assess transfer of motor skills to the clinical situation, nor did we examine same-subject test/retest consistency. Maximum practice time before knot testing was not a controlled variable. This study was not designed to compare performance differences according to knot type because we wanted each subject to select his or her own base knot based on personal preference and experience. Previous research has looked at biomechanical performance variation as a function of knot type, and it was not our purpose to examine this question. During the study design phase, we recognized and discussed the implications of variation in the base knot of each study participant. We wanted each subject, particularly the faculty and experienced surgeons, to pick the base knot that he or she would be most comfortable tying. We did not want to impose a particular knot choice because we were concerned that individual performance could be adversely affected by asking subjects to tie knots with which they were unaccustomed, thereby introducing greater data variability. This study was not designed to cross-correlate knot performance with knot "appearance." These are important study limitations and represent opportunities for further work.

Conclusions

The FAST workstation and knot tester offer a simple and reproducible educational approach for enhancement of arthroscopic knot-tying skills. Load displacement of the suture loop is a direct reflection of mechanical performance of the surgical knot. There is significant room for improvement in the quality and consistency of these important arthroscopic skills, even for experienced arthroscopic surgeons.

References

1. Mishra DK, Cannon WD Jr, Lucas DJ, Belzer JP. Elongation of arthroscopically tied knots. *Am J Sports Med* 1997;25:113-117.

2. Swan KG Jr, Baldini T, McCarty EC. Arthroscopic suture material and knot type: An updated biomechanical analysis. *Am J Sports Med* 2009;37:1578-1585.
3. Hanypsiak BT, DeLong JM, Simmons L, Lowe W, Burkhart S. Knot strength varies widely among expert arthroscopists. *Am J Sports Med* 2014;42:1978-1984.
4. Abbi G, Espinoza L, Odell T, Mahar A, Pedowitz R. Evaluation of 5 knots and 2 suture materials for arthroscopic rotator cuff repair: Very strong sutures can still slip. *Arthroscopy* 2006;22:38-43.
5. Baums MH, Sachs C, Kostuj T, Schmidt-Horlohé K, Schultz W, Klinger HM. Mechanical testing of different knot types using high-performance suture material. *Knee Surg Sports Traumatol Arthrosc* in press, available online 13 October, 2013. doi:10.1007/s00167-013-2711-9.
6. Ilahi OA, Younas SA, Alexander J, Noble PC. Cyclic testing of arthroscopic knot security. *Arthroscopy* 2004;20:62-68.
7. Ilahi OA, Younas SA, Ho DM, Noble PC. Security of knots tied with Ethibond, Fiberwire, Orthocord, or Ultrabraid. *Am J Sports Med* 2008;36:2407-2414.
8. Mahar AT, Moezzi DM, Serra-Hsu F, Pedowitz RA. Comparison and performance characteristics of 3 different knots when tied with 2 suture materials used for shoulder arthroscopy. *Arthroscopy* 2006;22:614.e1-614.e2.
9. Riboh JC, Heckman DS, Glisson RR, Moorman CT III. Shortcuts in arthroscopic knot tying: Do they affect knot and loop security? *Am J Sports Med* 2012;40:1572-1577.
10. Shah MR, Strauss EJ, Kaplan K, Jazrawi L, Rosen J. Initial loop and knot security of arthroscopic knots using high-strength sutures. *Arthroscopy* 2007;23:884-888.
11. Gallagher AG, Ritter EM, Lederman AB, McClusky DA III, Smith CD. Video-assisted surgery represents more than a loss of three-dimensional vision. *Am J Surg* 2005;189:76-80.
12. Gallagher AG, Ritter EM, Champion H, et al. Virtual reality simulation for the operating room: Proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg* 2005;241:364-372.
13. Gallagher AG, O'Sullivan GC. *Fundamentals of surgical simulation; principles & practices*. London: Springer Verlag, 2011.
14. Department of Health. A framework for technology enhanced learning. Available at http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_130924. Published 2011.
15. Angelo RL, Ryu RKN, Pedowitz RA, et al. A proficiency-based progression training curriculum coupled with a model simulator results in the acquisition of a superior arthroscopic Bankart skill set. *Arthroscopy* 2015;31:1854-1871.
16. Chan KC, Burkhart SS, Thiagarajan P, Goh JC. Optimization of stacked half-hitch knots for arthroscopic surgery. *Arthroscopy* 2001;17:752-759.
17. Lo IK, Burkhart SS, Chan KC, Athanasiou K. Arthroscopic knots: Determining the optimal balance of loop security and knot security. *Arthroscopy* 2004;20:489-502.
18. Burkhart SS, Wirth MA, Simonich M, Salem D, Lanctot D, Athanasiou K. Knot security in simple sliding knots and its relationship to rotator cuff repair: How secure must the knot be? *Arthroscopy* 2000;16:202-207.
19. Karam MD, Pedowitz RA, Natividad H, Murray J, Marsh JL. Current and future use of surgical skills training laboratories in orthopaedic resident education: A national survey. *J Bone Joint Surg Am* 2013;95:e4.
20. Pedowitz RA, Marsh JL. Motor skills training in orthopaedic surgery: A paradigm shift toward a simulation-based educational curriculum. *J Am Acad Orthop Surg* 2012;20:407-409.
21. Thomas GW, Johns BD, Marsh JL, Anderson DD. A review of the role of simulation in developing and accessing orthopaedic surgical skills. *Iowa Orthop J* 2014;34:181-189.
22. Cuschieri A. Whither minimal access surgery: Tribulations and expectations. *Am J Surg* 1995;169:9-19.
23. Krampe RT, Ericsson KA. Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *J Exp Psychol Gen* 1996;125:331-359.
24. Cheung JJ, Rojas D, Weber B, Kapralos B, Carnahan H, Dubrowski A. Evaluation of tensiometric assessment as a measure of skill degradation. *Stud Health Technol Inform* 2012;173:97-101.
25. Ching SS, Mok CW, Koh YX, Tan SM, Tan YK. Assessment of surgical trainees' quality of knot-tying. *J Surg Educ* 2013;70:48-54.
26. Horeman T, Blikkendaal MD, Feng D, et al. Visual force feedback improves knot-tying security. *J Surg Educ* 2014;71:133-141.
27. Huang E, Chern H, O'Sullivan P, et al. A better way to teach knot tying: A randomized controlled trial comparing the kinesthetic and traditional methods. *Am J Surg* 2014;208:690-694.
28. Hu Y, Tiemann D, Michael Brunt L. Video self-assessment of basic suturing and knot tying skills by novice trainees. *J Surg Educ* 2013;70:279-283.
29. Mashaud LB, Arain NA, Hogg DC, Scott DJ. Development, validation, and implementation of a cost-effective intermediate-level proficiency-based knot-tying and suturing curriculum for surgery residents. *J Surg Educ* 2013;70:193-199.
30. Rodrigues SP, Horeman T, Sam P, Dankelman J, van den Dobbelsteen JJ, Jansen FW. Influence of visual force feedback on tissue handling in minimally invasive surgery. *Br J Surg* 2014;101:1766-1773.
31. van Empel PJ, Verdam MG, Huirne JA, Bonjer HJ, Meijerink WJ, Scheele F. Open knot-tying skills: Resident skills assessed. *J Obstet Gynaecol Res* 2013;39:1030-1036.
32. van Leeuwen N, Trimbos JB. Strength of sliding knots in multifilament resorbable suture materials. *Gynecol Surg* 2012;9:433-437.
33. Wong IH, Denkers MR, Urquhart NA, Farrokhyar F. Systematic instruction of arthroscopic knot tying with the ArK Trainer: An objective evaluation tool. *Knee Surg Sports Traumatol Arthrosc* 2015;23:912-918.
34. Wong IH, Denkers M, Urquhart N, Farrokhyar F. Construct validity testing of the Arthroscopic Knot Trainer (ArK). *Knee Surg Sports Traumatol Arthrosc* 2015;23:906-911.
35. Committee on the Governance and Financing of Graduate Medical Education, Board on Health Care Services, Institute of Medicine, Eden J, Berwick D, Wilensky G, eds. *Graduate medical education that meets the nation's health needs*. Washington, DC: National Academies Press, 2014.
36. Asch DA, Weinstein DF. Innovation in medical education. *N Engl J Med* 2014;371:794-795.

Appendix 9 (Chapter 7)

A Proficiency-Based Progression Training Curriculum Coupled With a Model Simulator Results in the Acquisition of a Superior Arthroscopic Bankart Skill Set

Richard L. Angelo, M.D., Richard K. N. Ryu, M.D., Robert A. Pedowitz, M.D., Ph.D., William Beach, M.D., Joseph Burns, M.D., Julie Dodds, M.D., Larry Field, M.D., Mark Getelman, M.D., Rhett Hobgood, M.D., Louis McIntyre, M.D., and Anthony G. Gallagher, Ph.D., D.Sc.

Purpose: To determine the effectiveness of proficiency-based progression (PBP) training using simulation both compared with the same training without proficiency requirements and compared with a traditional resident course for learning to perform an arthroscopic Bankart repair (ABR). **Methods:** In a prospective, randomized, blinded study, 44 postgraduate year 4 or 5 orthopaedic residents from 21 Accreditation Council for Graduate Medical Education–approved US orthopaedic residency programs were randomly assigned to 1 of 3 skills training protocols for learning to perform an ABR: group A, traditional (routine Arthroscopy Association of North America Resident Course) (control, n = 14); group B, simulator (modified curriculum adding a shoulder model simulator) (n = 14); or group C, PBP (PBP plus the simulator) (n = 16). At the completion of training, all subjects performed a 3 suture anchor ABR on a cadaveric shoulder, which was videotaped and scored in blinded fashion with the use of previously validated metrics. **Results:** The PBP-trained group (group C) made 56% fewer objectively assessed errors than the traditionally trained group (group A) ($P = .011$) and 41% fewer than group B ($P = .049$) (both comparisons were statistically significant). The proficiency benchmark was achieved on the final repair by 68.7% of participants in group C compared with 36.7% in group B and 28.6% in group A. When compared with group A, group B participants were 1.4 times, group C participants were 5.5 times, and group C^{PBP} participants (who met all intermediate proficiency benchmarks) were 7.5 times as likely to achieve the final proficiency benchmark. **Conclusions:** A PBP training curriculum and protocol coupled with the use of a shoulder model simulator and previously validated metrics produces a superior arthroscopic Bankart skill set when compared with traditional and simulator-enhanced training methods. **Clinical Relevance:** Surgical training combining PBP and a simulator is efficient and effective. Patient safety could be improved if surgical trainees participated in PBP training using a simulator before treating surgical patients.

Changing work patterns and a reduction in hours available for training[1,2] have forced the surgical community to consider new methods to augment and enhance training. Surgical simulation-based training, first proposed by Satava[3] in 1993 as a potential solution to this problem, has developed in sophistication and adoption among the medical education and training communities.[4,5] The first prospective, randomized, double-blind clinical trial of simulation-based training for the operating room showed that surgical residents trained to a "proficiency benchmark" (Table 1) on a virtual reality simulator made

From ProOrtho Clinic (R.L.A.), Kirkland, Washington; The Ryu Hurvitz Orthopedic Clinic (R.K.N.R.), Santa Barbara, California; Private Practice (R.A.P.), Culver City, California; Tuckahoe Orthopedics (W.B.), Richmond, Virginia; Southern California Orthopedic Institute (J.B., M.G.), Van Nuys, California; Mid-Michigan Physicians (J.D.), East Lansing, Michigan; Mississippi Sports Medicine and Orthopaedic Clinic (L.F., R.H.), Jackson, Mississippi; Westchester Orthopaedics (L.M.), White Plains, New York, U.S.A.; and ASSERT, University College Cork (A.G.G.), Cork, Ireland;

The authors report the following potential conflict of interest or source of funding: R.L.A. receives support from DePuy Mitek. R.K.N.R. receives support from MedBridge, Rotation Medical, and Mitek. J.B. receives support from ConMed. M.G. receives support from Mitek, Smith & Nephew, and Wolters Kluwer. R.H. receives support from DePuy.

Received August 18, 2014; accepted July 9, 2015.

Address correspondence to Richard L. Angelo, M.D., 12911 120th Ave NE, Ste H-210, Kirkland, WA 98072, U.S.A. E-mail: rlangeloortho@comcast.net

0749-8063/14705/$36.00
http://dx.doi.org/10.1016/j.arthro.2015.07.001

1854 Arthroscopy: The Journal of Arthroscopic and Related Surgery, Vol 31, No 10 (October), 2015: pp 1854-1871

Table 1. Glossary

	Definition
Construct validity	A type of evidence that supports that specific test items identify the quality, ability, or trait they were designed to measure
Content validity	An estimate (opinion) by experts of the validity of a testing instrument based on a detailed examination of the contents of the test items
Damage to non-target tissue	Iatrogenic damage to tissues not intended to be addressed in the specific step (e.g., articular cartilage damage)
Definition	A definite, distinct, and clear objective characterization providing an accurate and reliable identification of whether an event was or was not observed to occur
Delphi panel (modified)	A structured communication technique originally developed as a systematic, interactive forecasting method that relies on the opinion of a panel of experts; in modified form, experts answer queries or vote in 2 or more rounds (cycles) on the appropriateness of the metric-based operational definitions of detailed aspects of procedure performance with the goal of achieving consensus; voting is not anonymous
Error	A deviation from optimal performance
Face validity	An estimate (opinion) by experts who review the content of an assessment or tool to see if it seems appropriate and relevant to the concept it purports to measure
Inter-rater reliability	The extent of agreement between 2 raters on the occurrence of a series of observed events; it ranges between 0, no agreement, and 1.0, complete agreement
Likert-scale assessment	A method of assessing a range of attitudes; it ascribes a quantitative value to qualitative data (e.g., 1, extremely unlikely; 3, likely; and 5, highly likely)
Metric	A standard of measurement of quantitative assessments used for objective evaluations to make comparisons or to track performance
Operational definition	Terms used to define a variable or event in terms of a process (or set of validation tests) needed to determine its existence, quantity, and duration
Performance metric	The features determining the accomplishment of a given task measured against preset known standards of accuracy and completeness
Procedure phase	A group or series of integrally related events or actions that, when combined with other phases, constitute a complete operative procedure
Proficiency/proficiency benchmark	A specific level of performance defined by a quantitative score (benchmark) on a standardized test or other form of assessment
Proficiency-based progression	A training program that dictates that skill performance be demonstrated, to a predetermined benchmark level, by the trainee before advancement to more complex techniques
Sentinel error	An event or occurrence involving a serious deviation from optimal performance during a procedure that either (1) jeopardizes the success or desired result of the procedure or (2) creates significant iatrogenic insult to the patient's tissues
Step	A component task, the series aggregate of which constitutes the completion of a specific procedure
Task deconstruction	To break down a procedure into constituent tasks, steps, or components

significantly fewer objectively assessed intraoperative errors when compared with the control group.[6] The reader is referred to Table 1 for a list of terms used throughout this article. Gallagher et al.[7] and Gallagher and O'Sullivan[8] have argued that simulation-based training is optimal when trainees are given precise feedback on their performance with specific recommendations for improvement, proximate to the performance. They have also suggested that trainees be provided a quantitative performance benchmark to work toward and that this benchmark should be a valid representation of a clinically important performance characteristic or task. Thus trainees must demonstrate the ability to meet specific performance benchmarks before they are permitted to progress in training (proficiency-based progression [PBP] training) (Table 1). The effectiveness of this methodology is well supported.[6,9,10]

We sought to study the effectiveness of PBP training plus simulation for the acquisition of surgical skills. For the patient with unidirectional anterior instability due primarily to a Bankart lesion (capsulolabral detachment from the anteroinferior glenoid) without significant bone loss, a suture anchor repair using 3 implants is a commonly accepted method used to obtain a successful patient outcome.[11-17] In addition, the essential components of the procedure are well outlined regardless of whether the patient is placed in the lateral decubitus or beach-chair orientation.[18,19] Thus an arthroscopic Bankart repair (ABR) was selected as the platform for this research.

The investigation into PBP training plus simulation required the development and validation of 3 separate, specific tools to conduct the analysis. The first component to be created was a "metric tool" that could objectively and accurately characterize an ABR by clearly defining the essential "steps," "errors," and "sentinel errors" (more serious errors) for a standard reference repair (Table 1). The metric tool created was shown to have face and content validity[20] using a modified Delphi panel methodology (Table 1).[21] Second, a "training tool" (a shoulder model simulator coupled with the ABR metrics) was proved to have construct validity (Table 1), showing the ability to distinguish between novice and experienced surgeon performance. A proficiency benchmark for the use of the metrics with the simulator was established.[22] Lastly, an "assessment tool" (a cadaveric shoulder coupled

Table 2. Thirteen Phases of Bankart Procedure (in Roman Numerals) and Brief Summary of 45 Steps of Procedure

I. Portals
1. Establish posterior portal
2. View posterior humeral head and extent of the Hill-Sachs when present
3. Introduce midanterior spinal needle immediately superior to the subscapularis and direct it toward the anteroinferior glenoid and labrum
4. Establish a cannula that abuts the superior border of the subscapularis near the lateral subscapularis insertion
5. Demonstrate instrument access to the anteroinferior glenoid/labrum
6. Introduce anterosuperior spinal needle at the superolateral aspect of the rotator interval and direct it toward the anterior glenoid
7. Establish an anterosuperior cannula, arthroscopic sheath, or switching stick
II. Arthroscopic instability assessment
View from posterior portal
8. View or probe the superior labral attachment onto the glenoid
9. View or probe articular surface of the cuff
10. Probe anteroinferior glenoid/Bankart pathology including rim fracture, articular defect
View from anterosuperior portal
11. View or probe the midsubstance of the anterior-inferior glenohumeral ligaments
12. View or probe the insertion of the anterior glenohumeral ligaments onto the anterior humeral neck
III. Capsulolabral mobilization/glenoid preparation
13. Elevate the capsulolabral tissue from the glenoid neck and articular margin
14. View the subscapularis muscle superficial to the mobilized capsule
15. With an instrument, grasp and perform an inferior to superior shift of the capsulolabral tissue (demonstrate restoring tension)
16. Obtain a view of the anterior glenoid neck
17. Mechanically abrade the glenoid neck
IV. Inferior anchor preparation/insertion
18. Seat the guide for the most inferior anchor hole at the inferior region of the anteroinferior quadrant
19. Drill anchor hole oblique to the glenoid articular face
20. Insert anchor
21. Test suture anchor
V. Suture delivery/management
22. Pass a cannulated suture hook or suture retriever through the capsular tissue inferior to the anchor
23. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula
VI. Knot tying
24. Deliver an arthroscopic sliding knot
25. Back up with 3 or 4 half-hitches
26. Cut suture tails
VII. Second anchor preparation/insertion
27. Seat the drill guide for the second anchor superior to the first anchor and inferior to the glenoid equator
28. Drill anchor hole oblique to the glenoid articular face
29. Insert suture anchor
30. Test anchor security by pulling on suture tails
VIII. Suture delivery/management
31. Pass a cannulated suture hook or suture retriever through the capsular tissue inferior to the suture anchor
32. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula
IX. Knot tying
33. Deliver an arthroscopic sliding knot
34. Back up with 3 or 4 half-hitches
35. Cut suture tails
X. Third anchor preparation/insertion
36. Seat the drill guide for the third anchor at or superior to the equator
37. Drill anchor hole oblique to the glenoid articular face
38. Insert suture anchor
39. Test anchor security by pulling on suture tails
XI. Suture delivery/management
40. Pass a cannulated suture hook or suture retriever through the capsular tissue
41. Pass anchor suture limb through the capsular tissue and deliver out the anterior cannula
XII. Knot tying
42. Deliver an arthroscopic sliding knot
43. Back up with 3 or 4 half-hitches
44. Cut suture tails
XIII. Procedure review
45. View and/or probe final completed repair

Table 3. Summary of 29 Different Bankart Procedure Metric Errors

Failure to maintain intra-articular position of the posterior cannula
Failure to maintain intra-articular position of the midanterior cannula
Failure to maintain intra-articular position of the anterosuperior cannula
Damage to the superior border of the subscapularis
Damage to the anterior border of the supraspinatus
Loss of intra-articular position of arthroscope/sheath or operating cannula (loss of each portal is scored only once for each phase (Roman numeral), i.e., up to a total of 3 for arthroscope + 2 portals)
Lacerate intact capsulolabral tissue (sentinel error)
Failure to maintain control of working instrument (sentinel error)
Guide is not located in the inferior region of the anteroinferior quadrant of the glenoid (for the first anchor position)
Entry of the completed tunnel lies outside safe zone of 0 to 3 mm from the bony glenoid rim (sentinel error)
Shallow undermining and deformation of articular cartilage (sentinel error)
Failure to maintain secure seating of the drill guide during anchor insertion
Implant breakage
Implant remains visibly proud (sentinel error)
Failure to insert the anchor with the inserter laser line (when present) to or beyond the laser line on the drill guide
Anchor fails to remain securely fixed within bone at the appropriate depth
Capsular penetration is at or superior to the anchor hole for anchor positions 1 and 2 (sentinel error)
Capsular penetration is not at or peripheral to the capsulolabral junction
Instrument breakage
Tearing of capsulolabral tissue
Uncorrected entanglement of shuttling device or suture
Off-loading of suture anchor
Breakage of suturing device
Failure to create and maintain indentation of the capsule or labral tissue with knot tying (sentinel error)
Visible void is present between throws of the completed primary knot (sentinel error)
Completed knot abuts articular cartilage
Visible void is present between throws of the completed half-hitches
Suture breakage
Guide is inferior to the equator of the glenoid (for the third and final anchor)

NOTE. Metric errors can be associated with multiple phases and steps of the procedure (n = 77 total errors).

with the ABR metrics) was evaluated and also shown to have construct validity (Table 1).[23]

The purpose of this study was to determine the effectiveness of PBP training using simulation both compared with the same curriculum without the proficiency requirements and compared with a traditional Arthroscopy Association of North America (AANA) Resident Course for learning to perform an ABR. We hypothesized that a training protocol coupling PBP training with a shoulder model simulator would be superior to an identical curriculum using a simulator but without the need to show proficiency, as well as to a traditional curriculum with no simulator or proficiency requirements.

Table 4. Demographic and Baseline Perceptual, Visuospatial, and Psychomotor Assessment Data for Participants

Demographic Variable	Group A	Group B	Group C	P Value
Gender, n				
Male	13	13	14	
Female	1	1	2	
PGY of training, n				
PGY 4	6	8	7	
PGY 5	8	6	9	
Right hand dominant	82%	93%	94%	
Age, yr, mean (SD)	33 (4)	31 (2)	32 (3)	.54
Perceptual assessment (PicSOr), mean (SD)	0.92 (0.07)	0.93 (0.07)	0.93 (0.03)	.89
Visuospatial assessment, mean (SD)	25 (10)	26 (7)	24 (7)	.76
Psychomotor assessments, mean (SD)				
Dominant hand				
Correct incisions	12 (3)	11 (2)	10 (4)	.45
Incorrect incisions	0.3 (0.5)	0.9 (0.1)	0.1 (0.3)	.15
Nondominant hand				
Correct incisions	10 (4)	9 (5)	9 (5)	.66
Incorrect incisions	0.1 (0.5)	0.1 (0.3)	0.2 (0.6)	.6

PGY, postgraduate year; PicSOr, pictorial surface orientation.

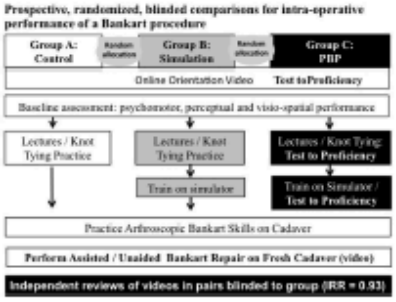


Fig 1. Study pathways for 3 separate training protocols: groups A (traditional training [control]), group B (simulation enhanced), and group C (proficiency-based progression [PBP] training plus simulation). All groups underwent baseline assessments followed by lectures on knot tying, but only group C was required to pass a cognitive examination on the metrics and to test to proficiency on knot tying. Groups B and C trained to perform an arthroscopic Bankart repair (ABR) on a model simulator, but only group C had to test to proficiency. All 3 groups trained to perform or practiced performing an ABR on a cadaver. All participants from each group performed an unaided ABR on a fresh cadaveric shoulder that was videotaped and scored in blinded fashion. (IRR, inter-rater reliability.)

Methods

Participants/Subjects

Forty-four postgraduate year (PGY) 4 or 5 residents from 21 Accreditation Council for Graduate Medical Education–approved orthopaedic residency training programs from across the United States participated. All subjects were assigned a unique identifying number that gave no indication of their PGY, residency program, or study group. The Western Institutional Review Board opined (No. 1-776362-1) that, as an educational curriculum study, this investigation was exempt from the need for full institutional review board approval [based on the criteria of 45 CFR 46.101(b)(1)]. The study protocol was registered with the National Institutes of Health (ClinicalTrials.gov No. NCT01921621) before initiation of the investigation.

Bankart Procedure Performance Metrics

The surgical residents were evaluated on their skill in performing an ABR on a cadaveric specimen. Previously validated "performance metrics" formed the basis of this evaluation and included 45 key steps with related steps grouped into 1 of 13 phases (Table 1) of the procedure[21] (Table 2). Seventy-seven potential errors to be avoided were specified (Table 3). Of these errors, 20 were designated as more serious, or sentinel, errors because either (1) the error's enactment had the potential to seriously compromise the success of the procedure or (2) the error had the potential to create significant iatrogenic damage to the shoulder. The metrics were clearly defined with beginning points and endpoints for each step, as well as precisely what did and did not constitute each potential error. All subject surgeons and faculty were provided a link on the AANA Web site to 2 full-length orientation videos, 1 each in the lateral decubitus and beach-chair orientations. Access was available 4 weeks before the course in which they were participating. Each video demonstrated all of the steps, in addition to either demonstrating or specifically identifying each of the potential

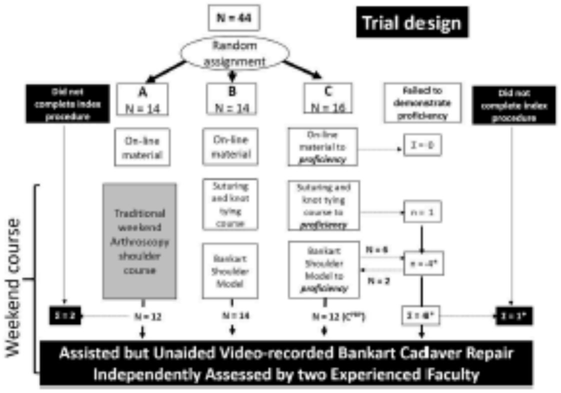


Fig 2. Participant flowchart detailing that 2 of 14 participants in group A were unable to complete the procedure, 1 group C subject could not complete knot tying to proficiency, and 4 group C residents were unable to show proficiency on the shoulder model simulator (failed 2 attempts). The group C subject who failed to show proficiency both on knot tying and on the model simulator also failed to complete the final Bankart repair. The remaining 12 subjects from group C who met all intermediate proficiency benchmarks, designated as group C^{PBP}, would be the only subjects allowed to progress per the PBP protocol. Asterisks indicate that the participant failed to demonstrate proficiency in one or more intermediate procedure components.

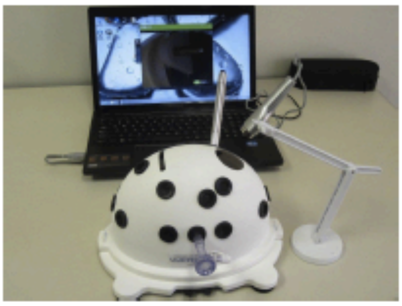


Fig 3. AANA Fundamentals of Arthroscopic Surgery Training workstation with multiple potential portal sites in an opaque dome covering a mandrel around which knots are tied through an arthroscopic cannula (inset on laptop screen). A USB camera (Ipevo, Sunnyvale, CA), directed through a window, projects an image of the inside of the dome onto a laptop computer screen.

errors (including sentinel errors) to be avoided in performing an ABR safely.

Baseline Assessments

To ensure homogeneity among the 44 subjects, all residents completed previously validated assessments of their visuospatial,[6,24] perceptual,[6,25] and psychomotor abilities[26-28] (Table 4). "Visuospatial ability" is one component of cognitive function that is related to the capacity to process and interpret visual information about where objects are in space. In this assessment, a pencil was used to create the shortest and most appropriate route between 2 specific points on a block-grid street map. A possibility of 20 correct routes between various points existed for each of two tests. The number of correct routes created in two 3-minute time periods was scored. Each registrant completed 2 tests (scores range from 0 to 40). "Perceptual ability" refers to the capacity to identify, organize, and interpret sensory information about visual depth of field. It was assessed with a computer-generated and -scored task requiring the subject to orient the axis of a spinning cone perpendicular to a designated face of a cube.[25] Each of 30 trials placed the cube in a different 3-dimensional orientation (scores range from 0.0 to 1.0). "Psychomotor ability" refers to the capacity for coordinated activity involving the arms, hands, and fingers (and, potentially, movement of the feet). Performance was assessed using a lighted endoscopic box trainer with a fixed overhead view projected onto a laptop screen. A 4 × 8–inch piece of paper had a series of 1-inch-long parallel lines drawn perpendicular to and along the long border of the sheet. Each of 30 parallel lines was separated by 10 mm. Instruments were passed through openings in the front of the box trainer. An endoscopic grasper was controlled by one hand and used to hold the paper within the box. Endoscopic scissors were controlled by the other hand and used to make cuts in the paper between the designated lines. The number of accurate paper cuts (between, but not touching, the parallel lines) able to be

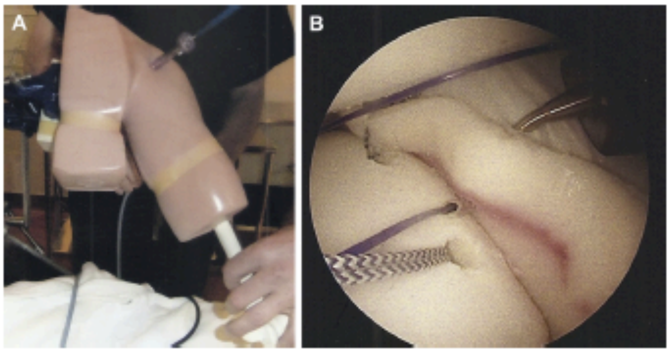


Fig 4. (A) Exterior view of the anterior aspect of a left shoulder model simulator supported in the beach-chair orientation; a hook probe is delivered through a midanterior portal. (B) Arthroscopic view from the anterosuperior portal of a left shoulder in the lateral decubitus orientation. The second anchor is in place with its sutures coursing toward the posterior portal; a suture hook device passes through the midanterior portal, pierces the labrum, and delivers a monofilament suture (to be used later for shuttling).

made in 60 seconds was tabulated. Two trials were run, one with the scissors in the dominant hand and the other with them in the nondominant hand.

Study Groups

During the weekend courses, all groups were provided similar background shoulder instability lectures that focused on indications, contraindications, and case-based examples. References to surgical technique were avoided in the lecture presentations. For each of the 3 groups, separate, dedicated, experienced Master and Associate Master AANA faculty members worked closely with that cohort of residents. The duration of training was similar for each of the 3 groups. The 3 training curricula are outlined in Figure 1. All training was conducted at the Orthopedic Learning Center (OLC) in Rosemont, Illinois.

Group A: Traditional (Control). Group A was derived from a cohort of PGY 4 and 5 residents who had independently registered for a 3-day AANA Resident Course at the OLC. The curriculum involved equal time spent on the practice of knee and shoulder procedures, which included an ABR. The registrants of the course were given the opportunity to be involved in the study but were not required to do so. A cohort of 14 residents (of a total of 48 registered residents), elected to participate in the research project, and completed an ABR at the end of their course of training.

The 14 subjects in group A served as the control group and followed an agenda typical for an AANA Resident Course (Fig 2). The curriculum included lectures on various topics including shoulder instability. Knot-tying skill was practiced under the direction of an experienced faculty member. Both sliding and non-sliding knots, as well as half-hitches to secure the primary knot, were practiced. Knot-tying boards provided the opportunity to tie knots around hooks using large cord and/or No. 2 FiberWire suture (Arthrex, Naples, FL). In addition, practice tying suture knots and delivering them down an 8.5-mm arthroscopic cannula was afforded. Finally, knots could be created and delivered through a cannula with the loop around a mandrel (smooth bar) using the AANA Fundamentals of Arthroscopic Surgery Training workstation (Fig 3). The opaque dome eliminated direct surgeon view of the knot being tied and required the use of triangulation skills. A small USB camera was directed through a window in the dome with the image viewable on a laptop computer screen. The field of view contained the knot being tied. The subject surgeon had the opportunity to spend as much time practicing knots as desired and until the participant believed he or she was proficient. A series of 5 knot trials were then tied and labeled in sequence for later testing.

Study in the laboratory using a cadaveric specimen then followed under the direct supervision of an experienced faculty member with knowledge of the Bankart metrics. Fresh-frozen specimens with a complete shoulder girdle from the scapula to the mid humerus with associated soft tissues were used. After appropriate thawing, the scapula was mounted with a clamp in the surgeon's orientation of preference (lateral decubitus *v* beach chair). Bony landmarks were identified and marked with a surgical pen, portals established, and a diagnostic arthroscopy performed. An ABR was generally the first intra-articular procedure studied. After a Bankart lesion was created, practice was conducted in the steps necessary to mobilize the capsulolabral tissue and complete a 3-anchor repair. Standard instruments for an ABR were made available. A 45° cannulated suture hook was the primary tool used to deliver sutures through the capsulolabral tissue. Single-loaded push-in anchors and a simple loop suture pattern were used. The residents were able to continue with guided study on the Bankart repair as long as desired and until they believed they were proficient. Subsequently, additional arthroscopic shoulder procedures could be electively studied as well.

Group B: Simulator. Multiple randomly selected residency program coordinators were notified of the opportunity for their residents to participate in this PBP training study. The first 32 PGY 4 and 5 residents to register became the study participants. Training for groups B and C was conducted concurrently at the OLC but during a different weekend from group A training. Of the 32 preregistered subjects, 16 were randomly assigned to 1 of 2 training protocols (group B or C) based on a computer-generated random allocation[29] (Fig 2). Of the 16 residents randomized to group B, 2 who were preregistered failed to show up for the weekend course; thus group B comprised 14 residents. The residents in group B engaged in knot-tying study and practice similar to that for group A until they believed they were proficient. A series of 5 knot trials were then tied and labeled in sequence for later testing. Group B participants were afforded the additional opportunity to train and practice an ABR using a dry shoulder model simulator (Fig 4) secured in the orientation of surgeon preference.[22] A standard equipment tower with a 30° arthroscope was provided along with all instruments necessary to complete an ABR. The simulator model was composed of a dense plastic endoskeleton palpable through simulated skin and soft tissues. Posterior, anterosuperior, and midanterior portals were created. A glenoid, humeral head, biceps, capsule, and labrum, in addition to Bankart and Hill-Sachs lesions, were present and provided the opportunity to complete all of the steps demonstrated in the orientation videos for an ABR. Work on the simulator continued as long as the residents desired and until they believed they were proficient with the steps and sequences of the capsulolabral repair. Further study and guided practice of the steps for an ABR on a cadaveric specimen then followed and continued as long as participants desired and until they believed they were proficient.

Group C: PBP. All residents randomized to group C attended the course (n = 16) and were exposed to a protocol identical to that of group B with the additional requirement of showing proficiency at various stages of the training (Fig 2). Each of the individual proficiency benchmarks for the procedural components was established based on the mean performance of separate groups of experienced surgeons on the specific exercise.[6-10,22,23] After arriving at the OLC, if the resident had not yet taken and passed the cognitive test online, covering the validated metric steps and errors demonstrated in the orientation videos, he or she was required to do so on site. A minimum score of 84% was required to pass. Those who initially failed continued to study the material and were provided additional faculty instruction.

Knot-tying study progressed in a manner similar to groups A and B. Once residents believed they had mastered the knot-tying skills, they had the opportunity for the integrity of their knots to be tested using the Fundamentals of Arthroscopic Surgery Training workstation knot tensiometer[30] if desired. To pass, the loop-knot construct had to elongate less than 3 mm when subjected to a static load of 15 lb for 15 seconds. The benchmark was set at a minimum of 3 of 5 knots meeting this standard. Once the formal testing process began, the subject tied 5 knot trials, which were labeled in sequence. All 5 were then tested using the tensiometer. Subjects who failed continued to practice until confidence was gained, and then the testing sequence was repeated with a new series of 5 knots being tied. This process continued until the resident achieved proficiency or was unable to do so and failed to show progressive improvement in the knot-tying skill set.[30]

Work and practice then began with the same shoulder model simulator used by group B. The model was oriented according to physician preference. Landmarks were identified, and posterior, midanterior, and anterosuperior portals were established. After a diagnostic examination was performed, the steps for a 3-anchor Bankart repair were practiced. The faculty instructors provided proximate feedback and recommended corrections based on the previously defined step and error metrics demonstrated in the orientation videos. Practice and faculty feedback continued through a complete procedure until the subject and his or her faculty instructor both believed the subject had adequately prepared for testing.

A new simulator model was then oriented in the resident's position of preference. Equipment representatives from multiple different vendors served as surgical assistants and were randomly assigned to participating surgeons. The assistants were instructed to act only at the specific direction of the operating surgeon. Prompting and coaching (of technique) were prohibited (the procedures were proctored by staff from the OLC). The resident surgeon then proceeded to complete a diagnostic evaluation and perform a 3-anchor ABR attempting to mimic the key steps identified in the orientation videos. By use of the Bankart metric score sheet, 1 of 6 designated faculty members, intimately familiar with the Bankart metrics, scored the subject in real time during the arthroscopic repair on the simulator model. The simulator benchmark for a passing score on the shoulder model was established from a prior study[22] and included a 3-anchor repair with no more than 4 total errors and no more than 1 sentinel error. If a resident failed to meet the benchmark, the faculty who scored the model test, the assigned training instructor, and the resident all conferred to identify the specific deficiencies exhibited and the appropriate corrections. The subject then worked toward acquisition of the requisite skills with instructor guidance. When confident, the subject was given 1 additional opportunity to repeat the scored procedure on a new model. In a normal PBP protocol, residents who fail to meet each of the intermediate proficiency benchmarks would not be allowed to progress in training and would require additional practice until the deficiencies were corrected (and would not be allowed to progress to working with the cadaver). However, given the artificial finite time constraints of the study weekend, all group C participants, regardless of persistent deficiencies, were allowed to proceed to practice with a cadaveric specimen and guided instruction similar to groups A and B.

Final Videotaped Bankart Repair Assessment

At the completion of their respective courses, the subjects from each group performed an assisted, unaided arthroscopic diagnostic survey and a 3-anchor Bankart repair on a fresh cadaveric shoulder. The cadaveric specimens were considered acceptable if (1) arthroscopic visibility of the target tissues was obtainable; (2) the specimen permitted adequate access to the target tissues (flexibility); and (3) the integrity of the capsulolabral tissues was sufficient to permit mobilization, suture delivery, and knot tying. All necessary instrumentation and implants were made available. Residents participating in the course served as assistants for each other. They were instructed to act only at the request of the operating surgeon and were prevented from coaching or prompting. The OLC staff proctored the procedures for compliance.

The procedure was videotaped in its entirety beginning with the initial view from the posterior portal. The resident surgeon mapped the bony landmarks and then created his or her preferred portals. All or a portion of the diagnostic examination was completed. The arthroscope was withdrawn, and a red card was videotaped for 5 seconds to signal that the subject surgeon was no longer operating. One of 4 designated faculty members then reintroduced the arthroscope and, using a sharp elevator from either the anterosuperior portal, midanterior portal, or both, created a standard Bankart lesion, 6 to 9 mm deep (medial from the bony rim) and from the 2- to 6-o'clock position along the glenoid. Once the Bankart lesion was created, care was taken to avoid additional mobilization of the capsulolabral tissue. The arthroscope was then withdrawn, and a green card was videotaped for 5 seconds, signaling that the subject surgeon was operating for the balance of the procedure. The arthroscope was reintroduced into the glenohumeral joint by the subject surgeon, and any remaining elements of the diagnostic survey were completed. The subject then performed a 3-anchor Bankart repair, attempting to mimic the steps demonstrated in the orientation videos and practiced in the simulation model.

All subject surgeons used identical implants: single-loaded (2.8-mm) Gryphon push-in anchors with a single No. 2 Orthocord (DePuy Mitek, Raynham, MA). A 45° cannulated suture hook was used to deliver a shuttling device with retrograde passage of the anchor sutures through the capsulolabral tissues. Before work was begun on the final scored Bankart repair, instructions were given to all residents regarding the protocol for anchor pullout from cadaveric bone. If an anchor failed prior to completion of the index sliding knot, the surgeon was permitted to remove the anchor and suture and to replace it with a metal 5.5-mm screw-in anchor (Smith & Nephew, Andover, MA). The procedure then continued with no penalty. The time required for the reintroduction of the screw-in anchor and re-passage of the anchor suture through the capsulolabral tissue was subtracted from the total procedure time. If the anchor failed subsequent to the initial sliding knot being completed (i.e., efforts to back up the primary knot with half-hitches), the surgeon was instructed to abandon the first anchor position and proceed to the second anchor position. No time limit was imposed on the performance of the Bankart repair, and participants were able to continue to work as long as they believed they were making progress. At the point in the procedure when the subject surgeons did not believe they could make further progress in the Bankart repair, they could electively choose to terminate the procedure.

Video Reviewer Training

Once the construction of the metrics for an ABR was completed and face and content validity verified,[21] a final version of a score sheet was formatted. Ten Master/Associate Master AANA faculty surgeons formed the panel of reviewers designated to score the videos. This group included the 3 arthroscopic surgeons (R.L.A., R.K.N.R., and R.A.P.) who, in conjunction with a consultant experimental psychologist (A.G.G.), developed the arthroscopic Bankart metric definitions (Table 1). The 10 reviewers were randomly assigned to form 5 fixed pairs, which remained constant throughout the scoring of all videos. Reviewer training was initiated with an 8-hour in-person meeting, during which time each metric was studied in detail. Multiple video examples of live patient cases were shown to illustrate each particular metric. Videos of patients in both the lateral decubitus and beach-chair orientations were represented. Discussion helped to clarify how each step and error were to be scored, including the nuances and conventions to be used. Several weeks later, full-length practice videos 1 and 2 (one each in the lateral decubitus and beach-chair orientation) were sent to and independently scored by each of the 10 reviewers, and the scores were tabulated. During 2 subsequent 2-hour group phone conferences, the differences and discrepancies among all reviewers were compared and discussed, seeking conformity in scoring. In addition, each designated pair of reviewers conducted 1 to 3 additional phone conferences to analyze the specific instances in which the two of them scored particular events differently. Subsequently, all reviewers scored practice videos 3 and 4 (with each patient orientation again represented), and the results were tabulated. The scores for each of the 5 designated pairs of reviewers were compared for the second set of practice videos. In only 1 of 10 comparisons (2 videos × 5 reviewer pairs) did the inter-rater reliability (IRR) (Table 1) calculation (as discussed later) fall below an acceptable level[8] of 0.8, at 0.76. Thus, confidence was established that the future scoring of the videos generated by the participants of this study could be accomplished with a high IRR.

Video Scoring

The AANA research coordinator randomly assigned the 44 full-length study videos, each with only the designated unique identifying number attached, to a single pair of reviewers. Other than the research coordinator and the study consultant, all video reviewers remained blinded to the source of the video being reviewed. Each video was independently reviewed and scored by the 2 members of an assigned pair of reviewers. All scores were tabulated for each of the 13 phases of the procedure. Each step and error metric were scored as either yes or no, designating whether

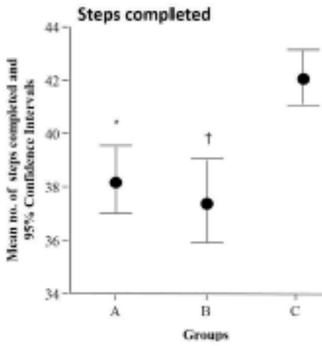


Fig 5. Mean number of Bankart procedure steps (and 95% confidence intervals) completed for each study group. The probability values observed for differences in performance were as follows: *$P < .001$ between group C and group A; †$P < .001$ between group C and group B.

the specific event was or was not observed to have occurred by the reviewer. In addition to scoring of steps and errors, each event characterized as "damage to non-target tissue" (Table 1) (e.g., gouging the articular cartilage or tearing of the capsule) was scored. There was no limit to the number of individual instances in which damage to non-target tissue could be scored, with each occurrence tallied as a single error event. The score sheet also contained a box for specific reviewer comments for each metric.

Score Tabulation

For each of the 13 separate phases of the procedure, the numbers of uncompleted steps and errors made were tabulated and the scores for the 2 reviewers averaged. Furthermore, for each subject, the step and error data were pooled for the 3 repetitive components of the procedure: (1) anchor preparation, (2) suture passage/management, and (3) knot tying. These data were used to determine which of the procedural phases showed the greatest differences in performance among the groups (1-factor analysis of variance analysis) (IBM SPSS statistical software program; IBM, Armonk, NY). Furthermore, for the entire procedure, the total numbers of steps completed, errors made, and sentinel errors enacted were also averaged for the pair of reviewers. The subject's operative time was obtained by subtracting the faculty time to create the Bankart lesion from the total recording time for the procedure.

The 2 raw score sheets from the designated pair of reviewers were compared for each of the individual steps ($n = 45$), and the number of agreements (either both reviewers documented that a step was performed or both scored the step as not being completed) was tabulated. In addition, the number of disagreements in scoring steps (one of the reviewers indicated that the step had been completed and the other indicated that the step had not) was tabulated. The IRR for the steps was calculated according to the following formula: Agreements/(Agreements + Disagreements).

In a similar manner, there was either agreement or disagreement in the 2 scores for each of the potential errors ($n = 77$). The IRR for error scoring was calculated in the same manner as that for the steps. Finally, the IRR for scoring the entire procedure was calculated using both the step and error agreements or disagreements for the complete procedure ($n = 122$). The acceptable IRR was defined as 0.80 or greater.[8]

Statistical Methods

The analysis was conducted as a series of multiple regressions. The exogenous variables (covariates) were the 3 intervention conditions, that is, PBP plus simulator (condition C), simulator (condition B), and traditional training (condition A). Group C was used as the reference condition within the analysis. As a check on the veracity and stability of the results, all of the analyses were also conducted using Poisson regression. The substantive interpretation remained unchanged regardless of the model used. All of the reported results are based on the analyses from the multiple regressions. Furthermore, a logistic regression analysis was performed to estimate the probability of those trainees from the different training curricula being able to attain the proficiency benchmark for the final ABR.

A secondary analysis was conducted to evaluate the subset of group C subjects who successfully met all of the intermediate benchmarks throughout training,

Table 5. Regression Analysis of Steps Completed

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	SE	β	*t*	*P* Value
Reference group: group C (constant)	42.167	0.618		68.225	< .001
Group A	−3.833	0.927	−0.599	−4.135	< .001
Group B	−4.702	0.890	−0.715	−5.286	< .001

B, beta; β, standardized beta; SE, standard error; t, test statistic.

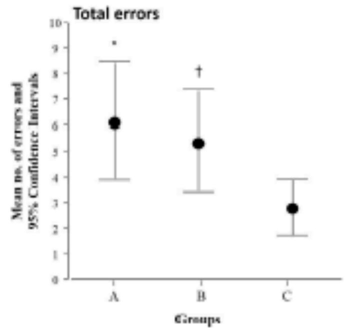


Fig 6. Mean number of procedure errors enacted for each study group. The probability values observed for differences in performance were as follows: *$P = .01$ between group C and group A; †$P = .049$ between group C and group B.

designated as group C^{PBP}. Group C^{PBP} was evaluated for the same performance metrics as the other groups, which included steps, errors, sentinel errors, and time, as well as the probability of attaining the benchmark on the final repair. All of the participants in group C followed the PBP training "curriculum." The PBP "protocol," in distinction, would only permit those individuals who meet each intermediate proficiency benchmark to progress in training (group C^{PBP}) (Fig 2).

Results

The mean and standard deviation scores on the baseline assessments of perceptual, visuospatial, and psychomotor performance are shown in Table 4. Although group A performed somewhat better on the psychomotor test than groups B and C, these differences were not statistically significant.

Intermediate Proficiency Training Benchmarks for Group C

All 16 participants in group C were able to obtain a passing score on the cognitive examination, although several required additional instruction after failing to achieve a passing score of 84% on their initial test. One subject from this group was unable to show proficiency in knot tying despite repeated training and practice. Six group C participants failed their first attempt to meet the benchmark for a Bankart repair on the simulator model. After additional guided training and practice, 2 of 6 were able to show proficiency on their second attempt with the shoulder model. Of the 4 participants who were unable to show proficiency on the model, one of whom was also the participant who failed to show proficiency at knot tying. Thus 12 of 16 group C subjects met all of the intermediate proficiency benchmarks during training. On the basis of the PBP protocol, these 12 (designated group C^{PBP}) would have been the only participants from group C allowed to progress to working on the cadaver.

Final Cadaveric Bankart Assessment

Two cadavers, 1 each from groups B and C, failed to meet the acceptability criteria and were replaced with better specimens. The video recording was restarted with the onset of work on the replacement specimen. For the 44 videos scored, the mean IRR for the total number of steps performed and errors made was 0.93 (range, 0.84 to 0.99; SD, 0.04).

Incomplete Final Procedures

Of all 44 subjects, only 3 failed to complete their final Bankart repair on a cadaveric shoulder. Two individuals from group A were only able to finish the first anchor with an average of 16.25 steps completed, 7 errors made, and 0.5 sentinel errors enacted. They worked for an average of 99 minutes. In group C, 1 subject had the first anchor pullout during efforts to deliver and secure the primary suture knot, and this subject elected not to replace that anchor (although the subject could have done so according to the anchor pullout protocol). The subject completed all of the second and third anchor components, thus performing only a 2-anchor final repair, which was deemed incomplete. During this procedure, an average of 37 steps were completed, along with 4 errors made and 0 sentinel errors enacted. The operative time was 92 minutes. This subject was the one who had previously failed to show proficiency both on knot tying and on the shoulder model repair components of the training curriculum and would not

Table 6. Regression Analysis of Errors Made

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	SE	β	*t*	*P* Value
Reference group: group C (constant)	2.600	0.819		3.173	.003
Group A	3.275	1.229	0.444	2.665	.011
Group B	2.400	1.179	0.339	2.035	.049

B, beta; β, standardized beta; SE, standard error; t, test statistic.

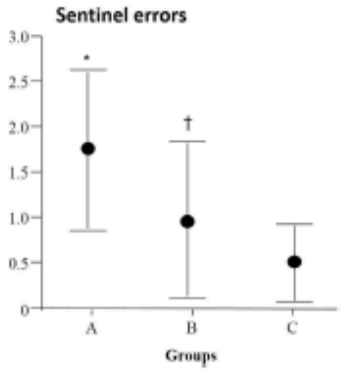


Fig 7. Mean number of sentinel errors enacted for each study group. The probability values observed for differences in performance were as follows: *$P = .023$ between group C and group A; †$P = .351$ between group C and group B.

normally have been allowed to progress to training with the cadaver. It was not possible to estimate or accurately extrapolate the number of errors, number of sentinel errors, or time for the 3 incomplete procedures. Thus, for the comparative analysis of the 3 groups for steps, errors, sentinel errors, and time, group A comprised 12 subjects, group B comprised 14, and group C comprised 15.

Steps Completed

Figure 5 shows the mean values and 95% confidence intervals (CIs) of procedure steps completed by groups A, B, and C. Groups A and B completed a similar number of procedure steps, whereas group C, on average, completed 4 more steps. The differences between the groups' performances using the regression model with group C as the reference group are summarized in Table 5. The results showed that group C completed, on average, 42.2 procedure steps. Subjects in group A completed 3.8 fewer steps, whereas those in group B completed 4.7 fewer steps. Both of these differences were statistically significant ($P < .001$ for group C *v* group A and $P < .001$ for group C *v* group B).

Procedure Errors

The mean number of errors and 95% CIs for each group are shown in Figure 6. The mean number of errors among subjects in group C was 2.6 (Table 6). On average, the subjects in group A made 3.3 more errors whereas those in group B made 2.4 more errors than those in group C. Both of these differences were statistically significant ($P = .011$ for group C *v* group A and $P = .049$ for group C *v* group B). Overall, the subjects in group C showed a 56% reduction in the mean number of errors over group A and a 41% reduction over the number of errors made by those in group B. The participants in group C were also more consistent, with the range in number of errors much smaller when compared with groups A and B.

Sentinel Errors

The mean values and 95% CIs for sentinel errors are shown in Figure 7. On average, the subjects in group C made 0.53 sentinel errors, whereas those in group A made 1.175 more sentinel errors and those in group B made 0.43 more sentinel errors (Table 7). The difference between group A and group C for sentinel errors was statistically significant ($P = .017$), but the difference between group C and group B was not. Overall, the subjects in group C made 69% fewer sentinel errors than those in group A and 44% fewer than group B.

Bankart Performance Time

The mean values and 95% CIs for time taken by the groups to perform the index procedure are shown in Figure 8. Groups A, B, and C took a similar amount of time to complete the procedure, with no significant differences observed (Table 8).

Analysis of Group C^{PBP}

A secondary analysis was conducted to determine the performance of the 12 group C^{PBP} subjects. The differences between the mean scores of group C^{PBP} and those of group C were calculated. The difference in steps completed was marginal (42.46 for group C^{PBP} *v* 42.2 for group C). The error analysis showed that 12% fewer errors were made (2.29 for group C^{PBP} *v* 2.6 for group C), 6% fewer sentinel errors were made (0.5 for group C^{PBP} *v* 0.53 for group C), and 4% less time was required to complete the procedure (77.17 minutes for group C^{PBP} *v* 80.44 minutes for group C).

Table 7. Regression Analysis of Sentinel Errors Made

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	SE	β	*t*	*P* Value
Reference group: group C (constant)	0.533	0.313		1.706	.096
Group A	1.175	0.469	0.425	2.506	.017
Group B	0.431	0.450	0.162	0.958	.344

B, beta; β, standardized beta; SE, standard error; t, test statistic.

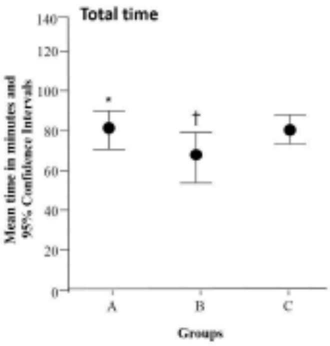


Fig 8. Mean procedure duration in minutes for each study group. There were no significant differences in the probability values between groups C and A (asterisk) or between groups C and B (dagger).

Final Bankart Proficiency Benchmark

The proficiency benchmark (set as the mean performance of an experienced group of surgeons) had previously been established for a cadaveric shoulder[23] as no more than 3 total errors (1 less error than the simulator model benchmark) and no more than 1 sentinel error. In addition, a 3-anchor Bankart repair must have been completed. Overall, 28.6% of group A subjects (4 of 14), 36.7% of group B (5 of 14), 68.7% of group C (11 of 16), and 75% of group C^{PBP} (9 of 12) were able to achieve the final proficiency benchmark.

Logistic regression analysis for the relative differences between the control condition (group A, traditionally trained group) and group B (simulator), group C (simulator + BPB curriculum), and group C^{PBP} (simulator + PBP protocol) was performed and used to determine the odds ratios for the comparisons. Relative to group A, group B subjects were 1.4 times ($P = .121$), group C subjects were 5.5 times ($P = .033$), and group C^{PBP} subjects were 7.5 times ($P = .024$) as likely to achieve the final quantitatively defined proficiency benchmark. Only the comparisons of proficiency between group A and group C, as well as between group A and group C^{PBP}, were statistically significant (Fig 9). Trainees in group C^{PBP} had a 36.4% greater probability of achieving the final benchmark than those in the entire group C.

Discussion

PBP Paradigm

Two primary conclusions can be drawn from the data in this study. First, the performance of the entire group C on the final Bankart evaluation shows that the PBP curriculum using simulation is superior both to the traditional curriculum (group A) and to the curriculum identical to that in group C (including the use of the simulator) but without the requirement to show proficiency (group B). Second, the performance of group C^{PBP} shows the superiority of the PBP protocol itself, in which only those trainees who meet each sequential intermediate proficiency benchmark during training are permitted to progress in the curriculum. The most important and revealing comparison of the 3 training protocols, therefore, compares groups A, B, and C^{PBP}. The subjects in group C^{PBP} performed more of the operative steps but did not take significantly longer to do so. They also made significantly fewer objectively assessed intraoperative errors and were over 7 times more likely to achieve the final benchmark than those in group A, who followed a traditional training pathway.

The performance of group B was only marginally better than that of group A, and this finding suggests that it is not simply access to working with the simulator that is important. Rather, it is the metric-dependent PBP curriculum coupled with the simulator that optimizes the effectiveness of the training. The findings of this investigation strongly support the "outcomes" (objective assessments)–based approach rather than the "process" (time spent/exposure gained)–based approach to graduate medical education advocated by the Institute of Medicine.[31]

Because of the artificial constraint of a limited period for training (a weekend course), all members of group C were allowed to progress and practice with the cadaver and participate in the final Bankart assessment whether they met all of the intermediate benchmarks or not. During the training process, 25% of the subjects in group C failed to show proficiency during the knot-tying phase or on the shoulder simulation model (1 subject failed both). These individual performances

Table 8. Regression Analysis of Time Taken

Model	Unstandardized Coefficients		Standardized Coefficients		
	B	SE	β	t	P Value
Reference group: group C (constant)	80.438	4.455		18.055	< .001
Group A	0.634	6.522	0.016	0.097	.923
Group B	−13.009	6.522	−0.331	−1.995	.053

B, beta; β, standardized beta; SE, standard error; t, test statistic.

are, thus, not representative of the PBP protocol and diminish the performance of group C as a whole. Had these 4 subjects had additional time and further opportunities to show proficiency, it is likely that they may have been able to do so.

Overall, confidence in the observed effects of the training methods was high. There were no statistically significant differences between the groups on pre-course, baseline visuospatial, perceptual, or psychomotor assessments. Furthermore, the blinded and objectively assessed videotaped performances of the subjects' final Bankart repairs were scored with a consistently high IRR (>90% agreement between the raters for all assessments, with none falling below 80%).

Performance Outliers

For the 2 group A subjects who effectively completed only a 1-anchor Bankart repair, it is probable that they would have enacted substantially more errors during a complete 3-anchor procedure. It was not possible to accurately estimate the total number of errors that would have been enacted for a full repair. Consequently, all of their data (including steps completed and errors enacted) had to be excluded from the statistical analysis. The necessary exclusion of these 2 subjects' data results in an overestimation of the performance of group A as a whole. The 1 group C subject whose repair was considered incomplete abandoned the effort on the first anchor repair because this subject was unable to complete the knot-tying steps. This subject did, however, complete the second and third anchor components of the repair and, along with performing 37 steps, made only 4 errors and no sentinel errors. Thus the subject's overall performance was not substantially different from group C as a whole.

PBP Superiority

An important finding from this study is that the training process must be more than an educational experience. Simple knowledge of the metrics (steps and errors) and the opportunity to practice with expert feedback (group A, traditional AANA Resident Course) resulted in an inferior demonstration of arthroscopic Bankart skills. Furthermore, the addition of the opportunity to work with the simulator (group B) resulted in a modest improvement in performance over the control (group A). The use of the metric-dependent PBP curriculum coupled with the simulator (group C) resulted in the acquisition of a statistically superior ABR skill set. Multiple potential reasons exist for the superiority of this protocol:

1. The requirement to obtain a passing score on the cognitive examination at the outset ensures that the trainee is very familiar with the steps to be completed and errors to be avoided for the reference repair.
2. Proximate feedback linked to established and validated metrics facilitates the prompt, specific, and effective correction of errors.
3. Deliberate practice[12] in attempting to mimic the specific skills demonstrated in the orientation videos ensures uniformity in acquiring the essential skills needed to perform the reference procedure.
4. The medium-fidelity simulator provides the opportunity for practice and repetition of the important skills necessary for effective performance of a Bankart repair.
5. The validated performance benchmark for the simulator serves as an intermediate assessment tool and helps identify individual trainee deficiencies requiring correction.
6. The trainee's knowledge of the requirement to show proficiency by meeting each of the intermediate benchmarks to be able to progress in training helps the trainee focus on acquiring the necessary skills.
7. Trainee performance at a quality-assured performance level based on validated metrics must be shown.[7,8,33,34]

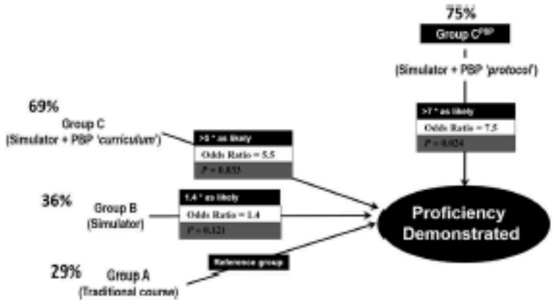


Fig 9. Odds ratios and statistical significance of differences between group A and groups B, C, and C^{PBP} for final Bankart proficiency demonstration. (PBP, proficiency-based progression.)

Metric-Based Curriculum

There are a number of unique aspects to this study, and they primarily relate to the development and use of the procedural metrics. This is the first simulation-based study, to our knowledge, in which the metric characterization and validation of a complete procedure have been carried out. This effort sought to investigate the merits of the emerging paradigm shift in surgical skills training from the apprenticeship model to a PBP format and, consequently, became known as the AANA Copernicus Initiative (Nicolaus Copernicus is credited with the paradigm shift from the earth to the sun being considered the center of the universe). In the approach reported on in this study, the simulation was simply one of the vehicles for the delivery of a metric-based training curriculum. Much of the effort focused on the development and validation of metric-based performance characteristics that appropriately captured a reference approach to the performance of an ABR.[21] Face validity and content validity for the Bankart metrics were verified using a modified Delphi panel meeting with Master and Associate Master AANA shoulder faculty.[21] The construct validity of the metrics coupled with the shoulder simulation model[22] and separately with a cadaveric shoulder[23] was confirmed. On the basis of these results, specific performance benchmarks were established separately for the shoulder model and for the cadaveric specimens.

The approach to the assessment of performance in this study uses precise metric definitions of performance and simply requires the scorer to determine whether a specific event did or did not occur. This binary approach to the measurement of performance has been shown to facilitate the reliable scoring of metric-based performance units across a variety of functions during skills training[10,24,35,36] with different experience levels.[37,38] In contrast, Likert-scale assessments (Table 1) result in a less focused approach to minimizing errors because the deviations from optimal performance are less clearly defined.[39] A Likert-type scale is a method of ascribing a quantitative value to qualitative data to make it amenable to statistical analysis and was originally designed to assess a range of respondent attitudes.[40] Because of the inherent subjectivity in this method of attempting to rate objective performance, it can be difficult to obtain acceptable levels of IRR (>80%) in the scoring of events.[18] It has been shown that Likert-scale scoring may be less reliable than metric-based assessments[41] and simply gives the trainee feedback information on the global aspects of his or her performance.

Simulation Platforms

The detailed metrics enabled a simulation platform that already existed (an anatomically accurate shoulder model) to be used for training and assessment. This medium-fidelity platform is relatively inexpensive, is readily available, and serves as an accurate representation of the human shoulder joint. One shortcoming of a physical simulator, however, is that it is unable to capture any performance data or provide feedback to the trainee. A significant investment in time and effort for instructional faculty and video reviewers was required to obtain detailed data and formulate accurate performance assessments. The approach used for simulation-based training in this study, nevertheless, holds considerable promise in the short-term because the vast majority of surgical procedures (particularly for traditional open surgery) have no virtual reality platform. Relative to higher-fidelity computer-based simulators, physical simulation models are much easier to develop in an expedient manner. This capability affords the surgical community the opportunity to develop PBP simulation-based training programs in a reasonable amount of time for traditional and new surgical procedures. The crucial element in terms of the effectiveness of any simulator will be its coupling with appropriate and accurate metric-based characterizations and the "operational definitions" (Table 1) for those metrics.

Studies assessing the value of simulation in surgical skills education and training have begun to emerge. We were unable to find any studies in the literature with which to compare our investigation. Frank et al.[42] performed a systematic review of the published literature (19 studies) on modern arthroscopic simulator training models. The analysis suggested that practice on arthroscopic simulators improves performance on the simulators, but evidence that skills obtained during simulator training are transferred to the operating room is lacking. Cannon et al.[43] studied the impact on transfer of training using the ArthroSim (ToltTech, Aurora, CO) virtual reality arthroscopic knee simulator, which has previously been shown to have construct validity.[44] The PGY 3 orthopaedic residents who trained on the simulator (for an average of 11 hours) showed greater proficiency on a live diagnostic knee arthroscopy than the control group trained in a traditional fashion. They performed significantly better on the procedural checklist and assessment of probing skills but not on the assessment of visualization skills.

Implications

For well over a century, the apprenticeship model has been the predominant method used to assist surgical trainees in skill acquisition and preparation for the practice of surgery. A paradigm of repeated observation in addition to graded, enhanced responsibility and independence during operations of increasing technical complexity has been used. Although reasonably effective, this approach is inefficient and produces considerable variability in the skill sets obtained by trainees with equivalent time exposure and experience.[5] Alternatively, PBP training using simulation enables trainees to focus on the acquisition of specific procedural skills, measure their progress, and correct deficiencies. It is through the process of deliberate practice[6] that they learn not only what to do but, perhaps more importantly, what not to do. The trainee is thus able to enact errors and learn to correct them in an inconsequential manner and without risk to patients. In addition, this structured approach promotes the acquisition of a more homogeneous skill set at the completion of training.[33]

One of the concerns of the medical community prior to this study was the generalizability of simulation-based training. Even simulation enthusiasts harbor the concern that simulation-based training effectiveness may be, in part, a function of the effort that enthusiasts put into the training initiatives and the reported science. The results of this study showed that simulation-based training is very effective, even when applied across a large number of residents from training programs throughout the United States using faculty equally dispersed. One of the reasons for this success was our deliberate choice of a reference approach to a particular procedure. This method of standardization means that, at the outset of their learning a particular procedure, trainees do not have a myriad of approaches and techniques to master. They can develop their own surgical style once they have acquired safe operative skills for the reference approach.

Limitations

Other than their year in training, no additional information regarding the participants was available. The extent of an individual resident's arthroscopic surgery exposure and experience, particularly shoulder arthroscopy, was likely variable. The diverse group of residency programs represented (21) and the fact that residents from an individual program were randomized to different training protocols should have minimized any selection bias. Although the residents in group A performed the best overall on the baseline visuospatial, perceptual, and psychomotor assessments, the between-group differences were not statistically significant.

The number of participants enrolled was based, in part, on previous studies (e.g., Seymour et al.[6]) and secondarily on the logistic challenges of having more than 12 simultaneous recording stations at the OLC during the final Bankart repair. Because there were more than 12 participants in each of the groups, the residents were randomly assigned to flights to complete their videotaped repair. The 14 group A residents were volunteers from a normal AANA Resident Course, and their Bankart recording was performed over 2 flights. Training for the 14 group B and 16 group C participants was conducted on a separate study weekend. The final Bankart repair for these 2 groups required 3 flights of surgical procedures, some of which lasted 2 hours. With turnover, 8 hours were required to record the 3 flights (30 total recordings).

It is acknowledged that conducting an analysis of PBP training during the finite time period of a weekend course imposes an artificial constraint. PBP training dictates that the trainee continue to study and practice as long as it takes to master the requisite skills. It is probable that most of the participants in group C who did not reach all of the intermediate benchmarks during the weekend course would likely have been able to do so with additional training and practice. Furthermore, it was not the intent of this investigation to study the efficiency of individual resident's skill acquisition or residents' efficiency in performing the index procedure. The efficiency of the PBP training protocol, however, is clearly implied. In essentially the same time frame as the other groups, a substantially greater percentage of residents trained using the PBP protocol coupled with simulation achieved the final benchmark compared with those participating in the other training methods.

The final proficiency assessment was performed using cadaveric shoulders. The specimen size, tissue compliance, and extent of pre-existing glenohumeral pathology likely introduced some inherent variability in the cadaveric shoulders. The impact of these differences was minimized by using the previously described acceptability criteria: Those specimens deemed unsuitable for an ABR were replaced, and a new video recording was initiated.

Finally, all participants and faculty for each of the 3 study groups were provided a link to 2 full-length videos (lateral decubitus and beach chair) demonstrating each of the 45 steps and showing or identifying all of the errors and sentinel errors. Although the link was provided to all participants and faculty 1 month before the course in which they were involved, we have no record of how often they viewed or studied the videos.

Conclusions

A PBP training curriculum and protocol coupled with the use of a shoulder model simulator and previously validated metrics produces a superior arthroscopic Bankart skill set when compared with traditional and simulator-enhanced training methods.

References

1. Barden CB, Specht MC, McCarter MD, et al. Effects of limited work hours on surgical training. *J Am Coll Surg* 2002;195:531-538.
2. Pickersgill T. The European working time directive for doctors in training. *BMJ* 2001;323:1266.

3. Satava RM. Virtual reality surgical simulator. The first steps. *Surg Endosc* 1993;7:203-205.
4. Sachdeva AK, Pellegrini CA, Johnson KA. Support for simulation-based surgical education through American College of Surgeons—accredited education institutes. *World J Surg* 2008;32:196-207.
5. Pellegrini CA, Sachdeva AK, Johnson KA. Accreditation of education institutes by the American College of Surgeons: A new program following an old tradition. *Bull Am Coll Surg* 2006;91:8-12.
6. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: Results of a randomized, double-blinded study. *Ann Surg* 2002;236:458-463.
7. Gallagher AG, Ritter EM, Champion H, et al. Virtual reality simulation for the operating room: Proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg* 2005;241:364-372.
8. Gallagher AG, O'Sullivan GC. *Fundamentals of surgical simulation—Principles and practices*. London: Springer Verlag, 2011.
9. Ahlberg G, Enochsson L, Gallagher AG, et al. Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. *Am J Surg* 2007;193:797-804.
10. Van Sickle K, Ritter EM, Baghai M, et al. Prospective, randomized, double-blind trial of curriculum-based training for intracorporeal suturing and knot tying. *J Am Coll Surg* 2008;207:560-568.
11. Waterman BR, Burns TC, McCriskin B, et al. Outcomes after Bankart repair in a military population: Predictors for surgical revision and long-term disability. *Arthroscopy* 2014;30:172-177.
12. Petrera M, Dwyer T, Tsuji MR, Theodoropoulos JS. Outcomes of arthroscopic Bankart repair in collision versus noncollision athletes. *Orthopedics* 2013;36:e621-e626.
13. Harris JD, Bupta AK, Mall NA, et al. Long-term outcomes after Bankart shoulder stabilization. *Arthroscopy* 2013;29: 920-933.
14. Archetti Netto N, Tamaoki MJ, Lenza M, et al. Treatment of Bankart lesions in traumatic anterior instability of the shoulder: A randomized controlled trial comparing arthroscopy and open techniques. *Arthroscopy* 2012;28: 900-908.
15. Chahal J, Marks PH, Macdonald PB, et al. Anatomic Bankart repair compared with nonoperative treatment and/or arthroscopic lavage for first-time traumatic shoulder dislocation. *Arthroscopy* 2012;28:565-575.
16. Stein T, Linke RD, Buckup J, et al. Shoulder sport-specific impairments after arthroscopic Bankart repair: A prospective longitudinal assessment. *Am J Sports Med* 2011;39: 2404-2414.
17. Ee GW, Mohamed S, Tan AH. Long term results of arthroscopic Bankart repair for traumatic anterior shoulder instability. *J Orthop Surg Res* 2011;6:28.
18. Angelo R. Arthroscopic Bankart repair for unidirectional shoulder instability. *Instr Course Lect* 2009;58:305-313.
19. Ryu RK. Arthroscopic approach to traumatic anterior shoulder instability. *Arthroscopy* 2003;19:94-101.
20. Smith R. All changed, changed utterly. *BMJ* 1998;316: 1917-1918.
21. Angelo RL, Ryu RKN, Pedowitz RA, Gallagher AG. Metric development for an arthroscopic Bankart procedure: Assessment of face and content validity. *Arthroscopy* 2015;31:1430-1440.
22. Angelo RL, Pedowitz RA, Ryu RKN, Gallagher AG. The Bankart performance metrics combined with a shoulder model simulator create a precise and accurate training tool for measuring surgeon skill. *Arthroscopy* 2015;31:1639-1654.
23. Angelo RL, Pedowitz RA, Ryu RKN, Gallagher AG. The Bankart performance metrics combined with a cadaveric shoulder create a precise and accurate assessment tool for measuring surgeon skill. *Arthroscopy* 2015;31:1655-1670.
24. McClusky DA III, Ritter EM, Lederman AB, et al. Correlation between perceptual, visuo-spatial, and psychomotor aptitude to duration of training required to reach performance goals on the MIST-VR surgical simulator. *Am Surg* 2005;71:13-20.
25. Gallagher AG, Cowie R, Crothers I, et al. PicSOr: An objective test of perceptual skill that predicts laparoscopic technical skill in three initial studies of laparoscopic performance. *Surg Endosc* 2003;17:1468-1471.
26. Gallagher AG, McClure N, McGuigan J, et al. An ergonomic analysis of the fulcrum effect in the acquisition of endoscopic skills. *Endoscopy* 1998;30:617-620.
27. Crothers IR, Gallagher AG, McClure N, et al. Experienced laparoscopic surgeons are automated to the "fulcrum effect": An ergonomic demonstration. *Endoscopy* 1999;31: 365-369.
28. Gallagher AG, Smith CD, Bowers SP, et al. Psychomotor skills assessment in practicing surgeons experienced in performing advanced laparoscopic procedures. *J Am Coll Surg* 2003;197:479-488.
29. Randomization.com Web site. Available at http://www.randomization.com. Accessed November 23, 2013.
30. Pedowitz RA, Nicandri GT, Angelo RL, et al. Objective assessment of knot-tying proficiency with the fundamentals of arthroscopic surgery training program workstation and knot tester. *Arthroscopy* 2015;31:1872-1879
31. Asch DA, Weinstein DA. Innovation in medical education. *N Engl J Med* 2014;371:794-795.
32. Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 1993;100:363-406.
33. Zendejas B, Brydges R, Hamstra SJ, et al. State of the evidence on simulation-based training for laparoscopic surgery: A systematic review. *Ann Surg* 2013;257: 586-593.
34. Gallagher AG. Metric-based simulation training to proficiency in medical education:- What it is and how to do it. *Ulster Med J* 2012;81:107-113.
35. Van Sickle K, Smith B, McClusky DA III, et al. Evaluation of a tensiometer to provide objective feedback in knot-tying performance. *Am Surg* 2005;71:1018-1023.
36. Van Sickle K, Gallagher AG, Smith CD. The effect of escalating feedback on the acquisition of psychomotor skills for laparoscopy. *Surg Endosc* 2007;21:220-224.
37. Neary PC, Boyle E, Delaney CP, et al. Construct validation of a novel hybrid virtual-reality simulator for training and assessing laparoscopic colectomy; results from the first course for experienced senior laparoscopic surgeons. *Surg Endosc* 2008;22:2301-2309.

38. Nicholson WJ, Cates CU, Patel AD, et al. Face and content validation of virtual reality simulation for carotid angiography: Results from the first 100 physicians attending the Emory NeuroAnatomy Carotid Training (ENACT) program. *Simul Healthc* 2006;1:147-150.
39. Gallagher AG, Ritter EM, Satava RM. Fundamental principles of validation, and reliability: Rigorous science for the assessment of surgical education and training. *Surg Endosc* 2003;17:1525-1529.
40. Likert R. A technique for the measurement of attitudes. *Arch Psychol* 1932;140:44-53.
41. Gallagher AG, O'Sullivan GC, Leonard G, Bunting BP, McGlade KJ. Objective structured assessment of technical skills and checklist scales reliability compared for high stakes assessments. *ANZ J Surg* 2014;84:568-573.
42. Frank RM, Erickson B, Frank JM, et al. Utility of modern simulator training models. *Arthroscopy* 2014;30: 121-133.
43. Cannon WD, Garrett WE Jr, Hunter RE, et al. Improving residency training in arthroscopic knee surgery with use of a virtual-reality simulator. A randomized blinded study. *J Bone Joint Surg Am* 2014;96:1798-1806.
44. Cannon WD, Nicandri GT, Reinig K, et al. Evaluation of skill level between trainees and community orthopaedic surgeons using a virtual reality arthroscopic knee simulator. *J Bone Joint Surg Am* 2014;96:e57.