Received: 17 June 2016 | Revised: 5 October 2016 | Accepted: 15 November 2016
DOI 10.1002/rcs.1799

WILEY The International Journal of Medical Robotics and Computer Assisted Surgery

ORIGINAL ARTICLE

A hierarchical task analysis of shoulder arthroscopy for a virtual arthroscopic tear diagnosis and evaluation platform (VATDEP)

Doga Demirel[1,2] | Alexander Yu[1] | Seth Cooper-Baer[1] | Aditya Dendukuri[1] | Tansel Halic[1] | Sinan Kockara[1] | Nizamettin Kockara[3] | Shahryar Ahmadi[4]

[1] Computer Science Department, University of Central Arkansas, Conway, Arkansas, USA

[2] Computer Science Department, University of Arkansas, Little Rock, Arkansas, USA

[3] Department of Orthopedics and Traumatology, Erzincan University Medical School, Erzincan, Turkey

[4] Department of Orthopaedic Surgery, University of Arkansas for Medical Sciences, Little Rock, Arkansas, USA

Correspondence Tansel Halic, Computer Science Department, University of Central Arkansas, 201 Donaghey Ave, Conway, AR 72035, USA.
Email: tanselh@uca.edu

Funding information
National Institute of General Medical Sciences, (NIGMS), Grant/Award Number: P20 GM103429: National Institutes of Health (NIH), Grant/Award Number: NIH/NHLBI 1R01HL119248-01A1, NIH/NIBIB 2R01EB005807, 5R01EB010037, 1R01EB009362, and 1R01EB014305

Abstract

Background Shoulder arthroscopy is a minimally invasive surgical procedure for diagnosis and treatment of a shoulder pathology. The procedure is performed with a fiber optic camera, called arthroscope, and instruments inserted through very tiny incisions made around the shoulder. The confined shoulder space, unintuitive camera orientation and constrained instrument motions complicates the procedure. Therefore, surgical competence in arthroscopy entails extensive training especially for psychomotor skills development. Conventional arthroscopy training methods such as mannequins, cadavers or apprenticeship model have limited use attributed to their low-fidelity in realism, cost inefficiency or incurring high risk. However, virtual reality (VR) based surgical simulators offer a realistic, low cost, risk-free training and assessment platform where the trainees can repeatedly perform arthroscopy and receive quantitative feedback on their performances. Therefore, we are developing a VR based shoulder arthroscopy simulation specifically for the rotator cuff ailments that can quantify the surgery performance. Development of such a VR simulation requires a through task analysis that describes the steps and goals of the procedure, comprehensive metrics for quantitative and objective skills and surgical technique assessment.

Methods We analyzed shoulder arthroscopic rotator cuff surgeries and created a hierarchical task tree. We introduced a novel surgery metrics to reduce the subjectivity of the existing grading metrics and performed video analysis of 14 surgery recordings in the operating room (OR). We also analyzed our video analysis results with respect to the existing proposed metrics in the literature.

Results We used Pearson's correlation tests to find any correlations among the task times, scores and surgery specific information. We determined strong positive correlation between cleaning time vs difficulty in tying suture, cleaning time vs difficulty in passing suture, cleaning time vs scar tissue size, difficulty passing vs difficulty in tying suture, total time and difficulty of the surgery.

Conclusion We have established a hierarchical task analysis and analyzed our performance metrics. We will further use our metrics in our VR simulator for quantitative assessment.

KEYWORDS
arthroscopy, hierarchical task analysis and video analysis, rotator cuff, virtual surgical simulation

1 | BACKGROUND

Shoulder arthroscopy is a minimally invasive surgery for diagnosis and treatment of the tissues/joints in the shoulder area.[1] Surgical treatment of rotator cuff tears is one of the most common shoulder arthroscopy procedures. Rotator cuff is a group of muscles and tendons that connect the shoulder blade to the upper arm. These muscles enable rotational motion of the shoulder and provide stability. An injured rotator cuff can cause pain, movement constraints and weakness in the arm and shoulder. If the problems are persistent and non-surgical treatments are insufficient, surgical intervention is necessary to avoid severe pain and to regain shoulder mobility; e.g. cases with substantial rotator cuff muscle tears.[2] Over the years, arthroscopy has been adopted for the repair technique as opposed to open surgery.[3,4]

During arthroscopy, the whole procedure is performed with a fiber optic camera called an arthroscope and instruments inserted through very tiny incisions made at the shoulder. Surgeons perform the whole operation with 2D arthroscopic view on a screen. Arthroscopy, unlike open surgery, is very challenging because of the constrained instrument motions attributed to the limited and narrow space, unconventional hand–eye coordination, and limited field of view arising from the angle of the arthroscope.[5] Arthroscopy requires surgeons equipped with a high level of psychomotor skill to perceive a 3-dimensional environment from a 2-dimensional camera image and manipulate instruments at the same time.[6] In addition to basic hand–eye coordination, arthroscopic rotator cuff repair requires proficiency in a diverse set of specialized skills including but not limited to: suture-to-bone fixation, suture-to-tendon fixation, and restoration of the anatomic rotator cuff footprint (the surface area of bone to which the cuff tendons are attached). The skill competence in these tasks requires significant training, technical ability and experience.

Despite the need for rigorous training, there is no standard in teaching or in performance assessment. Although associations such as The Arthroscopy Association of North America (AANA) and American Board of Orthopedic Surgery (ABOS) require a minimum number of surgery cases for surgeon membership eligibility or certain specialty, they do not mandate any guidelines and regulations about objective measurement of the proficiency and training level of surgeons. The medical curriculum in residency also does not specify any certification or accreditation that is based on any objective examination.[7] Therefore, the minimum training and skills concerning the patient health and safety in arthroscopy still remains unclear.

Traditionally, surgical training uses the 1-to-1 apprenticeship model. Recent changes in surgical training to a more streamlined approach in conjunction with a restriction in working hours have resulted in difficulties in maintaining this apprenticeship model.[8] Learning arthroscopic skills for novice surgeons in the OR with this model is time-consuming and might also cause iatrogenic injuries to the patients.[9] Also, arthroscopy training using cadavers and animals are costly and limited with one-time use only.[9] In contrast, VR based surgical simulators could offer a low-cost, realistic risk-free training and assessment platform.[10] Existing studies noted that continuous training and receiving objective feedback in VR simulators can enhance surgical performance.[11–14] Quantitative performance assessment with VR simulation needs to be translated to actual OR experience. This can be only viable with derivation of the surgery metrics validated with actual OR performance. However, the metrics respecting the measurable or less subjective scores for training as well as OR performance are still lacking in arthroscopy especially for the shoulder. This inadequacy in the literature affects the efficacy of any VR based arthroscopy simulation, as there is no consensus and validation for the ground truth in quantifiable rubric and proficiency level.[15]

Our ultimate goal is to develop a VR based arthroscopic rotator cuff tear diagnosis and repair surgery simulator platform (VATDEP) integrated with a valid scoring rubric, which could objectively identify the skill level of a shoulder arthroscopy surgeon. As a first step of development of the VATDEP, we analyzed the procedure and derived tasks/subtasks (Actions) and goals associated with ideal and optional actions.[10] For objective evaluation of technical ability, we have developed surgery specific metric in collaboration with expert arthroscopy surgeons to minimize subjectivity. The derivation scoring metrics was used in surgery video time analysis. We also analyzed our metrics with previously proposed rubrics.[16,17] Our rubrics will be used to perform construct validation studies to evaluate the effectiveness of the VATDEP as a training platform.

In this work, our contribution is as follows: (a) derivation of hierarchical task analysis (HTA) tree specific to rotator cuff surgery; (b) development of quantifiable metrics for performance assessment; (c) time and performance analysis of actual surgery videos of expert surgeons at different skill levels with unbiased raters; further (d) analysis of our results with early proposed work.

2 | METHODS

2.1 | Procedure analysis

Hierarchical task analysis (HTA) describes the details of all surgery necessary and optional tasks and actions taken to achieve goals. In addition, HTA demonstrates task relations and their execution sequence. The HTA in this work was derived via consultation with expert surgeons that routinely perform shoulder arthroscopy. In HTA, three main phases were identified: (a) pre-procedure; (b) start of examination; and (c) start of repair procedure. In Figure 1, these three main phases are illustrated. The rectangle diagrams refer to the tasks needed to be performed. The forward arrows indicate linear progression of the procedure. The branching in subtasks simply specifies optional paths that can be taken based on preference.

The shoulder arthroscopy procedure starts with preparation of the patient. The patient is first given general anesthesia. This is followed by positioning in the lateral decubitus or the ‘beach chair’ position.[18] The next step is preparation and draping the patient. In this phase, the patient's shoulder is disinfected and draped. Once draping is completed, the mobility of the patient arm is externally examined in detail. After examination, skin markings for landmarks (e.g. acromion, clavicle, and the coracoid possible portal locations) are marked on the shoulder with a sterile pen. The next phase is the examination phase where the shoulder undergoes detailed inspection with the arthroscope. First, arthroscope vision is established by making a 5 mm incision posteriorly with continuous irrigation of the joint with saline to create expansion in the shoulder volume. This continuous fluid flow is provided with either gravity or a pump system attached to the arthroscope. The incoming and outgoing fluid flow creates pressure differences which are important to establish clear arthroscopic view. Otherwise, the view can be obstructed with bleeding, air bubbles, and body fluids such as synovial fluids or debris (during the cleaning task). Preceding the examination or during intra-examination, additional portals can be needed. These portals can be established either using inside-out or outside-in technique. Inside-out technique is performed by pushing the

WILEY The International Journal of Medical Robotics and Computer Assisted Surgery

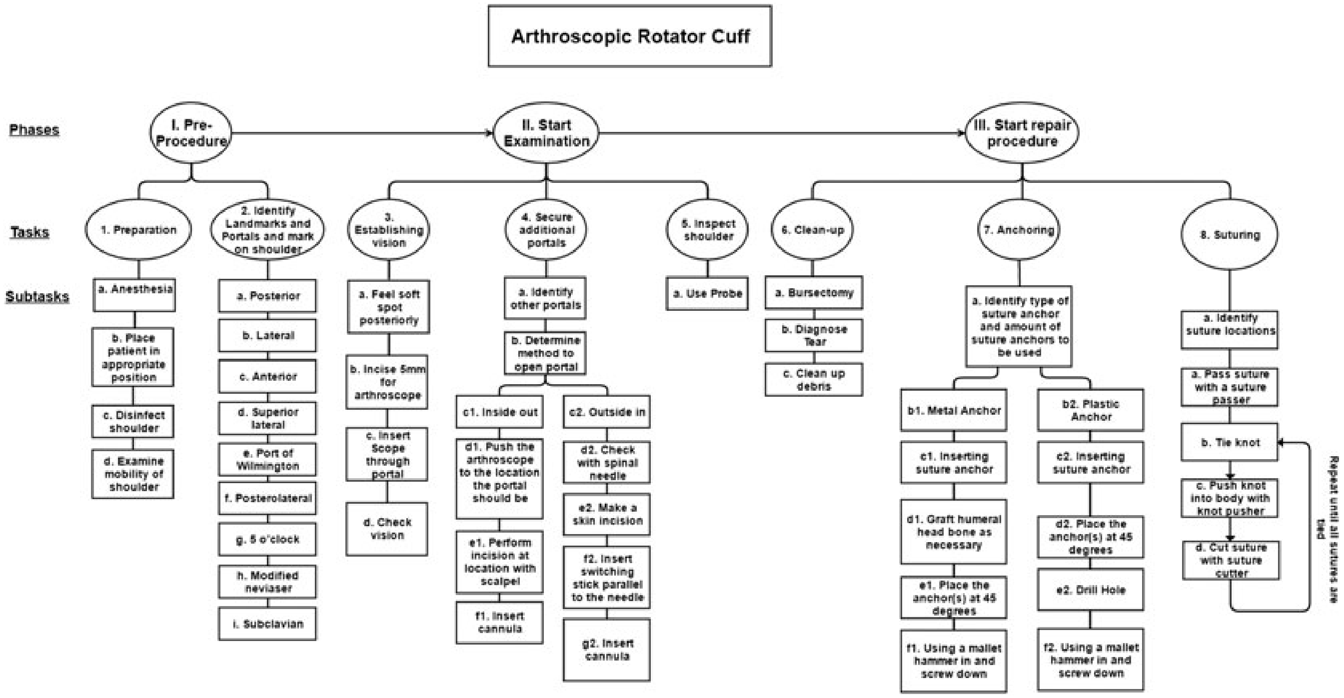

FIGURE 1 Hierarchical task analysis tree

arthroscope to the designated portal location through an incision using a scalpel and inserting a cannula. Outside-in technique is performed by first checking the location with a spinal needle, then creating a skin incision at the location of the needle. A switching stick is inserted parallel to the needle and then the cannula is inserted through the incision. During the examination of the patient anatomy, a probe is used to assist with the inspection to manipulate the muscles, tendons, and ligaments to have a better arthroscopic view. It is essential to examine the injured area and assess the size of the tear if present. The next phase following the detailed examination is the repair phase. Bursal tissue is cleaned using a shaver and/or an electrocautery. The cleaning phase allows clear view and better assessment of the tear. Figure 2A and B shows the view before and after cleaning, respectively. After assessment, the type of suture anchors and number of sutures are determined.

As part of the tear repair, the suture anchor is inserted through the lateral port or a separate incision. The anchor is placed on the footprint at a 45° angle away from the tear. After the suture anchor is placed, the suture locations need to be identified. The suture location should be distributed evenly to ensure that the muscle with tear can be secured to the anchor as firmly as possible. This is critical for uniform distribution of the tension forces to the anchors. The sutures are deployed using a suture passer and then tied in square knots. This process is repeated until the knots are secured. After knots are secured, the sutures are cut using a suture cutter. After repair the arthroscope and the instruments are taken out and the incisions are closed to complete the procedure.

2.2 | Scoring metric

Our surgery metrics, unlike other existing metrics[1617] which are generic or subjective, focuses on the task specific scoring with minimal subjectivity. The score is based on a likert scale[19]; the best/preferred action is scored as 5, an acceptable but less ideal as 3, not adequate as 1, and a failing maneuver is given a score of zero. Details of the scoring system are shown in Table 1–4.

In arthroscopy, correct pressure of the fluid is essential. High pressure can cause swelling while low pressure can make it difficult to control the bleeding. Too much bleeding and swelling can make the surgery more challenging or even impede the surgeon from completing the procedure. Control of the pressure and avoiding bleeding attains the highest score, which otherwise could easily obstruct the view. Failing to complete the procedure owing to excessive swelling and/or bleeding will result in an overall failing score.

A posterior portal is generally used for the arthroscope. After opening of the posterior portal using the outside-in technique an anterior portal is established.[20] Any other portals can be established if necessary. Excess portals or not properly established portals will lower the performance score. During the inspection phase, the surgeon needs to identify major anatomical landmarks. This is important to determine the spatial location and navigation. Failing to identify each one of these landmarks results in loss of points on the task checklist in the grading metric.

During bursectomy, if cleaning is not sufficient (see Figure 2A, and B for inadequate and adequate cleaning), the surgeon must perform a re-cleaning process. Excessive aggressive burring of the bone can weaken the stability of suture anchor, affect healing of the tendons, cause bleeding and prolong the procedure time.

Placement of the anchors is very critical for tear treatment. The anchors should be placed at a 45 degree angle to the footprint, which is the optimal angle for the stability of anchors. Improper placement of the anchor, extreme angles (e.g. 60–75°) receive one point. Once the sutures are passed through the tendon, square knots should be used to secure the repair. A second row of sutures can also be used to perform double row rotator cuff repair.

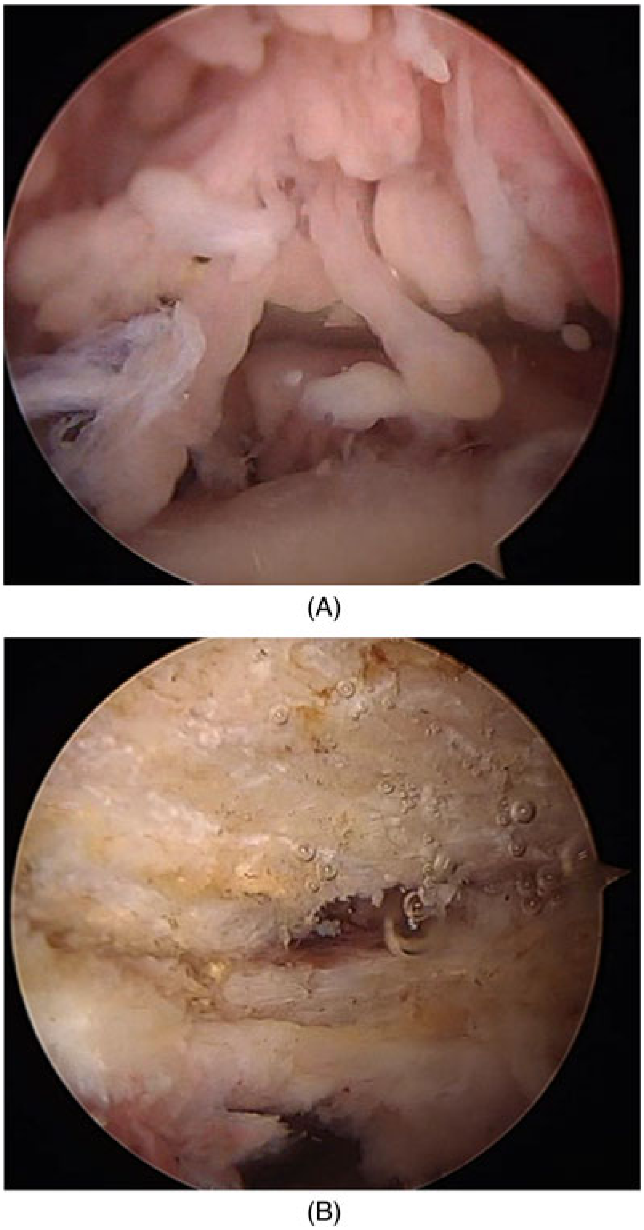

FIGURE 2 A, Before cleaning; B, after cleaning

In our study, we considered two existing rubrics for further analysis of our metrics; the arthroscopic surgical skill evaluation tool (ASSET) global rating scale[16] and the basic arthroscopic knee skill scoring system (BAKSSS).[17] We have merged these two scoring systems to create one unified metric to analyze our metrics.

2.3 | Video timing analysis methodology

In the timing analysis, each surgery video was analyzed using Windows and VLC media players. We determined each major task time defined in HTA. In order to have consistent timings among the raters, we identified the start and end times for the timed tasks (Table 5). Scores and times for each rater were collected for statistical analysis. An inter-rater reliability test was performed to see the degree of agreement between the raters using IBM SPSS 22 software.

TABLE 1 Preparation for the surgery

Technique	Scoring
Pressure of fluid	
No swelling or bleeding	5
High pressure- swelling	
Able to finish	3
Not able to finish	0 (Fail)
Low pressure- bleeding	
Able to stop bleeding directly and promptly	5
Able to stop bleeding and finish surgery	3
Not able to finish	0(Fail)
Number of portals	
Necessary portals	5
Unnecessary portals	3
Unable to establish appropriate portals	1
Identifying / establishing portals	
Proper technique	5
Improper technique	3
Damaging tendon or other muscles	0

TABLE 2 Task checklist for inspection of the rotator cuff area

Task checklist	1 or 0
Inspect Superior Labrum and Biceps	
Inspect Anterior Labrum and Capsule	
Inspect rotator cuff muscles (Supraspinatus, Infraspinatus, teres minor, subscapularis)	
Inspect Glenohumeral ligaments (middle, inferior)	
Inspect Rotator Interval (space situated between the supraspinatus and subscapularis tendons)	
Inspect Glenoid	
Inspect Humeral Head	

TABLE 3 Repairing procedure for the rotator cuff surgery

Technique	Scoring
Bursectomy	
Adequate cleaning	5
Aggressive bursectomy(bleeding)	3
Not enough cleaning (need to go back and reclean)	1
Debris cleaning	
Sufficient cleaning	5
Not adequate cleaning (e.g. increase healing time)	3

2.4 | Surgeon questionnaire

After completion of each procedure, expert surgeons were asked to fill out a questionnaire regarding the details of the procedure. The aim of the questionnaire is to systematically categorize the surgeries to understand the impact of the type, location, size of the tear, suturing and cleaning on difficulty levels, time and also performance ratings. Moreover, we would like to identify any rare cases (e.g. surgeons were asked to put additional feedback such as any extreme cases, unusual treatment or procedure applied) in the surgery that might have had

TABLE 4 Suturing procedure for the rotator cuff surgery

Technique	Scoring
Burring the bone	
Adequate burring	5
Aggressive burring	0
Position of anchor	
Evenly distanced	5
Not evenly distanced	1
Placing of the anchor	
45 degrees	5
30–45 or 45–60 degrees	3
15–30 or 60–75 degrees	1
Other	0
Identification of suture locations	
Proper repair	5
Sutures are unable to repair the tear completely	3
Unable to establish suture on rotator cuff within 3 attempts	0
Knot tying	
Square knot	5
Not square knot	1
Unable to create knot	0

TABLE 5 Start and end times for each timed task

Task name	Start time	End time
Diagnostic	Start using probe	Stop using probe
Pre-cleaning	End of diagnostic stage	Placement of suture anchor
Anchor suture	Hammering of anchor	Final suture is tied
Cleaning	Final suture is tied	End of Surgery
Total cleaning	Pre- cleaning and cleaning time	
Total surgery	Start of procedure	End of procedure

negative impact on the task times (e.g. unexpected prolonged tasks times). The sample questionnaire is shown in Table 6.

3 | RESULTS

Fourteen shoulder arthroscopy videos (total ≈10 h length) performed by expert surgeons were used in the analysis. Of 14 surgeries, two were full thickness, one was L-shaped and 11 were crescent shaped tears. All tears were located on the Supraspinatus and the tear sizes varied from 1 cm to 3 cm.

TABLE 6 Surgery questionnaire form

Type of tear: ___________________________________
Location of tear: _________________________________
Size of tear: ____________________________________
Un-anchored sutures required (if any): ____
Anchored sutures required: ___
Cleaning done (before / after) suturing (circle one).
Overall difficulty of procedure: 1 2 3 4 5
(1: significantly easier than usual, 3: as expected, 5: significantly harder than usual)
Amount of scar tissue: 1 2 3 4 5
Amount of cleaning necessary: 1 2 3 4 5
Difficulty passing sutures: 1 2 3 4 5
Difficulty tying sutures: 1 2 3 4 5
Any additional comments/feedback about the video (any information that you would like to note):

All the videos were examined by four raters. Raters had shadowed the surgery in OR many times and were competent in the details of procedure and the HTA. Precise guidelines were given to the raters for the task time measurements. Once timings of each video were completed, raters were asked to measure the scores using our metrics.

Pearson's correlation test was conducted to find any correlation between the times and surgeon questionnaires. The correlation test results showed a strong positive correlation ($R = 0.8823$) with a statistical significance ($P = 0.00003$) between the difficulty of procedure and total surgery time. Similarly, more tasks had strong positive correlations. These tasks are shown in Table 7.

Figures 3–6, respectively, show time and rating comparisons in percentage values for difficulty of procedure and total surgery time, difficulty of passing sutures and cleaning time, amount of scar tissue and cleaning time, difficulty of tying suture and cleaning time. Figure 7 shows a comparison of ratings in percentage values for difficulty passing suture and difficulty tying suture.

In our study, we measured diagnostic time; the average time that the probe is active in the scene is 118.8 s. The pre-clean/prep time, that is the time after the diagnostic and includes debris cleaning, the bursectomy and any bone burring that is performed prior to placement of the suture anchors. The average time for this task is measured as 607 s. The tool efficiency refers to the active time of the tool while inside the shoulder divided by the total time that the tool is in the body. We recorded the efficiency of the shaver tool and the electrocautery tool at 86.4% and 87.9%, respectively. Suture timing consists of several substeps; time to place the anchor, time to pass the suture, time to secure the knot, and transition time. These timings indicate the per suture timings and in most cases there exist three sutures per anchor. Per suture average timings are time to establish anchor, 50.13 s; time to punch anchor, 38.63 s; time to pass sutures, 85.1 s; time to secure suture knots, 136.56 s; transition time per suture, 14.2 s; and transition time per suture. Transition time, the percentage time spent not performing a notable task during the suturing and includes tasks that may not be required for every single suture such as time to grab the suture within the body and additional cleaning. Our results have shown that anchor punching time is relatively independent, with times spanning 17 to 71 s with little to no relation to other tasks. Anchor punching time and anchor placement time, respectively, indicate the time spent to hammer in the anchor and time spent to place the anchor in the cavity created on the humeral head.

TABLE 7 Correlation results

	Correlation task/s and time/s	Pearson correlation (R)	Significance (P)
Strong Positive Correlation	Difficulty of passing sutures and cleaning time	0.8359	0.000196
	Amount of scar tissue and cleaning time	0.8136	0.0004
	Difficulty of tying suture and cleaning time	0.9065	<0.00001
	Difficulty of passing suture and difficulty tying suture	0.8343	0. 000207

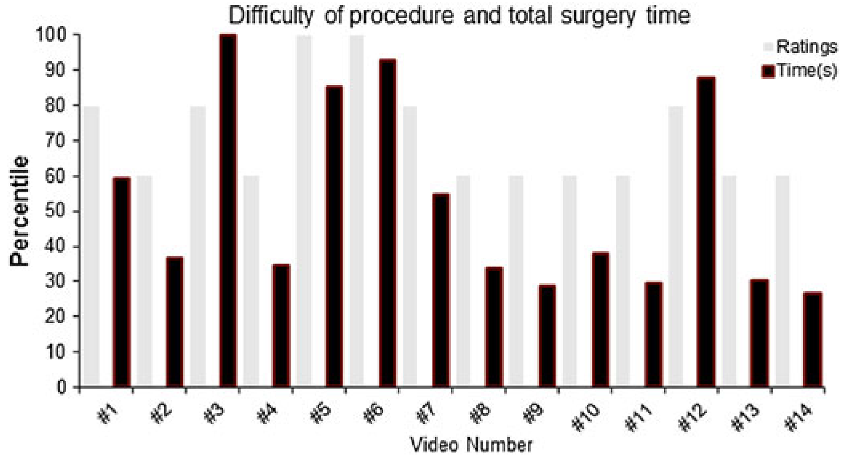

FIGURE 3 Time and rating comparison plot in percentage values for difficulty of procedure and total surgery time

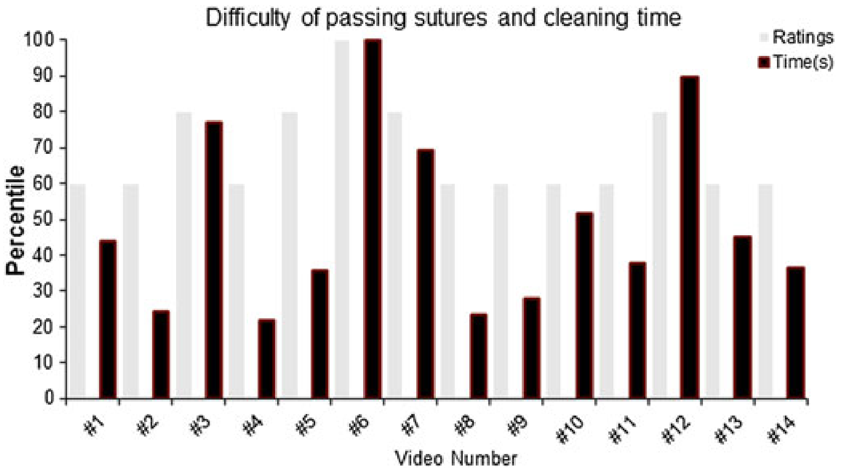

FIGURE 4 Time and rating comparison plot in percentage values for difficulty of passing sutures and total cleaning time

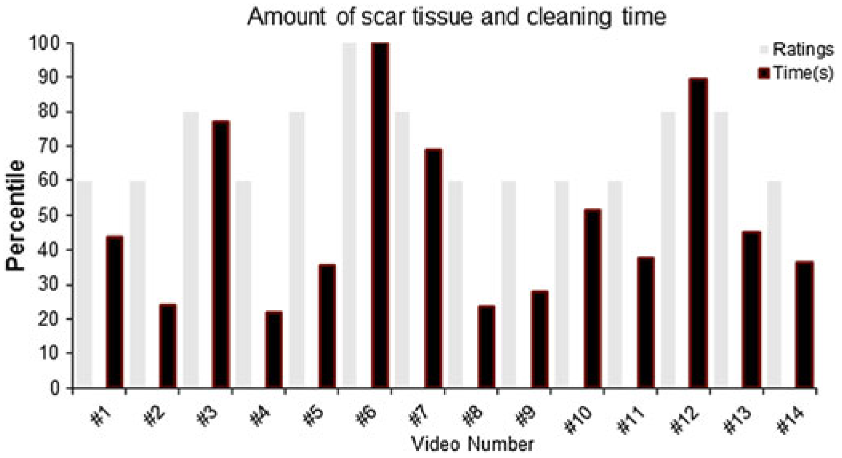

FIGURE 5 Time and rating comparison plot in percentage values for amount of scar tissue and total cleaning time

Task efficiency for the anchor suture task is calculated by the total time spent for anchor placement and suturing tasks divided by total anchor suture task time. The average task efficiency for the anchor suture task was 84.59%. The task efficiency measurement is computed as a measure to evaluate proficiency of the surgeon in the task. Figure 8 shows anchor suture task efficiency for each video.

Cleaning time represents the total time spent cleaning the bursal tissue and the footprint after the surgeon completes the suturing. Average total cleaning time was 786.36 s. Figure 9 shows the box plots for each timed task in seconds for expert surgeons.

In some certain cases, there were drawbacks in assessing the expert surgeon experience level using our grading metric and also in the existing metrics. To overcome this, we have extended our scoring

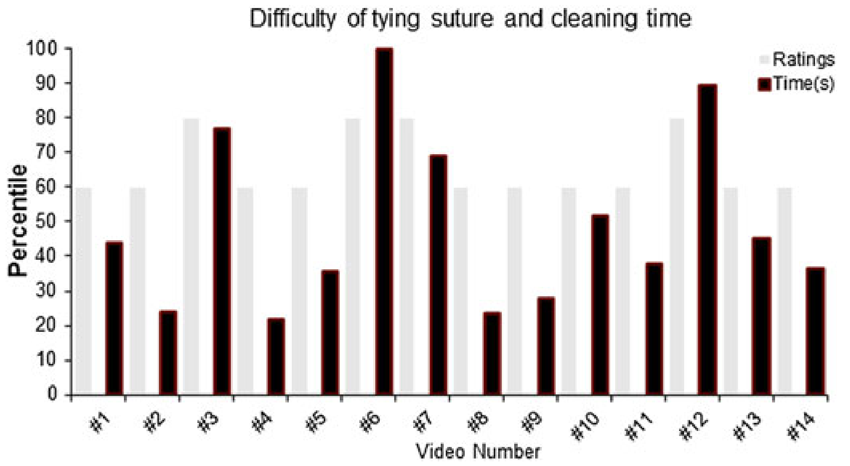

FIGURE 6 Time and rating comparison plot in percentage values for difficulty of tying suture and total cleaning time

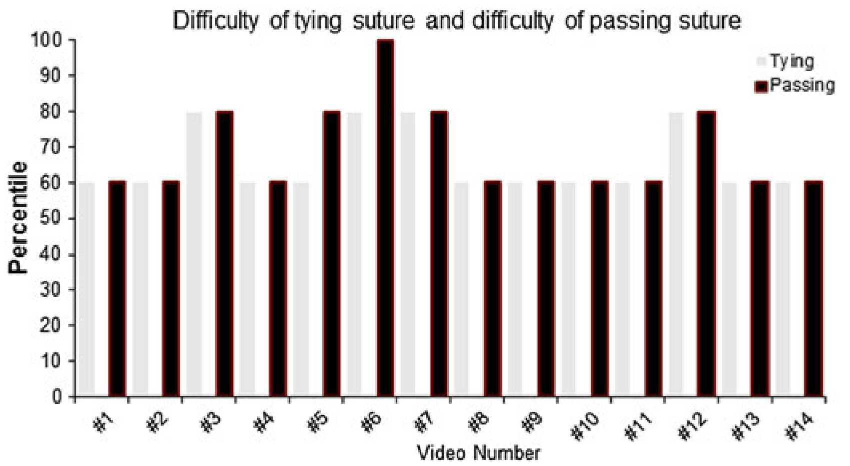

FIGURE 7 Rating and rating comparison plot in percentage values for difficulty of tying suture and difficulty of passing suture

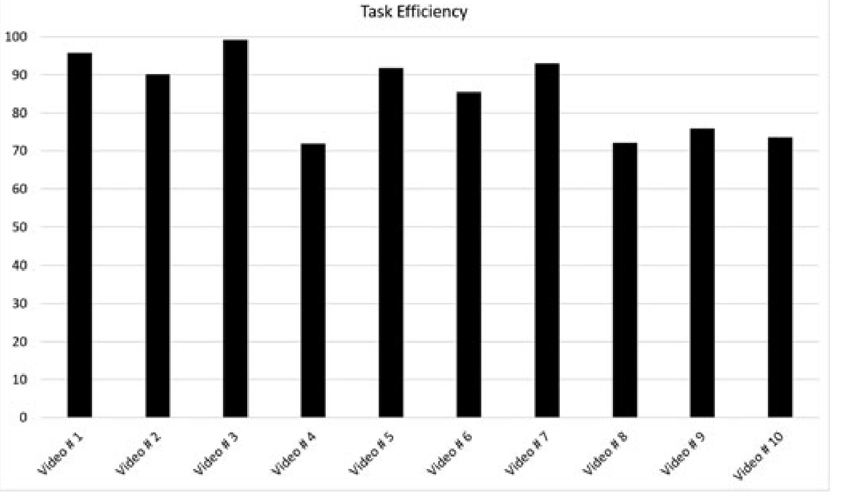

FIGURE 8 Anchor suture task efficiency for each video

system. The extension to our metric was developed by relating the task scores with task times. Each task is clearly defined by a series of events that can be easily identified by a lay person with minimal training. Each task is assigned an optimal completion time, established based on the average task time. The optimal time required to complete a task was calculated by $A_T + (1.5^*\sigma)$, where A_T is the average time and σ is the standard deviation of the specific task. Time intervals are created based on the complexity and duration of the task. If the surgeon is able to complete the task within the optimal time, a score of 5 is given. If the amount of time exceeds the optimal time, 1 point is taken off for each time interval exceeding the optimal.

Each video received maximum scores in preparation step and inspection checklist received maximum scores for each video. The average score acquired for the preparation step, and inspection

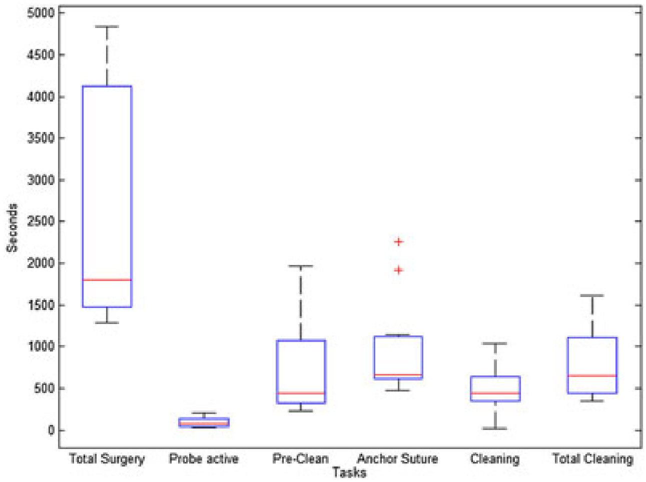

FIGURE 9 Box plots for expert surgeons for rotator cuff surgery tasks in seconds

checklist were 15 and 7, respectively. The average procedure score was 8.6 (max: 10, min: 2), while for the suturing the average grade was 28.5 (max: 30, min: 24). Overall average score for the videos were 59.1 (max: 62, min: 51). Figure 10 shows the box plots for each metric component scored for expert surgeons.

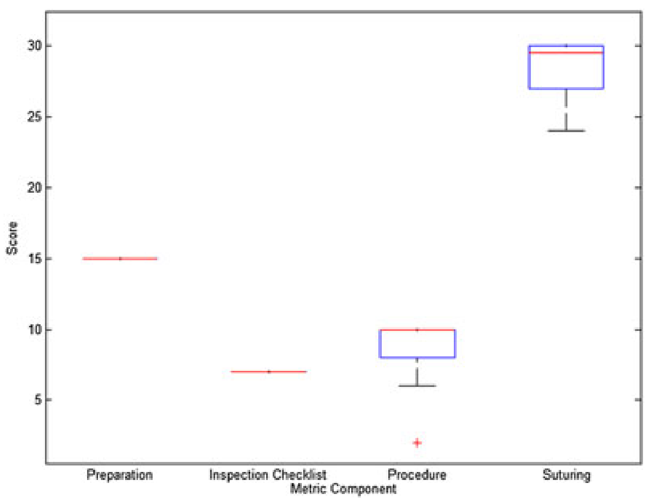

FIGURE 10 Box plots for each scored metric component for expert surgeons

3.1 | Analysis with BAKSSS and ASSET metrics systems

BAKSSS and ASSET metrics comprised of general categories rather than task specific categories. In order to perform a comparison study, we mapped each category to at least one of our scoring metrics. Based on this mapping matrix (see Table 8), our scoring metric results were populated in the unified (BAKSSS and ASSET) metric by averaging the correlated scoring metric categories for each merged metric category.

Field of view, camera dexterity, knowledge of procedure, quality of procedure and autonomy received maximum scores[5] for each video. The average score acquired for the instrument, and bi-manual dexterity were 4.62 and 4.5 respectively. The average efficiency score was 4.44 (max: 5, min: 3), while for flow of procedure the average grade was 4.27 (max: 5, min: 3). Overall average score for the videos were 52.72 (max: 55, min: 50.6). Figure 11 shows the box plots for merged metric components: instrument dexterity, bi-manual dexterity, efficiency and flow of procedure for expert surgeons.

4 | DISCUSSION

Interrater-reliability test results showed a perfect degree of agreement among raters for each task. This is due to the consistency among the rater's timings, which showed negligible deviation (e.g. couple of seconds for each task, average standard deviation per video, $\sigma = 1.124$). Pearson's correlation test shows there is strong correlation between the cleaning time and the difficulty in passing the sutures. It is expected that a longer time spent on adequate cleaning would streamline the subsequent processes and make the suture passing task easier. As seen in the time–score correlation, the difficulty of the procedure increases the cleaning time. Therefore, the cleaning time and suture passing time correlation stem from the difficulty of the task. In other words, both cleaning time and suture passing time increase as a result of a difficult case.

Correlation between the time and the amount of adequate cleaning in the questionnaire could conclude that cleaning more than an adequate amount does not affect the performance rating. Overall cleaning time had an indirect impact on the difficulty of passing sutures, the amount of scar tissue and the difficulty tying sutures. These correlations are results of excess scar tissue caused by large sized tears.

TABLE 8 The mapping of our scoring metrics and times for each BAKSSS and ASSET metric

	Scoring Metrics	Metrics based on BAKSSS and ASSET	Field of view	Camera dexterity	Instrument dexterity	Bi-manual dexterity	Efficiency	Flow of procedure
1	Identification of suture locations		X	X		X		
2	Inspection checklist		X	X				
3	Number of Portals			X				
4	Identifying/establishing portals			X				
5	Bursectomy				X		X	
6	Debris cleaning				X		X	
7	Preparing the bone						X	
8	Knot tying				X	X	X	X
9	Suture passing efficiency					X	X	X
10	Transition speed							X
11	Position of anchor				X			

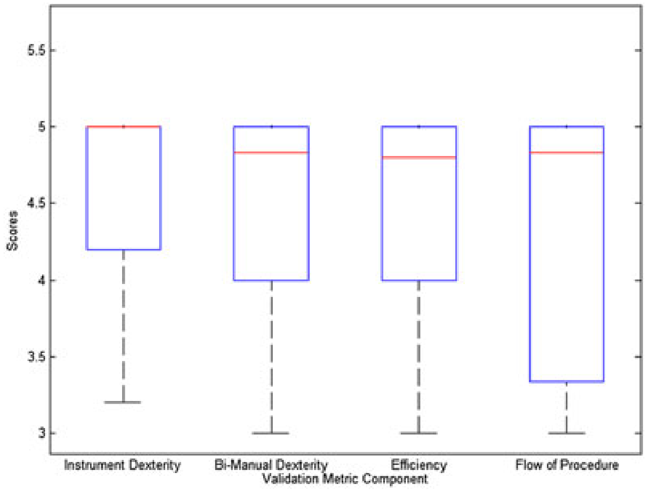

FIGURE 11 Box plots for merged matric components: instrument dexterity, bi-manual dexterity, efficiency and flow of procedure

Tool and task efficiency results were all >84%, evidencing that the surgeons are proficient in using the tools and completing the tasks. Higher percentages should indicate that the surgeon is more efficient with the tool and that the surgeon has a better awareness of the flow of the procedure and are an indication of smooth tool use. In cases where the tool is switched constantly and rapidly (< 10 s) the efficiency is not recorded.

In one of the surgery videos, knot tying score was computed as a score of one using the extension of our metric. This was due to technical difficulties with the suture passer tool and debris getting caught in between the knots. The suture passer failed to shoot up the suture multiple times and the suture passer had to be reloaded. As our extended metrics considers the average task times, the failure in suturing in this task for even an expert surgeon is innately determined.

Positioning of anchor was scored as 5 in every video due to the proper location and angle of the anchor. In the BAKSSS and ASSET metrics, knowledge of instruments, knowledge of specific procedure and quality of procedure were scored as 5 (in post timing and score computation phase) due to all expert surgeons having performed approximately 150 arthroscopic rotator cuff surgeries in the last 6 months. Autonomy knowledge was also scored as 5 due to each surgery having completed the examination phase and surgeries without any intermittence during the surgery.

Although this study involves only videos analysis of the rotator arthroscopic videos and does not contain any patient and private information at any phase, we had submitted the proposal of the study to the Institutional Review Board (IRB) at University of Arkansas for Medical Sciences. The IRB committee determined that the study is classified as not a human subject study as defined in federal regulations stated in 45 CFR 46.102, which had allowed the study.

5 | CONCLUSION

As the preliminary phase of developing VATDEP, we presented a HTA for arthroscopic rotator cuff surgery. We have derived a specific grading metric from the HTA for forthcoming use in our VR simulator for the construct validation studies. We performed video analysis of the actual surgery videos. As a result of analysis, with our metrics system, the surgery performance skills of a surgeon can be computed. In addition, the variation of the expert level can be determined with the time related task scores.

Currently, residents in orthopaedic training cannot perform the whole procedure without attending expert surgeon involvement. This precludes us to utilize surgery videos with resident involvement and compare with the expert videos. Therefore, we are currently timing the task of each resident involvement for that task. As a future work, this will enable us to perform unified objective metric framework to assess and compute the performance of surgeons with different skill levels.

COMPETING INTERESTS

The authors declare that they have no competing interests in regards to this study.

ACKNOWLEDGEMENTS

This publication was made possible by the Arkansas INBRE program, supported by a grant from the National Institute of General Medical Sciences, (NIGMS), P20 GM103429 from the National Institutes of Health. This project was also supported by National Institutes of Health (NIH) Grant NIH/NHLBI 1R01HL119248-01A1, NIH/NIBIB 2R01EB005807, 5R01EB010037, 1R01EB009362, and 1R01EB014305.

REFERENCES

1. Treuting R. Minimally invasive orthopedic surgery: Arthroscopy. *Ochsner J*. 2000 Jul;2(3):158–163.
2. Iannotti JP. Full-thickness rotator cuff tears: Factors affecting surgical outcome. *J Am Acad Orthop Surg*. 1994;2(2):87–95.
3. Abrams JS, Savoie III FH. Arthroscopic rotator cuff repair: Is it the new gold standard. 72nd Annual Meeting Proceedings, American Academy of Orthopaedic Surgeons. 2005;71.
4. Fineberg M, Wind W, Stoeckl A, Smolinski R. Current practices and opinions in shoulder surgery: Results of a survey of members of the American Orthopaedic Society for Sports Medicine and the Arthroscopy Association of North America. 22nd Annual Meeting Abstracts of Podium Presentations. 2003;19.
5. Lyu S-R, Lin Y-K, Huang S-T, Yau H-T. Experience-based virtual training system for knee arthroscopic inspection. *Biomed Eng Online*. 2013;12:63
6. Tashiro Y, Miura H, Nakanishi Y, et al. Evaluation of skills in arthroscopic training based on trajectory and force data. *Clin Orthop*. 2009 Feb;467(2):546–552.
7. ACGME > What We Do > Accreditation > Common Program Requirements [Internet]. [cited 2016 May 6]. Available from: http://www.acgme.org/What-We-Do/Accreditation/Common-Program-Requirements
8. Howells NR, Gill HS, Carr AJ, et al. Transferring simulated arthroscopic skills to the operating theatre – a randomised blinded study. *J Bone Joint Surg Br*. 2008;90(4):494–499.
9. Gomoll AH, O'Toole RV, Czarnecki J, Warner JJ. Surgical experience correlates with performance on a virtual reality simulator for shoulder arthroscopy. *Am J Sports Med*. 2007;35(6):883–888.
10. Demirel D, Butler KL, Halic T, et al. A hierarchical task analysis of cricothyroidotomy procedure for a virtual airway skills trainer (VAST) simulator. *Am J Surg* [Internet]. [cited 2015 Nov 11];

Available from: http://www.sciencedirect.com/science/article/pii/S0002961015005644

11. Fried MP, Satava R, Weghorst S, et al. The use of surgical simulators to reduce errors. In *Advances in Patient Safety: From Research to Implementation*, (Volume 4: Programs, Tools, and Products), Henriksen K, Battles JB, Marks ES, Lewin DI, eds. [Internet]. Rockville (MD): Agency for Healthcare Research and Quality (US); 2005 [cited 2015 Aug 23]. (Advances in Patient Safety). Available from: http://www.ncbi.nlm.nih.gov/books/NBK20604/
12. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance. *Ann Surg*. 2002 Oct;236(4):458–464.
13. Grantcharov TP, Kristiansen VB, Bendix J, et al. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg*. 2004 Feb;91(2):146–150.
14. Rosser JC, Rosser LE, Savalgi RS. Skill acquisition and assessment for laparoscopic surgery. *Arch Surg Chic Ill 1960*. 1997 Feb;132(2):200–204.
15. Hodgins JL, Veillette C. Arthroscopic proficiency: Methods in evaluating competency. *BMC Med Educ*. 2013;13:61
16. Koehler RJ, Amsdell S, Arendt EA, et al. The arthroscopic surgical skill evaluation tool (ASSET). *Am J Sports Med*. 2013 Jun 1;41(6):1229–1237.
17. Olson T, Koehler R, Butler A, et al. Is there a valid and reliable assessment of diagnostic knee arthroscopy skill? *Clin Orthop*. 2013 May;471(5):1670–1676.
18. Peruto CM, Ciccotti MG, Cohen SB. Shoulder arthroscopy positioning: lateral decubitus versus beach chair. *Arthroscopy*. 2009 Aug 1;25(8):891–896.
19. Maurer TJ, Pierce HR. A comparison of Likert scale and traditional measures of self-efficacy. *J Appl Psychol*. 1998;83(2):324–329.
20. Burkhart SS, Lo IK, Brady PC. Burkhart's view of the shoulder: A cowboy's guide to advanced shoulder arthroscopy. Lippincott Williams & Wilkins, ISBN: 978-0781780001, 2006; 10.

How to cite this article: Demirel D, Yu A, Cooper-Baer S, Dendukuri A, Halic T, Kockara S, Kockara N, Ahmadi S. A hierarchical task analysis of shoulder arthroscopy for a virtual arthroscopic tear diagnosis and evaluation platform (VATDEP). *Int J Med Robotics Comput Assist Surg*. 2016;e1799. doi: 10.1002/rcs.1799