TOPICS IN TRAINING

Shoulder Arthroscopy Simulator Performance Correlates with Resident and Shoulder Arthroscopy Experience

Kevin D. Martin, DO, Kenneth Cameron, PhD, MPH, ATC, Philip J. Belmont Jr., MD,
Andrew Schoenfeld, MD, and Brett D. Owens, MD

Background: The technical skills required to perform arthroscopy are multifaceted and require supervised training and repetition. Obtaining this basic arthroscopic skill set can be costly and time-consuming. Simulation may represent a viable training source for basic arthroscopic skills. Our goal was to evaluate the correlation between timed task performance on an arthroscopic shoulder simulator and both resident experience and shoulder arthroscopy experience.

Methods: Twenty-seven residents were voluntarily recruited from an orthopaedic residency program. Each subject was tested annually for three consecutive years on an arthroscopic shoulder simulator and objectively scored on time to completion of a standardized object localization task. Each subject's total number of shoulder arthroscopies, all arthroscopies, and cases were calculated according to postgraduate year from their Accreditation Council for Graduate Medical Education (ACGME) case log. Generalized estimating equation multivariate regression analysis was performed to determine the correlation between simulation performance and total numbers of shoulder arthroscopies, all arthroscopies, and cases.

Results: Univariate analyses revealed that postgraduate year, total number of shoulder arthroscopies, total number of arthroscopies of any joint, and total number of surgical cases performed during residency training prior to testing were associated with the mean time required to complete the simulator task. The number of prior shoulder arthroscopies performed ($r = 0.55$) and postgraduate year in training ($r = 0.60$) correlated most strongly with simulator basic task performance. In the multivariate analysis, the number of prior shoulder arthroscopies and postgraduate year remained independent predictors of faster completion of the simulator task. For every additional postgraduate year, there was a sixteen-second improvement in the time required to complete the simulator task ($p < 0.005$). Similarly, after controlling for the influence of postgraduate year, there was a twelve-second decrease in the time to complete the simulator task for every additional fifty shoulder arthroscopies performed during residency training ($p < 0.008$).

Conclusions: These results showed a significant relationship between performance of basic arthroscopic tasks in a simulator model and the number of shoulder arthroscopies performed. The data confirmed our hypothesis that simulator performance is representative of both resident experience and shoulder arthroscopy experience.

Clinical Relevance: This study suggests that greater resident clinical experience and shoulder arthroscopy experience are both reflected in improved performance of basic tasks on a shoulder simulator. These findings warrant further investigation to determine if training on a validated arthroscopic shoulder simulator would improve clinical arthroscopic skills.

Disclosure: None of the authors received payments or services, either directly or indirectly (i.e., via his or her institution), from a third party in support of any aspect of this work. One or more of the authors, or his or her institution, has had a financial relationship, in the thirty-six months prior to submission of this work, with an entity in the biomedical arena that could be perceived to influence or have the potential to influence what is written in this work. No author has had any other relationships, or has engaged in any other activities, that could be perceived to influence or have the potential to influence what is written in this work. The complete **Disclosures of Potential Conflicts of Interest** submitted by authors are always provided with the online version of the article.

Disclaimer: All of the authors of this manuscript are members of the United States military and/or employees of the U.S. government. The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or reflecting the views of the Department of the Army, Department of Defense, or the United States government.

The development of an orthopaedic resident's arthroscopic surgical skills requires supervised training and substantial repetition, both of which can be costly and time-consuming within the operating room[1,2]. The act of manipulating tissue and equipment while relying on a two-dimensional screen feels unnatural to most novice arthroscopists. This has led many educational bodies, including the American Academy of Orthopaedic Surgeons, to research the use of simulators and their potential role in training residents[3-5].

This research has been limited because of unrealistic simulators and the lack of validated simulation models. We previously compared the performance of basic arthroscopic tasks using the insightARTHRO VR shoulder simulator (GMV, Madrid, Spain) with the performance of the same tasks in a cadaveric model[6]. However, the link between clinical skill and simulator performance is still unclear. This previous investigation demonstrated a correlation between surgical experience and simulator performance; however, an important limitation, shared by other studies in this area, involved the ability to accurately define true surgical experience[6]. Many papers have described surgical experience with use of self-reported estimates of how many total arthroscopic procedures had been performed, without specifying their anatomic location[7,8]. Others have used generalizations such as "novice" or "expert," which are used to loosely define clinical experience[9-11]. Both methods of estimating surgical experience are fraught with subjective bias and inaccuracies, potentially compromising validity and reproducibility.

Orthopaedic surgical training has traditionally been based on the apprenticeship model, but the training of orthopaedic surgery residents has never been more regulated than under the current educational paradigm[12]. The Accreditation Council for Graduate Medical Education (ACGME) resident work hour guidelines have created training challenges that affect all residency programs in the United States. These work hour guidelines have brought with them a renewed focus on how to best address deficiencies in both the knowledge base and surgical dexterity skills of orthopaedic surgery residents. As the work hours and surgical experience of residents continue to decrease[13], the field of arthroscopy has arguably become increasingly challenging, with more technically demanding procedures being performed[14]. This phenomenon has forced educators to seek other means of training in basic arthroscopic skills in a time-effective manner.

The purpose of this study was to determine the association between surgical experience of residents, defined by the ACGME case logs of each participant, and basic arthroscopic task performance in a simulator environment. We hypothesized that basic arthroscopic task performance in the simulator model would positively correlate with both resident experience and shoulder arthroscopy experience.

Materials and Methods

Subjects

Twenty-seven orthopaedic residents in a range of postgraduate years took part in the study, which was conducted between August 2008 and May 2011. All subjects were within the same residency program and were voluntarily recruited to participate regardless of surgical experience. No subjects had prior direct simulation experience on enrollment in the current study. Test subjects were categorized by year in training, ranging from postgraduate year one through five, at the time of testing. Demographic information including postgraduate year, sex, and age were obtained at the initial testing session. Institutional review board approval was obtained before testing commenced.

Simulator

This study used the insightArthro VR shoulder simulator for all testing throughout the study. Fidelity, interactivity, objective properties, sensory input, and reactivity are all vital to simulation studies; the importance of these features has previously been described by Satava[15] and are present within this simulator model. The simulator has a high-definition monitor that provides high-fidelity resolution, allowing objects to appear real to the test subject. This simulator has two robotic arms equipped with force-reflective technology, providing interactive feedback via model arthroscopic equipment. A simple probe and a 30° arthroscopic camera were used for the purpose of this study. Each arm provides subjects with tactile and haptic feedback in response to interactive tissue manipulation and gravity. The tissue within the model also displays reactivity, i.e., it can be cut, handled, and deformed[15].

Testing was performed with use of the simulator's diagnostic "Blue Sphere" program. This program utilizes blue spheres located at key diagnostic anatomic landmarks within the glenohumeral joint. After the subject locates and palpates the sphere with the probe, the sphere then relocates to a different anatomic location, reproducing a diagnostic arthroscopy in which eleven key structures are identified. This program was selected because of its documented correlation with surgical experience and performance in cadaveric testing[6,7].

Testing

Each subject was given a two-minute hands-on orientation to the simulator and its functions and capacities before initial testing began. The subject was then given a five-minute practice session to become familiar with the haptics and anatomic structures. While the subject was practicing, the testing objectives and outcomes were explained and any adjustments to the equipment (including adjustment of the height of the table and the angle of the monitor) were made. Each subject then performed three repetitions of the diagnostic Blue Sphere program, which were evaluated on the basis of the time to completion (in seconds) as well as both the camera distance (in mm) and the probe distance (in mm) traveled during the simulation task. After completion of the three trials, the subject's scores for that year were averaged and recorded. This same testing protocol was repeated annually over a three-year period in late June (the end of academic year). Testing was conducted regardless of each resident's current clinical assignments, and the person administering the simulation task (K.D.M.) was blinded to the recent and cumulative arthroscopic experience of each subject.

Following the completion of the three-year simulation study, a retrospective analysis of each subject's surgical experience was performed. An extensive analysis of each resident's ACGME case log was conducted. For each subject, the total number of shoulder arthroscopies, arthroscopies regardless of anatomic location, and surgical cases were documented from entry into residency training until the date of testing each year. To ensure uniformity and accuracy, each resident's case log was reviewed quarterly throughout the study period by the same author (P.J.B.Jr.).

Statistical Analysis

Means and standard deviations were calculated for all continuous outcome measures and covariates according to program year. Pairwise Pearson correlations between all variables were calculated to examine the association between all outcome and predictor variables. All outcome variables, including the mean time to complete the simulator task, the mean distance that the probe traveled during the task, and the mean distance that the camera traveled during the task, were continuous and approximately normally distributed. All outcome measures were highly correlated, and as a result, we elected to use the mean time to

TABLE I Outcomes Measures and Covariates According to Postgraduate Year

Postgraduate Year	No. of Observations	Simulator Time to Completion* *(sec)*	Simulator Camera Distance* *(mm)*	Simulator Probe Distance* *(mm)*	Total No. of Shoulder Arthroscopies*	Total No. of Arthroscopies*	Total No. of Cases*
1	11	184.7 (69.4)	688.5 (351.7)	1418.7 (533.4)	3.8 (5.3)	18.8 (8.7)	114.3 (41.6)
2	11	120.2 (39.4)	513.4 (337.9)	1033.1 (453.3)	39.8 (23.9)	94.0 (44.9)	537.8 (149.8)
3	8	113.7 (49.7)	388.0 (199.6)	1015.4 (619.9)	68.3 (55.7)	147.0 (80.0)	1274.3 (228.8)
4	9	81.3 (16.9)	279.3 (137.1)	749.7 (153.9)	124.1 (46.5)	187.8 (137.9)	1706.3 (333.1)
5	12	87.6 (28.1)	266.8 (199.9)	957.8 (658.5)	98.8 (43.1)	308.0 (116.1)	2061.3 (298.6)

*Values are given as the mean and the standard deviation.

complete the simulator task as our primary outcome of interest in further analyses.

Because some subjects in the current study completed the simulator task in multiple years, we used the generalized estimating equation (GEE) version of linear regression to account for the correlation among these longitudinal observations of the time to complete the simulator task. The GEE version of linear regression specifies how the mean of an outcome variable for a given subject changes with covariates while accounting for correlations between repeated observations on the same subject over time[16]. We initially examined univariate models with the mean time to complete the simulator task as the primary outcome and postgraduate year, total number of shoulder arthroscopies performed, total number of arthroscopies performed on any joint, and total number of surgical cases performed as predictor variables. We subsequently evaluated multivariate models to determine the model that best fit the data in predicting the mean time to complete the simulator task. Model fit was evaluated with use of the model selection criteria originally described by Pan[17] and the quasi-likelihood under the independence model criterion (QIC). The QIC was also used to evaluate the working correlation structure for the GEE analyses with use of the methods described by Cui[16]. Semirobust standard errors were employed in calculating significance, to adjust for clustering within each subject. All statistical analyses were performed with use of STATA/SE software (version 10.1; StataCorp, College Station, Texas), and a p value of <0.05 was considered significant.

Source of Funding

No external funding was received for this investigation.

Results

The study sample of orthopaedic residents consisted of twenty men and seven women with a mean age of thirty-two years (range, twenty-six to forty-three years). Because of the progressive nature of residency, eleven subjects were tested in only one program year, eight were tested over two program years, and eight were tested over three program years (resulting in a total of fifty-one simulation testing sessions over the three-year study period). As a result, there were a total of eleven observations in the first postgraduate year of residency training, eleven in the second year, eight in the third year, nine in the fourth year, and twelve in the fifth year. Means and standard deviations for all outcome measures and covariates are presented according to postgraduate year in Table I. Unadjusted correlations between all pairs of variables are presented in Table II; notably, simulator time to completion was inversely correlated with the total number of prior shoulder arthroscopies performed ($r = -0.55$).

TABLE II Unadjusted Pairwise Correlation Coefficients Between Outcome Measures and Covariates*

Variable	Postgraduate Year	Total Shoulder Arthroscopies	Total Arthroscopies	Total Surgical Cases	Mean Simulator Time	Mean Camera Distance	Mean Probe Distance
Program year	1.00						
Total shoulder arthroscopies	0.70	1.00					
Total arthroscopies	0.76	0.74	1.00				
Total surgical cases	0.95	0.70	0.79	1.00			
Mean simulator time	−0.60	−0.55	−0.52	−0.62	1.00		
Mean camera distance	−0.52	−0.44	−0.45	−0.54	0.83	1.00	
Mean probe distance	−0.31	−0.31	−0.25	−0.36	0.82	0.81	1.00

*All presented correlations were significant ($p < 0.05$).

TABLE III Univariate and Multivariate GEE Models for the Mean Time to Complete the Simulator Task*

	UnivariateModel			Final Multivariate Model		
Variable	Slope	95% CI	P	Slope	95% CI	P
Postgraduate year	−22.8	−32.5, −13.1	<0.001	−15.9	−26.9, −4.9	0.005
Total shoulder arthroscopies	−0.6	−0.8, −0.3	<0.001	−0.3	−0.5, −0.1	0.008
Total arthroscopies	−0.2	−0.3, −0.1	<0.001			
Total cases	−0.04	−0.06, −0.03	<0.001			

*GEE = generalized estimating equation, and CI = confidence interval. The slope represents the change in simulator performance (measured in seconds) for a one-unit increase in each variable. For example, for every increase of one postgraduate year, the performance time decreased by 15.9 seconds (95% CI, 26.9 to 4.9 seconds) in the final multivariate model. Semirobust standard errors were used in calculating the p values and confidence intervals.

Univariate GEE analyses revealed that postgraduate year, total number of shoulder arthroscopies performed, total number of arthroscopies performed on any joint, and total number of surgical cases performed during residency training prior to testing were associated with the mean time required to complete the simulator task (Table III). Postgraduate year had the greatest impact on the time to complete the simulator task, followed by the total number of shoulder arthroscopies performed, total number of all arthroscopies performed, and total number of surgical cases performed. On average, for every one-year increase in postgraduate year, there was a twenty-three-second decrease in the time required to complete the simulator task ($p < 0.001$). Similarly, for every shoulder arthroscopy case performed as a resident during training, there was an average 0.6-second decrease in the time required to complete the simulator task ($p < 0.001$). The total number of arthroscopies performed and the total number of surgical cases completed during residency prior to completing the simulator task were also correlated with shorter times to complete the task ($p < 0.001$), but had a smaller effect.

Multivariate analysis revealed that the model that best fit the data included postgraduate year and the total number of shoulder arthroscopies performed during residency prior to completing the simulator task (Table III). On average, for every one-year increase in postgraduate year, there was a sixteen-second decrease in the time required to complete the simulator task ($p = 0.005$) when the number of shoulder arthroscopies performed during residency prior to testing was held constant. Similarly, after controlling for the influence of postgraduate year, there was a twelve-second decrease in the time to complete the similar task with every additional fifty shoulder arthroscopies performed during residency training ($p = 0.008$).

Discussion

The purpose of the current study was to evaluate the association of the surgical training and experience of orthopaedic residents with the performance of basic arthroscopic skills in a shoulder simulation model. Although some previous studies have provided evidence for the validity of individual simulator models on the basis of a subjective assessment of clinical experience[7-9,18-20], none of these studies has correlated simulation performance with an objective assessment of surgical experience. To the best of our knowledge, the present study is the first to use resident ACGME case logs along with other objective measures to assess arthroscopic simulator performance. Utilizing the ACGME case log allowed us to account for multiple variables, including the total number of shoulder arthroscopic procedures, total number of all arthroscopic procedures, total number of cases, and postgraduate year in training. To our knowledge, this study is also the first to prospectively follow its subjects over a three-year period, enabling the creation of a model with predictive validity. While controlling covariates (total number of shoulder arthroscopic procedures, total number of all arthroscopic procedures, total cases, and postgraduate year), we were able to establish that postgraduate year and the total number of shoulder arthroscopies performed were the most important factors predicting performance of basic arthroscopic skills on a shoulder simulator model.

Prior studies have compared subjects' simulator performance with experience, but consistency among these studies with regard to classification of experience and analysis of performance was poor[21,22]. In some previous studies, residents were categorized only according to postgraduate year, regardless of their individual surgical, and specifically arthroscopic, experience[6,9]. Variation in orthopaedic arthroscopic surgical training among programs and among residents is inevitable, making postgraduate year at best a poor generalization. Self-reporting of surgical experience[6,7,9] has also been utilized. Such methodology, however, is fraught with potential bias, variability, and inaccuracies. With no uniform measurement of surgical experience, arthroscopic simulation research has been unable to define predictive metrics capable of informing prospective resident evaluation or training. In the current study, we were able to clearly define surgical experience on the basis of the number of shoulder arthroscopies performed, and we subsequently followed the subjects, in some cases over multiple years, to establish a clear model with predictive validity based entirely on objective data.

It is not surprising that incremental experience in clinical arthroscopy correlated with improved performance on the simulator. Our finding that postgraduate year correlated with performance (which, to our knowledge, has not been reported previously) is also not surprising. It is also interesting to note that our postgraduate year-4 students performed better on the simulator compared with the year-5 students. This may be partially explained by the proximity of testing to the sports medicine rotation in postgraduate year 4, as manifested by the increased number of arthroscopic cases performed during that year.

The limitations of our study include the relatively small sample of subjects, all from a single program. The use of ACGME case logs is both a strength (since they clearly and objectively define surgical experience) and a weakness (since they are self-reported, without proven accuracy). This study was also performed prospectively over a three-year period during which the same simulator was used for all testing and one single test examiner, blinded to surgical experience, conducted all testing. Simulation confounders were also minimized by not allowing the subjects to have any additional simulation experience during the study, a factor that has not been accounted for in other research[18].

In conclusion, the present study demonstrated that performance of basic arthroscopic skills on a shoulder simulator model was independently associated with postgraduate year and with the total number of shoulder arthroscopies performed prior to testing. These findings warrant further investigation to determine if training on a validated arthroscopic shoulder simulator would improve basic shoulder arthroscopic surgical skills in the clinical setting. ■

Kevin D. Martin, DO
Philip J. Belmont Jr., MD
Andrew Schoenfeld, MD
William Beaumont Army Medical Center,
5005 North Piedras Street, El Paso, TX 79920

Kenneth Cameron, PhD, MPH, ATC
Brett D. Owens, MD
Keller Army Hospital, 900 Washington Road,
West Point, NY 10996

References

1. Farnworth LR, Lemay DE, Wooldridge T, Mabrey JD, Blaschak MJ, DeCoster TA, Wascher DC, Schenck RC Jr. 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 Orthop J. 2001;21:31-5.

2. Bridges M, Diamond DL. The financial impact of teaching surgical residents in the operating room. Am J Surg. 1999 Jan;177(1):28-32.

3. Ceponis PJ, Chan D, Boorman RS, Hutchison C, Mohtadi NG. A randomized pilot validation of educational measures in teaching shoulder arthroscopy to surgical residents. Can J Surg. 2007 Oct;50(5):387-93.

4. Howells NR, Gill HS, Carr AJ, Price AJ, Rees JL. Transferring simulated arthroscopic skills to the operating theatre: a randomised blinded study. J Bone Joint Surg Br. 2008 Apr;90(4):494-9.

5. Poss R, Mabrey JD, Gillogly SD, Kasser JR, Sweeney HJ, Zarins B, Garrett WE Jr, Cannon WD. Development of a virtual reality arthroscopic knee simulator. J Bone Joint Surg Am. 2000 Oct;82-A(10):1495-9.

6. 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 Nov 2;93(21):e1271-5.

7. 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 Jun;35(6):883-8. Epub 2007 Jan 29.

8. Pedowitz RA, Esch J, Snyder S. Evaluation of a virtual reality simulator for arthroscopy skills development. Arthroscopy. 2002 Jul-Aug;18(6):E29.

9. Srivastava S, Youngblood PL, Rawn C, Hariri S, Heinrichs WL, Ladd AL. Initial evaluation of a shoulder arthroscopy simulator: establishing construct validity. J Shoulder Elbow Surg. 2004 Mar-Apr;13(2):196-205.

10. 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 Jan;442:21-9.

11. McCarthy A, Harley P, Smallwood R. Virtual arthroscopy training: do the ''virtual skills'' developed match the real skills required? Stud Health Technol Inform. 1999;62:221-7.

12. Accreditation Council for Graduate Medical Education. Duty hours: ACGME standards. https://www.acgme.org/acWebsite/dutyHours/dh_index.asp. Accessed 1 Nov 2011.

13. Kairys JC, McGuire K, Crawford AG, Yeo CJ. Cumulative operative experience is decreasing during general surgery residency: a worrisome trend for surgical trainees? J Am Coll Surg. 2008 May;206(5):804-11; discussion 811–3.

14. Garrett WE Jr, Swiontkowski MF, Weinstein JN, Callaghan J, Rosier RN, Berry DJ, Harrast J, Derosa GP. American Board of Orthopaedic Surgery Practice of the Orthopaedic Surgeon: Part-II, certification examination case mix. J Bone Joint Surg Am. 2006 Mar;88(3):660-7.

15. Satava RM. Virtual reality surgical simulator. The first steps. Surg Endosc. 1993 May-Jun;7(3):203-5.

16. Cui J. QIC program and model selection in GEE analyses. Stata J. 2007;7(2):209-220.

17. Pan W. Akaike's information criterion in generalized estimating equations. Biometrics. 2001 Mar;57(1):120-5.

18. 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 Jun;35(6):883-8. Epub 2007 Jan 29.

19. Insel A, Carofino B, Leger R, Arciero R, Mazzocca AD. The development of an objective model to assess arthroscopic performance. J Bone Joint Surg Am. 2009 Sep;91(9):2287-95.

20. Michelson JD. Simulation in orthopaedic education: an overview of theory and practice. J Bone Joint Surg Am. 2006 Jun;88(6):1405-11.

21. Tashiro Y, Miura H, Nakanishi Y, Okazaki K, Iwamoto Y. Evaluation of skills in arthroscopic training based on trajectory and force data. Clin Orthop Relat Res. 2009 Feb;467(2):546-52. Epub 2008 Sep 13.

22. Sherman KP, Ward JW, Wills DP, Sherman VJ, Mohsen AM. Surgical trainee assessment using a VE knee arthroscopy training system (VE-KATS): experimental results. Stud Health Technol Inform. 2001;81:465-70.