Simulation Training Improves Surgical Proficiency and Safety During Diagnostic Shoulder Arthroscopy Performed by Residents

Brian R. Waterman, MD; Kevin D. Martin, DO; Kenneth L. Cameron, PhD, MPH, ATC; Brett D. Owens, MD; Philip J. Belmont Jr, MD

abstract

Although virtual reality simulators have established construct validity, no studies have proven transfer of skills from a simulator to improved in vivo surgical skill. The current authors hypothesized that simulation training would improve residents' basic arthroscopic performance and safety. Twenty-two orthopedic surgery trainees were randomized into simulation or standard practice groups. At baseline testing, all of the participants performed simulator-based testing and a supervised, in vivo diagnostic shoulder arthroscopy with video recording. The simulation group subsequently received 1 hour of total instruction during a 3-month period, and the standard practice group received no simulator training. After intervention, both groups were reevaluated with simulator testing and a second recorded diagnostic shoulder arthroscopy. Two blinded, independent experts evaluated arthroscopic performance using the anatomic checklist, Arthroscopic Surgery Skill Evaluation Tool (ASSET) score, and total elapsed time. All outcome measures were compared within and between groups. After intervention, mean time required by the simulation group to complete the simulator task (30.64 seconds) was 8±1.2 seconds faster than the time required by the control group (38.64 seconds; *P*=.001). Probe distance (51.65 mm) was improved by 41.2±6.08 mm compared with the control (92.83 mm; *P*=.001). When comparing ASSET safety scores, the simulation group was competent (3.29) and significantly better than the control group (3.00; *P*=.005) during final arthroscopic testing. This study establishes transfer validity for an arthroscopic shoulder simulator model. Simulator training for residents in training can decrease surgical times, improve basic surgical skills, and confer greater patient safety during shoulder arthroscopy. [*Orthopedics*.]

Graduate medical education in orthopedic surgery and other surgical disciplines has evolved rapidly in the past decade. With the implementation of major changes mandated by the Accreditation Council for Graduate Medical Education (ACGME), residents must now comply with significant work hour restrictions, case log requirements, and formal surgical skills training programs.[1,2] These changes have prompted residency program directors and orthopedic educators to critically appraise and further develop their current surgical skills training curriculum. Ide-

Figure: Photograph of the shoulder simulator model (Arthro VR, 3D Systems, Simbionix Products, Cleveland, Ohio).

ally, these programs should promote and accelerate skill acquisition within a safe and controlled simulation environment to ultimately improve surgical performance in the operating room. Although the link between technical ability on simulated models and actual surgical proficiency has been more firmly established in the general surgical literature,[3-6] many of these studies have not established the construct validity of the individual training model with an objective assessment of surgical experience, such as ACGME surgical case log data, or used previously validated assessment tools to quantify arthroscopic skill progression.

When incorporating the use of a simulator model as part of an orthopedic surgical skills curriculum, it is pragmatic to objectively assess its value prior to full integration. Knee and shoulder arthroscopy are among the most commonly performed surgeries in the United States,[7] and the ACGME Orthopaedic Surgery Residency Review Committee has selected shoulder

arthroscopy as 1 of the 15 procedures with minimum caseload requirements.[2] As the number and complexity of shoulder arthroscopic procedures has increased with a concomitant decrease in resident work hours, arthroscopic shoulder simulation may better facilitate the development of basic shoulder arthroscopy skills during residency training. In a previous study, the performance of basic arthroscopic exercises on a shoulder simulator model has shown a strong correlation with similar tasks in a cadaveric model.[8] In addition, technical performance on the same shoulder simulator model had a significant relationship with surgical experience as measured by total number of arthroscopic cases and year in training, thus establishing the construct validity of the arthroscopic shoulder simulator under study.[9] Based on these findings, the aim of the current study was to evaluate the transfer validity of this arthroscopic shoulder simulator model to actual shoulder arthroscopy.

To this end, the current investigation evaluated orthopedic surgery residents' performance before and after arthroscopic shoulder simulation training in comparison to a control group to determine whether the simulation training improved basic arthroscopic shoulder task performance in the operating room. The current authors hypothesized that orthopedic residents would show improved performance in basic shoulder arthroscopic skills following a supplementary arthroscopic simulation training curriculum.

MATERIALS AND METHODS
Participants

After receiving institutional review board approval, all orthopedic trainees (N=22) from within a single program with varying postgraduate year training were enrolled for this study between October 2012 and January 2013. All participants were recruited voluntarily to participate, and academic standing was not affected by study involvement or individual performance. Previous arthroscopic experi-

ence varied by the individual resident's postgraduate year in training and clinical rotation schedule. All residents had participated in diagnostic simulation testing, but no formal arthroscopic simulation training had been undertaken previously. Demographic information and arthroscopy experience were collected at the onset of testing and included age, sex, postgraduate year in training, and ACGME surgical case log data. Each resident's case log previously was reviewed quarterly by the senior author (P.J.B.) to ensure accuracy of total arthroscopic surgical cases.

Study Design

The study was a single-blinded, prospective randomized control trial using a parallel group design in which participants were randomized to either a simulation curriculum or a standard practice curriculum after stratification by year in training. Following randomization, both groups continued all rotation-specific clinical and surgical duties, which were uncontrolled but accounted for using ACGME case log data. Initially, both groups were briefed on the study design and shown a sample video recording of a 14-point diagnostic shoulder arthroscopy performed by an expert surgeon. In addition, all of the participants viewed a slide show presentation of designated anatomic structures and correlative arthroscopic images for later testing purposes.

Simulator Testing

All participants underwent standardized evaluation on the Arthro VR shoulder simulator (Simbionix Products, Cleveland, Ohio) (**Figure**). This simulator is arranged with 2 robotic arms that are equipped with force reflective technology to provide haptic feedback to the participant and produce a high-fidelity arthroscopic model.[10] The simulator allows participants to manipulate tissue and perform basic arthroscopic tasks using high-definition monitors and state-of-the-art simulation technology.

For the purposes of this study, pre- and postintervention simulator assessments were performed using a validated blue-sphere program,[8] which places blue spheres at anatomic locations within the joint that the participant must locate and palpate. The simulator objectively recorded task performance outcomes that included time to completion (in seconds), camera distance (in mm), and probe distance (in mm). In a previous validation study, a shorter time to completion, camera distance, and probe distance were associated with participants' expert performance, indicating a higher level of proficiency.[8]

Arthroscopic Testing

After orientation, both simulation and standard practice curriculum participants performed a baseline diagnostic shoulder arthroscopy in the operating room. All surgeries were performed in the beach chair position under staff supervision using a hooked probe, 30° arthroscope, and high-definition arthroscopic camera (Stryker Inc, Kalamazoo, Michigan) with video recording. Standard anterior working and posterior viewing portals were established by an unblinded, attending orthopedic surgeon, and participants performed standard 14-point diagnostic arthroscopy until completion (**Table 1**). All surgeries were performed without coaching or assistance from the attending surgeon, and testing could be terminated early if deemed unsafe.

All video recordings were viewed by 2 blinded, independent arthroscopy experts, which is a practice that has been proven to allow multiple raters to reliably and objectively assess surgical performance.[11] The shoulder arthroscopy performances were graded according to the Arthroscopic Surgery Skill Evaluation Tool (ASSET).[12] The ASSET global rating scale has 8 separate skill domains and has been shown to have acceptable validity as well as interobserver and test-retest reliability for assessing surgical

performance during diagnostic arthroscopy.[12] The primary diagnostic shoulder arthroscopy outcome measures included composite ASSET and individual domain scores, performance on the 14-point diagnostic arthroscopy anatomic checklist, and time to completion.

Simulator Training

All participants assigned to the simulation curriculum received 4 supplementary, one-on-one simulation training sessions lasting approximately 15 minutes with 1 senior resident during a 3-month period. Instruction emphasized camera orientation, probe manipulation, triangulation, and further psychomotor skill development. One hour of cumulative training was chosen based on a previous study demonstrating that residents improved 12 seconds on the simulator per 50 arthroscopic cases completed.[9] Thus, the current authors estimated that participating residents could perform approximately 12 to 15 diagnostic examinations in each simulator session to reach a critical threshold for performance evaluation. Training was concluded after a total of at least 60 minutes of supervised simulator activity.

Final Simulator and Arthroscopic Testing

After the simulation training intervention, both groups performed a second simulation examination using the blue-sphere program. In addition, a second diagnostic shoulder arthroscopy was performed, and a postintervention blinded ASSET score was generated from the video recording. Simulation testing, ASSET global score and individual domain scores, 14-point diagnostic arthroscopy anatomic checklist, and time to completion were compared within and between simulation and standard practice groups before and after the intervention.

Statistical Analysis

Mathematical means and standard deviations were calculated for continu-

Table 1
14-Point Anatomic Checklist[a]

No.	Task
1	Inspect and probe long head of the biceps tendon
2	Pull biceps from the intertubercular groove into the joint
3	Inspect the biceps sling
4	Inspect and probe the superior labrum
5	Inspect and probe the glenoid articular surface
6	Inspect and probe the humeral head articular surface
7	Inspect the articular surface of the supraspinatus muscle
8	Inspect the articular surface of the infraspinatus muscle
9	Inspect the posterior humeral head and bare area
10	Inspect the capsular attachment to the humerus (HAGL)
11	Inspect the inferior pouch of the humerus
12	Inspect and probe the posterior labrum
13	Inspect and probe the anterior labrum
14	Inspect the subscapularis recess and insertion

Abbreviation: HAGL, humeral avulsion glenohumeral ligament.
[a]Participants were video recorded completing the 14-point anatomic checklist (unassisted) arthroscopically, and the recording then was graded by 2 independent blinded experts.

ous variables, and frequencies and proportions were calculated for categorical variables by group (simulation vs standard practice curriculum) and time points (pre- vs postintervention). For continuous variables, independent *t* tests were used to evaluate between-group differences at baseline and at follow-up for all outcomes of interest. If the assumptions for the independent *t* test were not met, the nonparametric equivalent Kruskal-Wallis test was used for inter-group com-

Table 2

Demographic Data for the Simulation and Standard Practice Groups

Parameter	Standard Practice Group		Simulation Group		P
	Median	IQR	Median	IQR	
Age, y	33	4.0	32	2.5	
Sex, male:female	10:0	-	11:1	-	
Postgraduate year	3.0	2.0	3.0	2.5	.847
Arthroscopies performed[a]					
Baseline	38.5	99.0	33.5	36.0	.484
Follow-up	39.5	98.0	43.0	41.0	.816

Abbreviation: IQR, interquartile range.
[a]*Number of reported shoulder arthroscopy current procedural terminology (CPT) codes.*

parisons. Dependent *t* tests were used to evaluate within group differences from pre- to postintervention for continuous variables. If the assumptions for the dependent *t* test were not met, the nonparametric equivalent Wilcoxon rank sum test was used for within-group comparisons over time. All statistical analyses were completed using STATA/SE software version 10.1 (StataCorp, College Station, Texas), and a type I error rate of $P<.05$ was used to assess statistical significance.

RESULTS

Of the 22 residents who agreed to participate in the study, 12 were randomized to the simulation intervention group and 10 were randomized to the control or standard practice group. Both groups were similar in terms of postgraduate year in training, number of shoulder arthroscopy cases per-

formed at baseline testing (ie, preintervention), and number of shoulder arthroscopy cases performed at the time of postintervention testing (**Table 2**). There were no preintervention differences between the simulation and standard practice groups on any of the arthroscopic shoulder simulator task performance measures (**Table 3**).

Following the intervention, participants in the simulation group had significantly better ASSET safety scores ($P=.005$) during the diagnostic arthroscopy and performed better on the ASSET ($P=.061$) compared with participants in the standard practice group (**Table 3**). Participants in the simulation group also completed the simulation testing significantly faster ($P=.001$) and were much more efficient in using the probe ($P=.001$) during the simulation test following the intervention compared with the participants in the standard practice group. Participants in the simulation group also were more efficient in using the camera following the intervention compared with participants in the standard practice group; however, the difference between the 2 groups was not statistically significant ($P=.07$).

Table 3

Between-Group Comparisons for the ASSET Score and Simulation Variables Pre- and Postintervention

Variable	Pre					Post				
	Standard Practice Group		Simulation Group		P	Standard Practice Group		Simulation Group		P
	Mean	SD	Mean	SD		Mean	SD	Mean	SD	
Diagnostic examination										
ASSET total score	19.10	1.46	19.33	1.16	.674	21.25	1.47	22.50	1.46	.061
Safety score[a]	2.95	0.21	2.92	0.17	.721	3.00	0.19	3.29	0.24	.005
Anatomic checklist[b]	11.00	1.65	10.71	3.67	.711	11.35	2.32	11.17	3.21	.828
Time, sec	297.30	153.62	294.67	127.80	.966	232.0	106.03	205.92	105.23	.571
Simulation examination										
Time, sec	52.11	7.61	47.24	7.07	.140	38.64	4.21	30.64	1.12	.001
Probe,[c] mm	133.78	30.60	136.26	34.07	.859	92.83	23.61	51.65	6.08	.001
Camera,[c] mm	279.19	62.17	289.34	44.55	.672	207.8	24.04	191.64	9.19	.070

Abbreviation: ASSET, Arthroscopic Surgery Skill Evaluation Tool.
[a]*Safety score: an objective evaluation within the ASSET score.*
[b]*Anatomic checklist: number of anatomic locations identified during diagnostic examination.*
[c]*Simulation probe and camera indicate the distance the instrument traveled to complete a diagnostic examination.*

Table 4

Pre- to Postintervention Within-Group Comparisons for the ASSET Score and Simulation Variables

| | Standard Practice Group | | | | | Simulation Group | | | | |
| | Pre | | Post | | | Pre | | Post | | |
Variable	Mean	SD	Mean	SD	P	Mean	SD	Mean	SD	P
Diagnostic examination										
ASSET total score	19.10	1.46	21.25	1.47	.005	19.33	1.16	22.50	1.46	.002
Safety score[a]	2.95	0.21	3.00	0.19	.748	2.92	0.17	3.29	0.24	.119
Anatomic checklist[b]	11.00	1.65	11.35	2.32	.505	10.71	3.37	11.17	3.21	.548
Time, sec	297.30	153.62	232.00	106.03	.069	294.67	127.80	205.92	105.23	.026
Simulation examination										
Time to completion, sec	52.11	7.61	38.64	4.21	.082	47.24	7.07	30.64	1.12	.030
Probe distance[c], mm	133.78	30.60	92.83	23.61	.059	136.26	34.07	51.65	6.08	.040
Camera distance[c], mm	279.19	62.17	207.79	24.04	.180	289.34	44.55	191.64	9.19	.049

Abbreviation: ASSET, Arthroscopic Surgery Skill Evaluation Tool.
[a]Safety score: an objective evaluation within the ASSET score.
[b]Anatomic checklist: number of anatomic locations identified during diagnostic examination.
[c]Simulation probe and camera indicate the distance the instrument traveled to complete a diagnostic examination.

Both groups demonstrated significant improvements on the ASSET during the second diagnostic arthroscopy compared with the preintervention baseline assessments (**Table 4**). In addition, participants in the simulation group also performed the second diagnostic arthroscopy significantly faster following the intervention (P=.026). Finally, participants in the simulation group demonstrated significant improvements from pre- to postintervention for the time required to complete the simulator assessment (P=.03) and both probe (P=.04) and camera (P=.049) efficiency during the task, whereas none of the improvements observed in the standard practice group over time were statistically significant for the simulator task assessment (P>.05).

When examining performance on the 14-point diagnostic arthroscopy anatomic checklist in both the simulation and standard practice pre- and postintervention as well as between the 2 groups, no significant differences were found. On average, the residents completed each point of the 14-point diagnostic arthroscopic shoulder examination more than 70% of the time,

except when evaluating the subscapularis recess and insertion (56%) and the capsuloligamentous attachments on the humerus (27%).

DISCUSSION

The current study revealed that additional simulation training improved technical performance and patient safety measures among orthopedic surgery residents during shoulder arthroscopy compared with a standard teaching curriculum. Similarly, a strong correlation between improved simulation measures and greater arthroscopic surgical proficiency was demonstrated. After high-fidelity simulation training intervention, the simulation group was able to perform designated arthroscopic tasks faster and more efficiently than the standard practice group, although a comparatively higher overall ASSET score among the simulation cohort failed to achieve statistical significance. This study is the first investigation to establish transfer validity, or the translation of technical skill acquired from simulation training to improved performance during real-time arthroscopic surgery, in a shoulder simulator model.

Traditional surgical training curricula have long ascribed to an apprenticeship model, in which residents assume increasing levels of hands-on involvement during surgery with subjective feedback about surgical competency. However, with strict resident work hour restrictions and possibly inadequate arthroscopic training,[13] orthopedic surgical training programs have sought alternative training opportunities outside the operating room to respond to the increasing scope and complexity of arthroscopic surgery. Similarly, the implications of graduate medical training on patient safety and financial health care costs, particularly those related to increased operative time, also have been considered in developing simulation training programs.[14,15]

Despite the impetus for resident simulation-based learning, early concerns about the increased cost and validity of arthroscopic simulator training have persisted.[16] Construct validity, or the ability for a simulator to reflect relative arthroscopic experience, previously has been an area of focus in shoulder simulator models, including investigations at the current

authors' institution.[8,9,17,18] Gomoll et al[17] confirmed psychomotor skill assessment using a virtual reality shoulder simulator correlated with actual arthroscopic surgical experience among 43 participants, with the most notable differences evident between the least and most experienced groups in terms of time to completion, path length, hook collisions, and average probe velocity. Martin et al[9] found postgraduate year in training and total number of arthroscopic shoulder cases independently predicted total time to completion for simple arthroscopic tasks on a shoulder simulator, with a decrease of 12 seconds for every additional 50 cases performed during residency training.

Conversely, transfer validity has not been sufficiently investigated and is warranted to fully establish shoulder simulators as a useful adjunct with tangible results in arthroscopic surgical proficiency. In a knee model, Howells et al[19] reported junior orthopedic trainees undergoing a 1-week arthroscopic simulation program performed better on the Orthopaedic Competence Assessment Project score and on the global rating scale during actual knee arthroscopy than those with traditional training. To date, however, no study has confirmed transfer validity for an arthroscopic shoulder simulator model. Although both groups in the current study demonstrated significantly better overall ASSET scores on the second diagnostic arthroscopy compared with baseline arthroscopic evaluation, the simulation group had a superior performance on the ASSET safety domain, a measure of iatrogenic articular cartilage or soft-tissue injury, relative to the standard practice group (P=.005). In addition, the simulation group demonstrated a higher mean overall ASSET score than the standard practice group on final testing (P=.061). Ostensibly, these findings could be attributed to the statistically significant improvements on simulator-based outcome measures present only in the intervention group.

Within this study, certain limitations must be acknowledged. Only 22 orthopedic trainees were available for randomization, which represents all trainees from within the authors' single program. With only 22 participants, the postintervention between group comparisons may have been underpowered; however, a large effect size was observed between groups for the training intervention in total ASSET score (Cohen's d=0.853) and camera distance (Cohen's d=0.888), both of which approached significance (**Table 3**). Post hoc power estimates suggest 22 participants per group would have been needed to detect a significant between-group difference in postintervention ASSET score and 20 participants per group would have been needed to detect a significant between-group difference in postintervention camera distance. All of the participants in the simulation training group were given only 1 hour of arthroscopic shoulder simulation training according to the current protocol, but less experienced surgeons could have benefitted from further exposure. In addition, more prolonged arthroscopic simulation training could have accentuated existing findings in the absence of larger enrollment. Finally, surgical indications and concomitant shoulder pathology were not standardized, which may alter normal arthroscopic anatomy and otherwise affect individual testing performance.

CONCLUSION

A modified surgical curriculum with arthroscopic simulation training may accelerate the learning curve for performing basic tasks during shoulder arthroscopy. More importantly, adjunctive simulation training synergistically promotes patient safety vis-à-vis traditional methods, although both through improved psychomotor coordination and shorter operative times. Further large-scale, well-designed studies are required to better elucidate the merits of simulation-based training in orthopedic surgery residency and its downstream effects on arthroscopic profi-

ciency and ultimately on clinical patient outcomes. As simulation technology becomes more sophisticated, dedicated arthroscopic simulation training will be imperative for initial surgical development during orthopedic surgical residency. However, the cost effectiveness of virtual reality, high-fidelity simulation models also must be considered relative to other more lower-priced alternatives such as synthetic joint models or cadaveric specimens.

REFERENCES

1. Accreditation Council for Graduate Medical Education. ACGME common program requirements. http://www.acgme.org/Portals/0/PFAssets/ProgramRequirements/CPRs_07012015.pdf. Accessed October 4, 2013.

2. Accreditation Council for Graduate Medical Education. Orthopaedic surgery minimum numbers. http://www.acgme.org/Portals/0/PFAssets/ProgramResources/260_ORS_Case_Log_Minimum_Numbers.pdf. Accessed October 4, 2013.

3. Grantcharov TP, Kristiansen VB, Bendix J, Bardram L, Rosenberg J, Funch-Jensen P. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Br J Surg.* 2004; 91(2):146-150.

4. Gurusamy K, Aggarwal R, Palanivelu L, Davidson BR. Systematic review of randomized controlled trials on the effectiveness of virtual training for laparoscopic surgery. *Br J Surg.* 2008; 95(9):1088-1097.

5. Palter VN, Grantcharov T, Harvey A, MacRae HM. Ex vivo technical skills training transfers to the operating room and enhances cognitive learning: a randomized controlled trial. *Ann Surg.* 2011; 253(5):886-889.

6. Zendejas B, Cook DA, Bingener J, et al. Simulation-based mastery learning improves patient outcomes in laparoscopic inguinal hernia repair: a randomized controlled trial. *Ann Surg.* 2011; 254(3):502-509.

7. Garrett WE Jr, Swiontkowski MF, Weinstein JN, et al. American Board of Orthopaedic Surgery practice of the orthopedic surgeon: Part-II. Certification examination case mix. *J Bone Joint Surg Am.* 2006; 88(3):660-667.

8. 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.

9. 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.

10. Satava RM. Virtual reality surgical simulator: the first steps. *Surg Endosc.* 1993; 7(3):203-205.

11. Dath D, Regehr G, Birch D, et al. Toward reliable operative assessment: the reliability and feasibility of videotaped assessment of laparoscopic technical skills. *Surg Endosc.* 2004; 18(12):1800-1804.

12. Koehler RJ, Amsdell S, Arendt EA, et al. The Arthroscopic Surgical Skill Evaluation Tool (ASSET). *Am J Sports Med.* 2013; 41(6):1229-1237.

13. Hall MP, Kaplan KM, Gorczynski CT, Zuckerman JD, Rosen JE. Assessment of arthros-copy training in US orthopaedic surgery residency programs: a resident self-assessment. *Bull NYU Hosp Jt Dis.* 2010; 68(1):5-10.

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

15. Farnsworth LR, Lemay DE, Wooldridge T, et al. 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; 21:31-35.

16. Atesok K, Mabrey JD, Jazrawi LM, Egol KA. Surgical simulation in orthopaedic skills training. *J Am Acad Orthop Surg.* 2012; 20(7):410-422.

17. 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.

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

19. 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; 90(4):494-499.