Demonstrating the Effectiveness of the Fundamentals of Robotic Surgery (FRS) Curriculum on the RobotiX Mentor Virtual Reality Simulation Platform

John R. Martin, MD[1], Dimitrios Stefanidis, MD, PhD[1], Ryan P. Dorin, MD[2], Alvin C. Goh[3], MD, Richard M. Satava, MD[4], Jeffrey S. Levy, MD[5]

[1]Department of Surgery, Indiana University School of Medicine, Indianapolis, IN
[2]Center for Education, Simulation and Innovation, Hartford Hospital, Hartford, CT
[3]Urology Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY
[4]Department of Surgery, University of Washington Medical Center, Seattle, WA
[5]Department of Obstetrics/Gynecology, Drexel University College of Medicine, Institute for Surgical Excellence, Philadelphia, PA

Corresponding author:

John Rhodes Martin, MD
Department of Surgery
Indiana University School of Medicine
545 Barnhill Drive, Emerson Hall 125
Indianapolis, IN 46202
p: 317-274-4966
f: 317-274-8769
e: jrm32@iu.edu

This is the author's manuscript of the article published in final edited form as:

Martin, J. R., Stefanidis, D., Dorin, R. P., Goh, A. C., Satava, R. M., & Levy, J. S. (2021). Demonstrating the effectiveness of the fundamentals of robotic surgery (FRS) curriculum on the RobotiX Mentor Virtual Reality Simulation Platform. Journal of Robotic Surgery, 15(2), 187–193. https://doi.org/10.1007/s11701-020-01085-4

Demonstrating the Effectiveness of the Fundamentals of Robotic Surgery (FRS) Curriculum on the RobotiX Mentor Virtual Reality Simulation Platform

Abstract word count: 250

Background 72

Fundamentals of Robotic Surgery (FRS) is a proficiency-based progression curriculum developed by robotic surgery experts from multiple specialty areas to address gaps in existing robotic surgery training curricula. The RobotiX Mentor is a virtual reality training platform for robotic surgery. Our aims were to determine if robotic surgery novices would demonstrate improved technical skills after completing FRS training on the RobotiX Mentor, and to compare the effectiveness of FRS across training platforms.

Methods 86

An observational, pre-post design, multi-institutional rater-blinded trial was conducted at two American College of Surgeons Accredited Education Institutes certified simulation centers. Robotic surgery novices (n=20) were enrolled and trained to expert-derived benchmarks using FRS on the RobotiX Mentor. Participants' baseline skill was assessed before (pre-test) and after (post-test) training on an avian tissue model. Tests were video recorded and graded by blinded raters using the Global Evaluative Assessment of Robotic Skills (GEARS) and a 32-criteria psychomotor checklist. Post-hoc comparisons were conducted against previously published comparator groups.

Results 53

On paired-samples T tests, participants demonstrated improved performance across all GEARS domains ($p<0.001$ to $p=0.01$) and for time ($p<0.001$) and errors ($p=0.003$) as measured by psychometric checklist. By ANOVA, improvement in novices' skill after FRS training on the RobotiX Mentor was not inferior to improvement reported after FRS training on previously published platforms.

Conclusion **39**

Completion of FRS on the RobotiX Mentor resulted in improved robotic surgery skills among novices, proving effectiveness of training. These data provide additional validity evidence for FRS and support use of the RobotiX Mentor for robotic surgery skill acquisition.

Keywords: robotic surgery, simulation, proficiency-based training, virtual reality, benchmarks

Demonstrating the Effectiveness of the Fundamentals of Robotic Surgery (FRS) Curriculum on the RobotiX Mentor Virtual Reality Simulation Platform

Manuscript word count: 2425

Background 330

The practice of robotic surgery has spread widely since the technology was first introduced into clinical practice in the mid-1990s. In order to train surgeons in the safe, effective use of robotic surgery, a variety of skills curricula have been developed by individual investigators and organizations, and mandatory participation in an industry-sponsored robotic training pathway is required before use of robotic surgery in clinical practice.[1-5] However, despite increasing implementation, robotic training curricula vary in terms of structure, content, and proficiency requirements. This is especially true for curricula aimed at surgical novices. Green, et al recently described the current state of robotic surgery training curricula across 12 general surgery residency training programs in the United States. They identified critical gaps in the content of existing curricula for technical skill acquisition and in the teaching of important intra-operative principles such as tissue handling.[6]

As an optimal way to address these gaps, proficiency-based progression has been demonstrated to be effective in expediting skill acquisition and ensuring skill durability.[2, 4, 5, 7-13] The Fundamentals of Robotic Surgery (FRS) curriculum is a proficiency-based progression curriculum that was developed by 66 robotic surgery experts from multiple specialty areas using a Full Life-cycle Curriculum Development process.[14, 15] This curriculum consists of didactic modules and a simulation-based skills curriculum for the acquisition of basic robotic skills. In a recent multi-institutional, multi-national randomized controlled trial, the effectiveness of this curriculum in improving trainee basic robotic skills was demonstrated in a physical training model and in two virtual reality (VR) training simulators.[15] Since this trial was conducted, the RobotiX Mentor, which is a stand-alone VR simulation platform developed by 3D Systems (formerly Simbionix) for robotic training, was introduced to the market. Our aims in this study were to determine if robotic surgery novices would demonstrate improved technical skill performance after completing FRS proficiency-based psychomotor skills training using the RobotiX Mentor VR platform; and, to compare the effectiveness of the RobotiX Mentor VR platform to other previously published FRS training platforms.

Methods 879

The study conducted was an observational, pre-post design, multi-institutional rater-blinded trial. The participating sites were two tertiary institutions with American College of Surgeons Accredited Education Institutes (ACS-AEI) certified simulation centers. Both institutions received local IRB approval prior to execution of the study protocol.

Participants recruited to the study were residents in Accreditation Council for Graduate Medical Education (ACGME) accredited urology, gynecology, or general surgery training programs. With respect to robotic surgery, they were novices according to Dreyfus' Five-Stage Model of Adult Skill acquisition.[16] At the beginning of training, participants who went on to complete the study protocol averaged only 0.6 hours of prior experience on any robotic surgery simulator (range = 0 to 4 hours) and had performed fewer than 0.5 robotic surgery cases as bedside assistant (range = 0 to 2 cases). No participant had ever performed a robotic surgery case as the primary surgeon. The mean age of participants who completed the protocol was 30.5 years (standard deviation = 3.4 years). One participant self-identified as female. One participant was left-handed (Table 1: "RobotiX Mentor").

Participants first reviewed the online didactic modules of the FRS curriculum and were required to pass the FRS cognitive exam.[17] Participants' baseline robotic surgery skills were then assessed (pre-test) on a previously-described avian tissue model[15] using the da Vinci Si surgical system (Intuitive Surgical Inc., Sunnyvale, CA). Briefly, the avian tissue model consists of 5 technical tasks that mirror the FRS training tasks and are germane to the performance of robotic surgery (knot tying, suturing, 4th arm cutting, puzzle piece dissection, and vessel dissection & energy application). There is previously-published validity evidence for use of this model as an assessment tool for the FRS curriculum.[15]

After baseline assessment, participants trained in the proficiency-based progression curriculum on the FRS tasks (docking, ring transfer, knot tying, suturing, 4th arm cutting, puzzle piece dissection, and vessel dissection & energy application) using the RobotiX Mentor VR platform (Figure 1). Expert proficiency levels (benchmarks) were established by having six experts perform all seven FRS tasks on the RobotiX Mentor VR platform until no performance improvement was observed on two consecutive attempts. Outlier performances were excluded, and the remaining expert performance scores were then averaged for each task. Following the proficiency-based progression model of the FRS curriculum, the training goal for participants was to perform each successive task equally well as the benchmark set by the expert surgeons on two consecutive attempts before proceeding to the next task. After successful completion of all FRS curriculum training tasks on the VR platform, participants were tested again on the avian tissue model (post-test).

Participant pre- and post-test performances were video recorded using local video recording devices and uploaded to Crowd-Sourced Assessment of Technical Skills (C-SATS, Seattle, WA) HIPAA-compliant cloud storage.[18] At the conclusion of the study period, all videos were reviewed by two expert raters provided by C-SATS. Raters were blinded to pre- and post-test status, participant, and institution. The two C-SATS provided raters used the Global Evaluative Assessment of Robotic Skills (GEARS) rating scale to assess participants' performance. GEARS is a tool for assessing robotic operative skill across five domains (depth perception, bimanual dexterity, efficiency, use of force, and instrument control); it uses a 5-point, anchored Likert scale for each domain, and it has previously-published content validity evidence for the avian tissue model used in the pre- and post-test.[19] Inter-rater reliability (IRR) was calculated between these two raters and final GEARS scores were calculated as the average of the two raters' scores. These scores lacked normal distribution, so non-parametric paired T tests were used to compare GEARS pre-test scores to GEARS post-test scores. Both cumulative performance scores and domain-specific performance scores were compared.

At the conclusion of the study period, pre- and post-test performance videos were also assessed by one local rater at each of the two participating institutions. Both raters were expert robotic surgeons. These local raters used a 32-criteria psychomotor checklist that was developed to specifically assess performance on the avian tissue model tasks. The psychomotor checklist measures time and errors, and it has been previously associated with high inter-rater reliability (IRR = 0.82–0.97) among different rater pairs.[15] Parametric paired T tests were used to compare psychomotor checklist pre- and post-test scores.

Additionally, one-way analysis of variance (ANOVA) was performed to compare psychometric checklist scores of the RobotiX Mentor participants with psychometric checklist scores of robotic surgery novices (Table 2) from another recently published pre-post study of the FRS curriculum by Satava, et al. The robotic surgery novices of the Satava, et al study had either participated in the Intuitive da Vinci curriculum with *no* FRS training ("control"); they had completed FRS training on a physical practice model, the FRS Physical Dome ("Dome", Florida Hospital Nicholson Center, Celebration, FL); or they had completed FRS training on one of two VR practice platforms—the da Vinci backpack VR simulator ("DVSS", 3D Systems, Tel Aviv, Israel) or the Mimic robotic surgery VR simulator ("dV-Trainer", Mimic Technologies, Inc., Seattle, WA). Of important note, because only residents were recruited to the RobotiX Mentor group, fellow and attending data from the Dome, DVSS, dV-Trainer, and control groups of the Satava, et al study were intentionally excluded from the comparative analysis so that only the performance of residents was compared across all groups.

Results **387**

Twenty residents who were robotic surgical novices were initially enrolled to participate in the RobotiX Mentor group, thirteen of whom (n = 13) completed the FRS curriculum and the avian tissue model post-test (35% attrition rate). Two participants' GEARS post-test data were incomplete, so these participants were excluded from the GEARS analysis. Another participant's pre-test time was a significant outlier—greater than five standard deviations beyond the mean—and therefore that participant was excluded from analysis of the psychomotor checklist data. On dependent-samples sign-test, all participants who completed the study protocol demonstrated improved performance from pre- to post-test as measured by GEARS scores. This was true both for overall cumulative score ($p<0.001$) and across specific GEARS domains ($p=0.01$ to $p<0.001$) as reported in Table 2. "Median of the differences" is interpreted as the median improvement in GEARS score from pre- to post-test. This pre- to post-test improvement is further visualized in the box & whisker chart of GEARS scores in Figure 2. There was a non-significant positive correlation ($r=0.58$, $p=0.06$) between pre- and post-test cumulative GEARS scores. Among specific GEARS domains, the only significant positive correlation between pre- and post-test scores was for efficiency ($r=0.71$, $p=0.01$). Between the two C-SATS provided raters, GEARS inter-rater reliability was 0.43 (0.8) for the pre-test scores and 0.41 (1.0) for the post-test scores.

On paired-samples T tests of psychomotor checklist scores, participants again demonstrated significant pre- to post-test improvement. This was true both for time to task completion ($t(11)=7.42$, $p<0.001$) and number of errors ($t(11)=3.80$, $p=0.003$). Number of errors on the pre-test was a significant predictor of number of errors on the post-test ($r=0.64$, $p=0.03$), while pre-test time was not a significant predictor of post-test time ($r=0.55$, $p=0.07$). By one-way ANOVA, there were no significant baseline differences in pre-test time ($F(4,49)=1.86$, $p=0.13$) or pre-test errors ($F(4,49)=1.78$, $p=0.15$) between the RobotiX Mentor group and the Dome (n=11), DVSS (n=9), dV-Trainer (n=12), and control groups (n=10) of the Satava, et al study. Similarly, no significant differences in performance improvement from pre- to post-test were observed between any of the groups from either study for either time ($F(4,49)=1.20$, p=0.32) or errors (F(4,49)=1.11), p=0.36). The pre- to post-test changes in time and errors for each group are illustrated in Figures 3 & 4; more negative values (reduction in time, reduction in errors) indicate greater improvement.

Discussion 829

Robotic surgery novices in our study demonstrated improved performance of robotic surgery technical skills after completing the FRS proficiency-based progression curriculum on the RobotiX Mentor VR training platform. The skills acquired during training on this VR simulator transferred to a realistic model (avian tissue model) that is relevant to use of the actual robot in clinical practice. These results are consistent with other studies which have shown trainee improvement in confidence level and in the performance of technical skills following participation in organized robotic surgery skills curricula.[20, 21] Additionally, every individual RobotiX Mentor participant demonstrated score improvement from pre- to post-test for time, errors, and across every unique GEARS domain following completion of the curriculum. Among GEARS items, observed improvements in dexterity, use of force, and instrument control potentially address Green, et al's identified gap in the teaching of tissue handling among currently existing curricula.[6] Importantly, the improvement in novices' robotic skill observed in the RobotiX Mentor VR platform group was not inferior to improvement observed among similar novices who also trained in the FRS curriculum using two other VR platforms (dV Trainer, DVSS) and who were assessed using the same avian tissue model and psychometric checklist.[15] This suggests that the quality of the curriculum may be more important for effective skill acquisition than the actual simulation platform used. Indeed, Satava, et al found only minor participant performance differences across all training platforms they evaluated.[15]

A limitation of this study is the lack of a true parallel control group against which to compare pre- to post-test score improvement. Although the control group whose data were borrowed to serve as a comparator participated in the Intuitive da Vinci curriculum *without* FRS training, they may have been influenced by FRS curriculum participants at their own institutions to pursue additional practice, thus leading to more-than-expected improvement as previously reported by Satava, et al.[15] In theory, the pre-test itself could have been enough of an inoculation to the avian tissue model to improve post-test performance, with or without participation in any training curriculum. However, the lack of significant, positive correlations between pre- and post-test time and GEARS scores does not support that the pre-test led to increased familiarity or comfort with the post-test. Moreover, it is unlikely to expect significant performance improvements in the absence of training.

The study also suffered from a high attrition rate, with 7 of the initial 20 RobotiX Mentor group participants failing to complete the protocol. The study administrators attribute this to trainees' prioritization of clinical obligations over robotic skills practice and to rotations across multiple training sites separate from the location of the simulators. This is consistent with a recently published paper by Tam, et al, in which 13 of 24 residents similarly started but failed to complete a voluntary robotic skills training curriculum (54% attrition rate). When surveyed, 10 of the residents from that study attributed their attrition to prioritization of clinical responsibilities and research activities, and to a lack of physical access to the simulator due to rotations at other facilities.[22] In ongoing studies conducted by members of our group, we have moved simulators from a central skills lab out to the hospitals where residents are rotating in hopes of removing at least this one barrier to participation.

In a post-hoc analysis of GEARS scores, participants in our study who completed the curriculum and post-test did not have significantly different pre-test scores compared with participants who only completed the pre-test but did not complete the rest of the training protocol ($W = 33.5$, $p = 0.68$). In other words, all initial participants had similar baseline performance, suggesting that attrition was not due to early poor performance. There were likewise no notable demographic predictors of performance among RobotiX Mentor participants, and despite the fact that every participant who completed the curriculum improved, we were only able to show non-inferiority to other platforms. Therefore, we must also consider the possibility that attrition led to a Type II error.

Finally, the IRR of the two C-SATS provided raters using the GEARS tool was moderate. The GEARS tool also suffered less than desirable IRR in the study by Satava, et al.[15] We can only speculate as to why this was the case. Other authors have reported high IRR using the GEARS tool after a "proficiency phase" to develop initial agreement among expert raters, as well as high Cronbach's alpha between expert raters and C-SATS crowd-sourced GEARS scores.[23] Perhaps IRR in this instance could have been improved if raters had undergone more rigorous rater training prior to conducting the assessments.

In conclusion, completion of the proficiency-based progression FRS curriculum on the RobotiX Mentor VR training platform resulted in improved performance of robotic surgery skills among robotic surgery novices with results similar to those on previously-reported VR platforms. These data provide further validity evidence for the FRS curriculum itself and support the use of the FRS curriculum on the RobotiX Mentor as an additional VR simulation platform for robotic skill acquisition.

Declarations

Sources of Funding

The Institute for Surgical Excellence (ISE), a 501(c)(3) public charity that supports surgical education, received a grant award and provided administrative support for conducting this study.

Conflicts of Interest

The authors declare that they have no relevant conflicts of interest.

Ethics Approval

The study was approved by local IRB and conducted at two American College of Surgeons Accredited Education Institutes (ACS-AEI) certified simulation centers.

Consent to Participate

Informed consent was obtained from all individual participants included in the study.

Authors' Contributions

All authors whose names appear on the submission made substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data; drafted the work or revised it critically for important intellectual content; approved the version to be published; and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

References

1. (2018) da Vinci Training. Intuitive Surgical, Inc.

2. Arain NA, Dulan G, Hogg DC, Rege RV, Powers CE, Tesfay ST, Hynan LS, Scott DJ (2012) Comprehensive proficiency-based inanimate training for robotic surgery: reliability, feasibility, and educational benefit. Surg Endosc 26:2740-2745

3. Connolly M, Seligman J, Kastenmeier A, Goldblatt M, Gould JC (2014) Validation of a virtual reality-based robotic surgical skills curriculum. Surg Endosc 28:1691-1694

4. Dulan G, Rege RV, Hogg DC, Gilberg-Fisher KM, Arain NA, Tesfay ST, Scott DJ (2012) Developing a comprehensive, proficiency-based training program for robotic surgery. Surgery 152:477-488

5. Dulan G, Rege RV, Hogg DC, Gilberg-Fisher KM, Arain NA, Tesfay ST, Scott DJ (2012) Proficiency-based training for robotic surgery: construct validity, workload, and expert levels for nine inanimate exercises. Surg Endosc 26:1516-1521

6. Green CA, Chern H, O'Sullivan PS (2018) Current robotic curricula for surgery residents: A need for additional cognitive and psychomotor focus. Am J Surg 215:277-281

7. Ahlberg G, Enochsson L, Gallagher AG, Hedman L, Hogman C, McClusky DA, 3rd, Ramel S, Smith CD, Arvidsson D (2007) Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. Am J Surg 193:797-804

8. Angelo RL, Ryu RK, Pedowitz RA, Beach W, Burns J, Dodds J, Field L, Getelman M, Hobgood R, McIntyre L, Gallagher AG (2015) A Proficiency-Based Progression Training Curriculum Coupled With a Model Simulator Results in the Acquisition of a Superior Arthroscopic Bankart Skill Set. Arthroscopy 31:1854-1871

9. Cates CU, Lönn L, Gallagher AG (2016) Prospective, randomised and blinded comparison of proficiency-based progression full-physics virtual reality simulator training versus invasive vascular experience for learning carotid artery angiography by very experienced operators. BMJ Simulation and Technology Enhanced Learning

10. Gallagher AG (2012) Metric-based simulation training to proficiency in medical education:- what it is and how to do it. Ulster Med J 81:107-113

11. Gallagher AG, O'Sullivan GC (2011) Fundamentals of Surgical Simulation: Principles & Practice, Springer, New York, NY

12. Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G, Smith CD, Satava RM (2005) Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. Ann Surg 241:364-372

13. Stefanidis D (2010) Optimal acquisition and assessment of proficiency on simulators in surgery. Surg Clin North Am 90:475-489

14. Zevin B, Levy JS, Satava RM, Grantcharov TP (2012) A consensus-based framework for design, validation, and implementation of simulation-based training curricula in surgery. J Am Coll Surg 215:580-586 e583

15. Satava RM, Stefanidis D, Levy JS, Smith R, Martin JR, Monfared S, Timsina LR, Darzi AW, Moglia A, Brand TC, Dorin RP, Dumon KR, Francone TD, Georgiou E, Goh AC, Marcet JE, Martino MA, Sudan R, Vale J, Gallagher AG (2019) Proving the Effectiveness of the Fundamentals of Robotic Surgery (FRS) Skills Curriculum: A Single-blinded, Multispecialty, Multi-institutional Randomized Control Trial. Ann Surg

16. Dreyfus SE (2004) The Five-Stage Model of Adult Skill Acquisition. Bulletin of Science, Technology & Society 24:177-181

17. (2018) Fundamentals of Robotic Surgery Curriculum. Case Network

18. (2018) C-SATS. C-SATS, Inc.

19. Goh AC, Goldfarb DW, Sander JC, Miles BJ, Dunkin BJ (2012) Global evaluative assessment of robotic skills: validation of a clinical assessment tool to measure robotic surgical skills. J Urol 187:247-252

20. Schommer E, Patel VR, Mouraviev V, Thomas C, Thiel DD (2017) Diffusion of Robotic Technology Into Urologic Practice has Led to Improved Resident Physician Robotic Skills. J Surg Educ 74:55-60

21. Bertolo R, Garisto J, Dagenais J, Sagalovich D, Kaouk JH (2018) Single Session of Robotic Human Cadaver Training: The Immediate Impact on Urology Residents in a Teaching Hospital. J Laparoendosc Adv Surg Tech A 28:1157-1162

22. Tam V, Lutfi W, Novak S, Hamad A, Lee KK, Zureikat AH, Zeh HJ, 3rd, Hogg ME (2018) Resident attitudes and compliance towards robotic surgical training. Am J Surg 215:282-287

23. White LW, Kowalewski TM, Dockter RL, Comstock B, Hannaford B, Lendvay TS (2015) Crowd-Sourced Assessment of Technical Skill: A Valid Method for Discriminating Basic Robotic Surgery Skills. J Endourol 29:1295-1301

Table 1. Participant Demographic Data.

	RobotiX Mentor	dV Trainer[1]	DVSS[1]	Physical Dome[1]	Control[1]
Self-identified gender (M:F)	12:1	6:5	8:5	9:4	9:2
Handedness (R:L)	12:1	9:2	13:0	12:1	10:1
Mean age (SD)	30.5 (3.4)	31.7 (3.1)	32.6 (9.0)	30.2 (3.1)	30.4 (2.6)
Mean simulator experience, hours (SD)	0.6 (1.2)	2.1 (3.6)	0.7 (1.2)	4.4 (8.9)	2.5 (4.3)
Prior robotic experience:					
Mean no. of cases as bedside assistant (SD)	0.5 (0.9)	19.0 (30.6)	10.5 (16.7)	12.1 (15.5)	7.2 (9.1)
Mean no. of cases as primary surgeon (SD)	0 (0)	0 (0)	0 (0)	1.2 (2.1)	0.1 (0.3)

[1]From Satava, et al, 2019. These demographic data include residents *only*. Attendings and fellows were excluded from our comparative analysis.

Table 2. Pre- to Post-Test Performance Improvement, by GEARS Domain.

GEARS Domain	Median of differences	95% CI	p-value
Depth perception	0.7	0.3–1.0	p<0.001
Dexterity	0.5	0.1–1.0	p<0.01
Efficiency	0.8	0.3–1.0	p<0.001
Use of force	0.8	0.4–1.0	p<0.01
Instrument control	0.6	0.2–1.0	p=0.01
Cumulative score	3.55	1.4–4.8	p<0.001