DOI:
10.1016/j.jsurg.2017.09.005

The validation of a novel Robot-Assisted Radical Prostatectomy virtual reality module

Authors

Patrick Harrison BSc, Nicholas Raison, MRCS,[†] Takashige Abe, PhD, MD,[‡] William Watkinson BSc,[*] Faizan Dar, MBBS,[†] Ben Challacombe, MS, FRCS(Urol),[§] Henk Van Der Poel M.D, PhD,[|] Muhammad Shamim Khan, FRCS(Urol),[†,§] Prokar Dasgupta, MD, FEBU, FRCS(Urol),[†,§] and Kamran Ahmed, FRCS (Urol), PhD[†,§]

*King’s College London, London, UK

[†] MRC Centre for Transplantation, Kings College London, London, UK

[‡]Hokkaido University, Sapporo, Hokkaido, Japan

[§]Department of Urology, Guy’s and St Thomas’ NHS Foundation Trust, London, UK

[|]Department of Urology, Netherlands Cancer Institute, The Netherlands

Key words: Virtual Reality, Simulation, Medical Education, Robotic surgery

Objective: To perform the first validation of a full procedural virtual reality (VR) robotic training module and analysis of novice surgeon's learning curves.

Design: Participants completed the bladder neck dissection task (BND) and urethrovesical anastomosis task (UVA) as part of the prostatectomy module. Surgeons completed feedback questionnaires assessing the realism, content, acceptability and feasibility of the module. Novice surgeons completed a 5.5-hour training programme using both tasks.

Setting: King's College London, London.

Participants: 13 novice, 24 intermediate and 8 expert surgeons completed the validation study.

Results: Realism was scored highly for BDN (mean 3.4/5) and UVA (3.74/5), as was importance of BDN (4.32/5) and UVA (4.6/5) for training. It was rated a feasible (3.95/5) and acceptable (4/5) tool for training. Experts performed significantly better than novice group in 6 metrics in the UVA including time (p=0.0005), distance by camera (p=0.0010) and instrument collisions (p=0.0033), as well as task-specific metrics such as number of unnecessary needle piercing points (p=0.0463). In novice surgeons, a significant improvement in performance after training was seen in many metrics for both tasks. For BND, this included time (p<0.0001), instrument collisions (p=0.0013) and total time instruments are out of view (p=0.0251). For UVA, this included time (p=0.0135), instrument collisions (p=0.0066) and task specific metrics such as injury to the urethra (p=0.0032) and bladder (p=0.0189).

Conclusions: Surgeons found this full procedural VR training module to be a realistic, feasible and acceptable component for a robotic surgical training programme. Construct validity was proven between expert and novice surgeons. Novice surgeons have shown a significant learning curve over 5.5 hours of training, suggesting this module could be used in a surgical curriculum for acquisition of technical skills. Further implementation of this module into the curriculum and continued analysis would be beneficial to gauge how it can be fully utilised.

ACGME competencies: Practice Based Learning and Improvement

Key words: Virtual Reality, Simulation, Medical Education, Robotic surgery

[1] Abbreviations

[1] Abbreviations: VR, Virtual reality; UVA, Urethrovesical anastomosis; BND, Bladder neck dissection.

Introduction

There are a number of virtual reality (VR) robotic simulators commercially available. Basic VR modules have been well validated for these simulators thus far. This current simulator, the RobotiX Mentor ™ (3D systems; Simbionix Products, Cleveland, OH, USA) is a robotic surgery virtual reality simulator that has been developed to train surgeons for robotic surgery performed using the da Vinci® Surgical System. The simulator platform consists of a height adjustable headset containing stereoscopic visors, free floating hand controls and adjustable foot pedals integrated into a single console. It has been proven by Whittaker et al. to be effective for training using the Fundamentals of Robotic Surgery (FRS) curriculum, a basic VR module.[1] With continued validation of robotic VR simulators, it is becoming increasingly recognised that VR simulation is integral to the surgical curriculum. [1,2] Now, developments in VR technology have enabled production of full procedural VR training modules. Procedural modules can replicate a real-life environment with increasingly accurate anatomy. They have the potential to be used to develop cognitive skills, team and non-technical skills and more advanced technical skills that may not offered in basic VR modules.

Before implementing procedural VR modules into robotic urological training curriculum, their usefulness and accuracy need to be established. The aim of this prospective study was to validate this novel full procedural "Robotic Radical Prostatectomy module" and to assess the feasibility and acceptability of the modules into a training curriculum.

Material and Methods

Study design

This was a prospective, observational and comparative study that was conducted at King's College London, further with data collected at the European Association of Urology (EAU) hands-on-training (HOT) courses.

Participants

Subjects were categorised into 3 groups (novice, intermediate, and expert). Opinion is divided on the number of procedures required to reach proficiency in robotic prostatectomy.[3,4] Experts were defined as having performed 50 cases or more independently. The intermediate group included subjects receiving surgical training who have performed up to 49 independent cases. The novice group was defined as having no previous operative experience.

Module

The Robotic Radical Prostatectomy module is comprised of 3 tasks representing key steps during a robotic-assisted radical prostatectomy (RARP):

1. Bladder neck dissection
2. Neurovascular Bundle Dissection (Nerve-Sparing)

3. Urethrovesical anastomosis

Successful completion of the neurovascular bundle dissection (nerve sparing) required advanced surgical and anatomical knowledge beyond that of the novice group. Therefore, it was excluded from the preliminary validation.

Process

The novice surgeons initially underwent basic robotic training based on 3 FRS tasks:

1. Ring Tower Transfer
2. Railroad Track
3. Vessel Energy Dissection

The training consisted of guiding the participant through the controls and teaching basic robotic skills. Intermediates and experts were offered the opportunity to use the familiarisation tasks prior to use of the procedural modules. No data was collected from the familiarisation tasks. All participants then performed the guided BND task followed by the guided UVA task. Post-completion of these tasks, experts and intermediates were asked to fill in a questionnaire assessing their experience, opinion on realism, importance, acceptability and feasibility of the modules and simulator. Novice surgeons went on to complete a mean 5.5-hour supervised training programme over 5 weeks that consisted of 1-hour time slots. During each training session, participants performed each task in no particular order. The prostatectomy module was the only module participants were permitted to use during training. The study process is illustrated by Figure 1.

Following completion of the tasks, experts and intermediates were asked to complete a questionnaire that assessed their perceived realism of the simulator in general like that of the da Vinci robot (including the controls and graphics etc.), and their view of the degree of realism of the bladder, urethra and prostate. A 5-point Likert scale was used to measure each included element, with 1 being “not similar”, and 5 being “very similar”.

Outcome measures and performance evaluation

After each task, the simulator produced a performance evaluation report, incorporating various generic and task-specific metrics. Between the two tasks there were 13 generic metrics in common: clutch usage, distance by camera, instrument collisions, number of movement of left instrument, number of movement of right instrument, number of times instruments are out of view, total time instruments are out of view, total path of instruments travelled, respect for tissue, path length of right instrument, path length of left instrument, suspected injury to the ureteral orifices and total time. In total, there were 24 metrics produced in the bladder neck task and 27 in the uerethrovesical anastomosis task.

Statistical analysis

Statistical analysis was conducted with JMP 13 (SAS Institute Inc., Cary, NC, USA). Differences between the groups were analysed using a Wilcoxon Each Pair test. Learning curve analysis was analysed using the Kruskal-Wallis test. Graphs were made with GraphPad Prism v5.03 (GraphPad Software, Inc., La Jolla, CA, USA).

Results

45 participants completed the validation study. Demographics of each group are shown in Table 1.

Both modules were scored highly in the post-training questionnaires. Mean scores for simulator controls, organs and overall experience can be seen in Figure 2. Realism of the BND task was scored mean 3.44/5 and UVA task 3.74/5. When asked to score the importance of individual tasks in training for RARP, both tasks were scored highly with bladder neck dissection 4.32/5 and urethrovesical anastomosis 4.6/5. These results, along with the feasibility and acceptability of the RobotiX Mentor simulator and the results can be seen in Figure 3.

65% of expert and intermediate surgeons agreed that full procedural training on the RobotiX Mentor could form an important component of surgeon accreditation or recertification.

Bladder neck dissection technical performance

Experts performed better than the novice group in: time to complete the task, distance by camera, instrument collisions, percentage of cuts that were performed with optimal catheter positioning during posterior dissection, number of movements and path length of the 4^{th} arm. However, these differences did not reach significance.

Urethrovesical anastomosis technical performance

Experts performed better than novices in 13 metrics, of which 6 were statistically significant. Experts performed better than intermediates, who performed better than novice in 7 metrics. Of these 7 metrics, 2 were statistically significant between all groups: time taken to complete task and distance travelled by camera. There was a significant difference between experts and novices in several other metrics displayed in Table 2.

Learning curves

The novice group completed an average of 5.5 hours of training using both procedural tasks. In the bladder neck task, a significance improvement in performance was seen in 11 metrics including time ($p<0.0001$), instrument collisions ($p=0.0013$), total time instruments were out of view ($p=0.0251$), path length of left ($p=0.0038$), number of movements with the left ($p=0.0003$), right ($p=0.0002$) and 4^{th} arm ($p=0.0008$) instruments. Path length for the right instrument was not significant ($p=0.1170$). Similar improvements over the course of training were seen for the UVA task for 13 metrics in total, of which some are illustrated in Figure 4.

Discussion

We have performed the first validation of a full procedural VR training module for robotic surgery. The training modules were rated highly by surgeons for their anatomical accuracy and realism as well as their utility as training tools. The realism of the simulator controls scored highly, an observation that confirms the results published by Whittaker et al. [1] Surgeons were impressed with the realism of the graphics and when asked to rate the realism of the organs (bladder, urethra and prostate) they did so highly. We have proven construct validity for both BND and UVA tasks. For the UVA task, a significant difference was seen between expert and novice in 6 performance metrics, including both generic and task specific performance metrics. Further to this, a statistical improvement was seen among many performance metrics with training. Although the BND task was only able to show a significant difference between the experts and novice surgeons in number of movements of the 4th arm (p=0.0393), novice surgeons showed a statistical improvement over 5.5 hours of training.

In several metrics, differences between groups of surgeons was not shown or was insignificant. For the BND task, the use of the guided modules may have been the reason for this. Some intermediates and experts questioned the computer's technique for the task and were hesitant to perform the task as per the instructions provided by the simulator. The novice group, with no operative experience, followed the instructions readily. This difference in technique may be a reason for some metrics not reaching significance in the BND task.

In the UVA task, experts grasped the tissue more than novice, but this did not reach significance (p=0.0833). However, the difference between intermediates and novices did (p=0.0173) despite the mean score for experts being higher. This may be due to the lower number of expert participants included in the study. Experts triggered the metric "suspected injury to the bladder" significantly more than the novice group for the UVA task (p=0.0093). As experts were confident with their technique, this was often ignored during the task, however, novice surgeons were receptive to the metric and unlikely to trigger it again. This metric may need some revision by the developers to improve the accuracy of the task compared to the real-life procedure. Novice surgeons scored better than intermediate surgeons in "total time instruments are out of view" (p=0.046). This finding was unusual and can not be explained by the authors.

The novice group showed a significantly improved score in 13 performance metrics for the UVA task and 11 for the BND task over 5.5 hours of training. Some performance metrics were very specific, which meant that it was not applicable to the performance of the task. An example of this was "Suspected injury to the Neurovascular Bundle", where damage occurred so infrequently that this performance metric was not useful in terms of assessing a learning curve. This could be said for several others and therefore could be a reason significant improvement was not seen in a greater number of metrics. Of the metrics that showed a significant improvement, generally a step wise progressive improvement was observed after each attempt, even between attempt 5 and 6. Many of the curves have yet to plateau after 6 attempts. This was also observed in the BND task. This could suggest that novice surgeons could still benefit from continuing to use the module, having not fully developed the technical skills needed to complete the module.

VR robotic simulators are widely accepted in the literature through studies that have established face, content, construct, and predictive validity for basic VR modules. [1,5–8] However, there have been far fewer studies investigating procedural VR simulation despite modules existing in robotic hysterectomy, vaginal cuff closure, lobectomy and prostatectomy. Kang et al. [9] have shown face, content and construct, and Kim et al. [10] have shown concurrent and predictive validity for the "Tubes 3" module that simulates a urethrovesical anastomosis part of a RARP. This module differs to the "Tubes 3" module in that it this is the first robotic full procedural VR training module incorporating not only an anastomosis task but a BND task and neurovascular bundle dissection task available for individual and team training. This module also provides a number of additional metrics relating to the suture and approximation of tissue (defined as wound separation as a performance metric) for the UVA task in comparison to "Tubes 3". This module has scored highly regarding realism but like "Tubes 3", there were some comments made by surgeons regarding the realism of the module despite it being scored highly for both tasks and organs individually. Some described the tissue texture as "jelly like" and described "floating needles/sutures" for the UVA task. This could be another area for improvement for module designers.

In urology, basic VR has been adopted alongside other types of simulation to be included in a robotic training curriculum. [11] This curriculum prescribes the use of procedural simulation for training in robotic surgery and also suggests the need to break down operations into individual tasks.[11] The evidence from this study supports that this module could be used for procedural urology training to develop cognitive skills, but also task specific technical skills. To fully understand the benefits of this procedural module, further investigation into the learning curve for intermediates and experts would be necessary. In doing so, the results may give more evidence to suggest where in the curriculum (i.e. what stage) trainee surgeons should be using this procedural module.

65% of intermediates and experts thought that simulation should be part of accreditation and/or (re)certification. There is potential use for this module and other procedural modules in accreditation for trainee robotic surgeons, for example, performing to a certain proficiency prior to performing on a real patient in the operating room. However, before this can happen, benchmarks and predictive validity must be established. Raison et al. [2] have shown that benchmarking can provide an objective target for trainee surgeons allowing a more efficient competency based curriculum. Further use and analysis of this module will provide data that can be used to develop benchmarks. Alongside this, further studies in robotic VR simulation should separate basic and procedural VR training in order to compare the effectiveness of each. This would allow the identification of what specific technical skills trainees develop from each type of VR, and provide evidence for when to implement each during surgical training to allow maximum skill development. It is important to mention that procedural VR modules not only allow the trainee to develop robotic surgical skills, but also allows the user to develop cognitive, team and non-technical skills. [12]

This could be another area of potential research in the future, especially as this module provides team training applications.

There were a number of limitations that were identified in this study. The sample size, in particular expert robotic surgeons, was low. This was due to the limited number of robotic surgeons that were available to participate in this study. Although intermediate surgeon data was included from EAU HOT courses, further expert data from other institutions would make this analysis more robust. A further limitation was the wide variation of experience within both expert and intermediate groups. However, on closer analysis of the expert group, surgeons who had performed 50-120 independent robotic procedures (n=4) and surgeons who had performed >120 independent robotic procedures (n=4) were compared statistically. There was no significant difference when splitting the expert surgeons, and this therefore supports the use of ≥50 independent procedures as an expert definition.

Conclusion

We have established this novel procedural VR module for the RobotiX Mentor surgical simulator to be a valid training tool. Significant improvements with training for a group of novice surgeons was observed. Further studies investigating intermediates and expert surgeons learning curves and scores is essential to identify the benefits of this module for use in surgical curriculums.

Conflict of interest

No authors have any conflict of interest to disclose.

Acknowledgements

The authors thank 3D Systems Simbionix Products for loan of the RobotiX Mentor simulator. The authors received no financial payments of any kind from the company and declare no conflicts of interest. Patrick Harrison acknowledges financial support from the Royal College of Surgeons of England in the form of an Intercalated BSc research award. Kamran Ahmed and Prokar Dasgupta acknowledge the support of financial support from the National Institute for Health Research (NIHR) Biomedical Research Centre based Guy's and St Thomas's NHS Foundation trust. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR, or the Department of Health. Prokar Dasgupta also acknowledges the support of the Medical Research Council Centre for Transplantation, London Deanery, London School of Surgery, Guy's and St. Thomas' Charity, European Union, Vattikuti Foundation, and Olympus.

References

1. Whittaker G, Aydin A, Raison N, Kum F, Challacombe B, Khan MS, et al. Validation of the RobotiX Mentor Robotic Surgery Simulator. J Endourol. 2016 Mar;30(3):338–46.

2. Raison N, Ahmed K, Fossati N, Buffi N, Mottrie A, Dasgupta P, et al. Competency based training in robotic surgery: benchmark scores for virtual reality robotic simulation. BJU Int. 2017 May;119(5):804–11.

3. Abboudi H, Khan MS, Guru KA, Froghi S, de Win G, Van Poppel H, et al. Learning curves for urological procedures: a systematic review: Learning curves for urological procedures. BJU Int. 2014 Oct;114(4):617–29.

4. Sammon J, Perry A, Beaule L, Kinkead T, Clark D, Hansen M. Robot-assisted radical prostatectomy: learning rate analysis as an objective measure of the acquisition of surgical skill: RARP: LEARNING RATE ANALYSIS TO MEASURE ACQUISITION OF SURGICAL SKILL. BJU Int. 2010 Sep;106(6):855–60.

5. Perrenot C, Perez M, Tran N, Jehl J-P, Felblinger J, Bresler L, et al. The virtual reality simulator dV-Trainer® is a valid assessment tool for robotic surgical skills. Surg Endosc. 2012 Sep;26(9):2587–93.

6. Hung AJ, Zehnder P, Patil MB, Cai J, Ng CK, Aron M, et al. Face, Content and Construct Validity of a Novel Robotic Surgery Simulator. J Urol. 2011 Sep;186(3):1019–25.

7. Liss MA, Abdelshehid C, Quach S, Lusch A, Graversen J, Landman J, et al. Validation, Correlation, and Comparison of the da Vinci Trainer ™ and the da Vinci Surgical Skills Simulator ™ Using the Mimic ™ Software for Urologic Robotic Surgical Education. J Endourol. 2012 Dec;26(12):1629–34.

8. Whitehurst SV, Lockrow EG, Lendvay TS, Propst AM, Dunlow SG, Rosemeyer CJ, et al. Comparison of Two Simulation Systems to Support Robotic-Assisted Surgical Training: A Pilot Study (Swine Model). J Minim Invasive Gynecol. 2015 Mar;22(3):483–8.

9. Kang SG, Cho S, Kang SH, Haidar AM, Samavedi S, Palmer KJ, et al. The Tube 3 Module Designed for Practicing Vesicourethral Anastomosis in a Virtual Reality Robotic Simulator: Determination of Face, Content, and Construct Validity. Urology. 2014 Aug;84(2):345–50.

10. Kim JY, Kim SB, Pyun JH, Kim HK, Cho S, Lee JG, et al. Concurrent and predictive validation of robotic simulator Tube 3 module. Korean J Urol. 2015;56(11):756.

11. Ahmed K, Khan R, Mottrie A, Lovegrove C, Abaza R, Ahlawat R, et al. Development of a standardised training curriculum for robotic surgery: a consensus statement from an international multidisciplinary group of experts. BJU Int. 2015 Jul;116(1):93–101.

12. Våpenstad C, Buzink SN. Procedural virtual reality simulation in minimally invasive surgery. Surg Endosc. 2013 Feb;27(2):364–77.

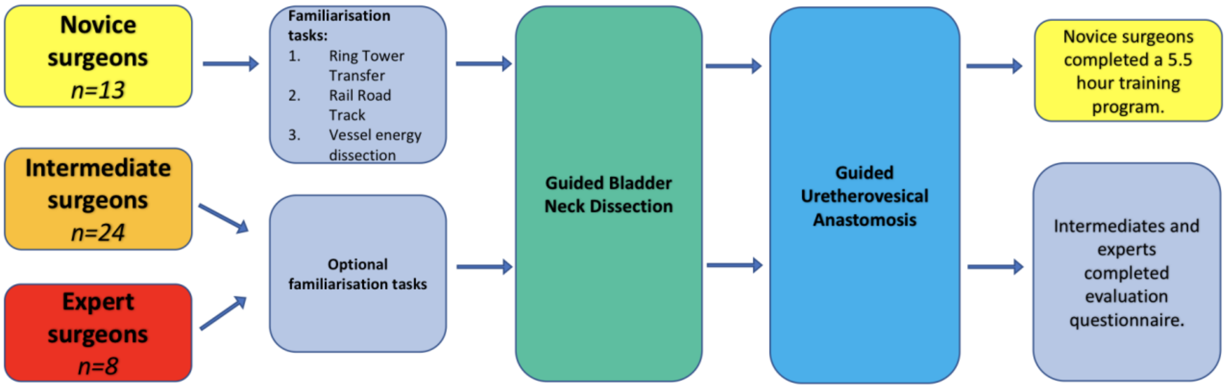

Figure 1- Flow chart illustrating the study process

Table 1- Participant demographics

	Novice	*Intermediate*	*Expert*
No. of Participants	13	24	8
Mean age (range)	23 (21-34)	36 (30-50)	45 (35-62)
No. of grade/specialty			
Novice	13	-	-
Resident	-	16	1
Consultant	-	8	7
Experience: Mean no. of robot-assisted nephrectomy, prostatectomy or cystectomy			
Observed (range)	4 (0-15)	-	-
Assisted (range)	1 (0-8)	33 (n/a-205)	-
Performed (range)	-	6 (0-40)	610 (50-2720)

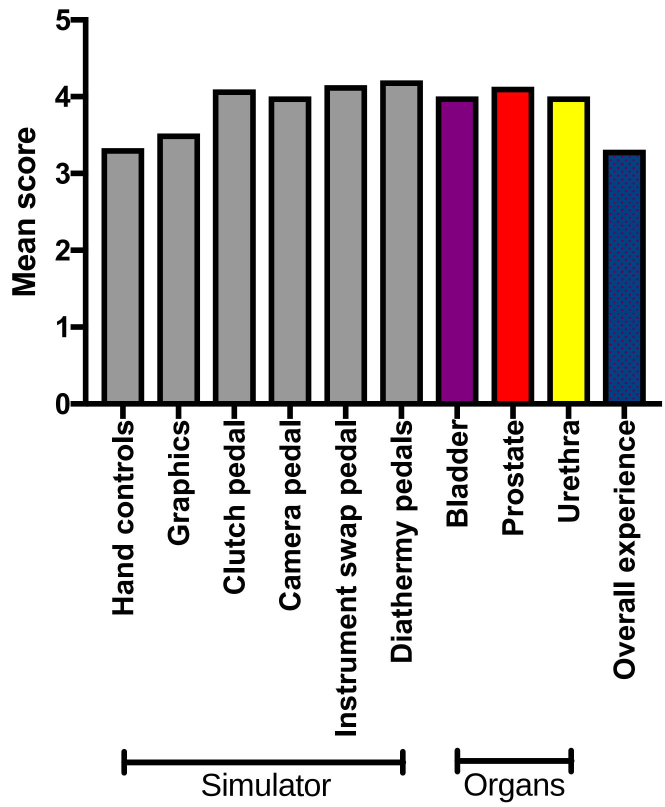

Figure 2- Assessment of realism of simulator controls compared to the da Vinci surgical robot

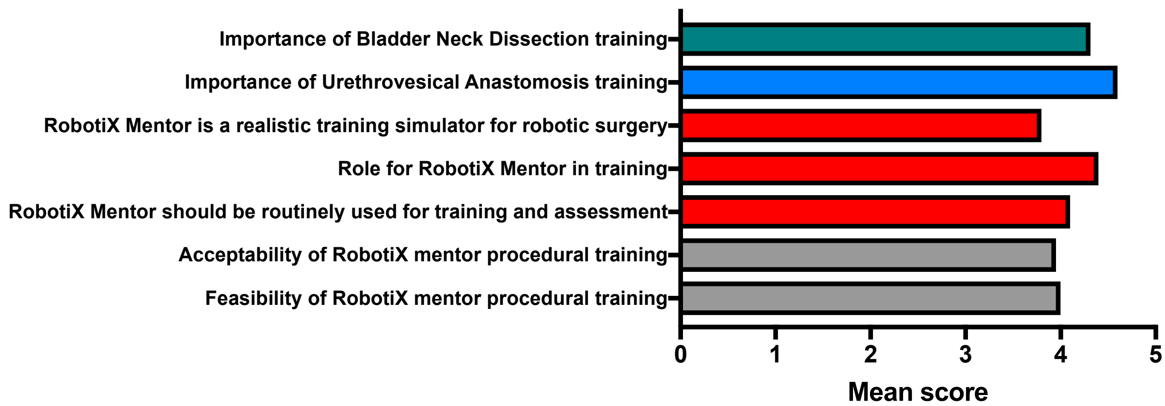

Figure 3- Participant evaluation of importance, feasibility and acceptability of RobotiX mentor procedural training.

Group	Mean Score (Range) Expert	Intermediate	Novice	p-Value Expert vs Novice	Expert vs Intermediate	Intermediate vs Novice	Comments
Generic metrics							
Total time (seconds)	1158.11 (944.95-1486.74)	1617.73 (911.84-2219.56)	2055.83 (1480.67-3033.39)	0.0005**	0.0120*	0.0172*	Experts were significantly faster than intermediates and novices. Intermediates were significantly faster than novice.
Distance by camera	3547.11 (1794.59-8461.53)	1865.62 (419.14 3306.44)	1360.1 (347.05-2040.37)	0.0010*	0.0157*	0.0443*	Experts had significantly more movements than novice and intermediates.
Instrument collisions	30 (16-51)	65(14-129)	73 (26-132)	0.0033**	0.0239*	0.5139	Experts were significantly better than intermediates and novices.
Total time instruments are out of view (Seconds)	120.11 (58.79-230.65)	206.11 (86.94-580.79)	142.57 (11.35-569.02)	0.9368	0.0829*	0.046*	Both novice and expert had a significantly shorter time than intermediates.
Number of movements (left instrument)	1228 (797-2402)	1367 (833-1987)	1414 (1031-1576)	0.0648	0.0844	0.7012	
Number of movements (right instrument)	1321 (824-2750)	1454 (859-2218)	1520 (1245-1871)	0.0043**	0.0280*	0.5393	Experts had more efficient movements than intermediates and novices.
Path length (left instrument)	13040.9 (8297.53-25917.4)	13654.1 (9180.41-19125.8)	13218.5 (9306.69-18077.5)	0.3282	0.1933	0.6452	
Path length (right instrument)	11428 (8588.63-14142.8)	13400.3 (9014.37-19268.1)	13874.3 (10625.1-18939.8)	0.0324*	0.1186	0.6452	Experts had a significantly lower distance travelled than novices
Task-specific metrics							
Number of times the tissue was grasped	128 (2-421)	77 (7-215)	26 (3-58)	0.0833	0.5239	0.0173*	Experts grasped tissue with the free hand when placing the suture, whereas novices did not.
No. of unnecessary needle piercing points	74 (37-164)	67 (37-145)	109 (55-184)	0.0463*	0.9576	0.0023**	Experts and intermediates had significantly fewer unnecessary needle piercing points than novices.
Respect for tissue	4.5 (0-14)	7 (1-19)	7 (1-18)	0.2292	0.2293	0.9232	
Suspected injury to the bladder	1 (0-2)	0 (0-1)	0	0.0093**	0.1973	0.0605	Novices had significantly less injuries to the bladder compared with experts, and marginally less than intermediates
Injury to urethra	4 (0-10)	8 (1-19)	9 (2-19)	0.0535	0.0933	0.5900	Experts had marginally significantly less than novice.

p-Values are displayed *<0.05 **p<0.01 ***p<0.0001

Table 1- UVA task technical performance and group comparison.

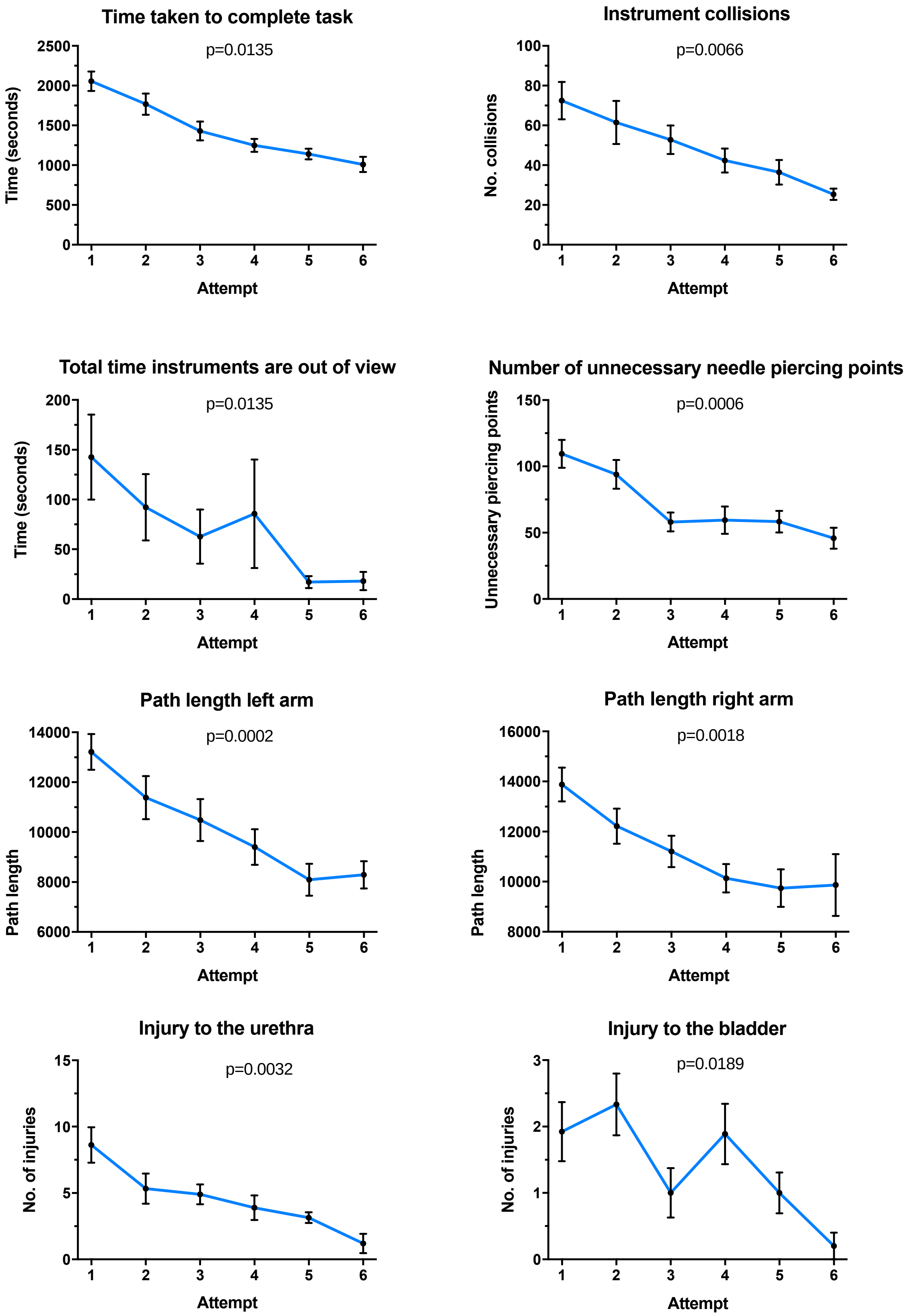

Figure 4- Learning curves for selected performance metrics for the UVA task.