Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.JournalofSurgicalResearch.com

Surgeons With Five or More Actual Cricothyrotomies Perform Significantly Better on a Virtual Reality Simulator

Di Qi, PhD,[a,*] *Emil Petrusa, PhD,*[b] *Uwe Kruger, EngD,*[a,c] *Nicholas Milef, BS,*[a] *Mohamad Rassoul Abu-Nuwar, MD,*[d] *Mohamad Haque, MD,*[e] *Robert Lim, MD,*[f] *Daniel B. Jones, MD,*[d] *Melih Turkseven, PhD,*[a] *Doga Demirel, PhD,*[g] *Tansel Halic, PhD,*[h] *Suvranu De, ScD,*[a] *and Noelle Saillant, MD*[b]

[a] Center for Modeling, Simulation, and Imaging in Medicine, Rensselaer Polytechnic Institute, Troy, New York
[b] Department of Surgery, Massachusetts General Hospital, Harvard School of Medicine, Boston, Massachusetts
[c] Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, New York
[d] Department of Surgery, Beth Israel Deaconess Medical Center, Harvard School of Medicine, Boston, Massachusetts
[e] Anderson Simulation Center, Madigan Army Medical Center, Tacoma, Washigton
[f] Tripler Army Medical Center, Honolulu, Hawaii
[g] Department of Computer Science, Florida Polytechnic University, Lakeland, Florida
[h] Computer Science Department, University of Central Arkansas, Conway, Arkansas

ARTICLE INFO

Article history:
Received 27 November 2019
Received in revised form 28 January 2020
Accepted 8 March 2020
Available online 15 April 2020

Keywords:
Surgical simulator
Simulation performance
Cricothyrotomy
Virtual reality

ABSTRACT

Background: Discriminating performance of learners with varying experience is essential to developing and validating a surgical simulator. For rare and emergent procedures such as cricothyrotomy (CCT), the criteria to establish such groups are unclear. This study is to investigate the impact of surgeons' actual CCT experience on their virtual reality simulator performance and to determine the minimum number of actual CCTs that significantly discriminates simulator scores. Our hypothesis is that surgeons who performed more actual CCT cases would perform better on a virtual reality CCT simulator.
Methods: 47 clinicians were recruited to participate in this study at the 2018 annual conference of the Society of American Gastrointestinal and Endoscopic Surgeons. We established groups based on three different experience thresholds, that is, the minimal number of CCT cases performed (1, 5, and 10), and compared simulator performance between these groups.
Results: Participants who had performed more clinical cases manifested higher mean scores in completing CCT simulation tasks, and those reporting at least 5 actual CCTs had significantly higher ($P = 0.014$) simulator scores than those who had performed fewer cases. Another interesting finding was that classifying participants based on experience level, that is, attendings, fellows, and residents, did not yield statistically significant differences in skills related to CCT.
Conclusions: The simulator was sensitive to prior experience at a threshold of 5 actual CCTs performed.

* *Corresponding author.* Center for Modeling, Simulation, and Imaging in Medicine, Rensselaer Polytechnic Institute, Troy, NY 12180. Tel.: +1(518)421-9091; fax: +1(518)-276-4002
E-mail address: qid@rpi.edu (D. Qi).
0022-4804/$ – see front matter © 2020 Elsevier Inc. All rights reserved.
https://doi.org/10.1016/j.jss.2020.03.021

Introduction

With recent advances in computer graphics and haptic interface technology, virtual reality (VR) simulation has become an important component of medical education and training reforms.[1-7] In addition to allowing surgeons to practice in a safe and controlled environment with no risk to patients, well-constructed VR surgery simulators also standardize training, save instructor resources, and assure uniformity of training.

Evidence to claim the validity of a VR simulator, however, should occur before stand-alone use in training programs. The ability of a simulator to accurately discriminate the performance of learners with varying experience is a component of validity evidence.[8,9] This is often evaluated by comparing the performance between groups with different experience.[8-10] In this regard, superior performance by the group with more actual experience is pivotal in validating the use of the simulator for training and assessment.[9] To this end, creating groups that are likely different in the ability on the simulator tasks is necessary.

It is important to note that the way of creating groups with varying experience varies across studies.[3,9] A common approach is based on their medical experience.[9,11,12] In a comprehensive survey[9] regarding the methodologies of validating surgical simulators, Van Nortwick *et al.* point out that attendings are frequently identified as more experienced, while fellows or senior residents also serve as more experienced in a third of the studies. Moreover, the number of actual clinical cases managed is another criterion for establishing groups with different expertise. Individuals who have extensive clinical experience in terms of performing a sufficiently large number of surgeries are regarded as more experienced.[3,13,14] However, for some rare and emergent surgeries such as cricothyrotomy (CCT), the opportunity for performing such procedures is infrequent,[15-17] which renders creating groups that are likely different on simulator tasks a difficult task.

CCT has been considered as a life-saving emergency procedure for patients who cannot be intubated by conventional means (e.g., rapid sequence intubation) and would otherwise risk death.[16] To assure competency with CCTs given so few actual opportunities, we need to supplement actual experience with simulated training mechanisms. Traditionally, this approach can be seen in the Advanced Trauma Life Support course[18] using human cadavers or plastic mannequins. Despite the high cost of these training models, none of the aforementioned simulations offer repetitive practice and anatomic variants.[19,20] VR simulation, on the other hand, that offers repetitive and unlimited practice and automated objective assessment can provide an alternative that addresses these problems in a cost-effective manner.[21] However, no focus has been placed on evaluating the validity of existing CCT simulators [19,20] in terms of discriminating groups with varying experience. In fact, there are no studies indicating the impact of the surgeon's actual CCT experience on the simulator performance. This may result in establishing comparison groups showing no difference in simulator scores.

In this article, we seek to determine the minimal number of actual CCTs that significantly discriminates performance on a VR CCT simulator. Our hypothesis is that surgeons with more actual CCT experience perform better on a VR CCT simulator than those who performed fewer cases. To determine the minimum number of actual CCTs that significantly discriminates simulator performance, we explored the impact of three different actual experience thresholds (AET), that is, the minimum number of actual CCT cases performed (1, 5, and 10), on simulator scores. We also compared the simulation performance between groups established based on the participants' medical experience (e.g., attending or nonattending) to confirm our hypothesis.

Material and methods

Participants

To identify a sufficiently large sample of subjects with varied clinical experience, we conducted this study at the 2018 annual conference of the Society of American Gastrointestinal and Endoscopic Surgeons. The study was approved by the institutional review board at Beth Israel Deaconess Medical Center. 47 participants were voluntarily recruited to perform CCT tasks on a VR simulator. All participants were informed about the purpose of this study, and all of them gave their informed consent before participating in the study. The demographic data are shown in Table 1.

VR CCT simulator

Configuration

The haptic-enabled VR CCT simulator, *VAST-CCT*,[22] consists of a desktop computer and a 24-inch computer monitor connected to two Geomagic Touch haptic devices, each of which

Table 1 – Demographic data.

Age, average (range)	36.8 (22-88)
Sex, female:male	10:35 (2 missing responses)
Hand dominance, left:right:ambidextrous	2:41:2 (2 missing responses)
Medical experience	17 attendings, 2 fellows, 12 residents, 2 medical students, 11 combat medics, 3 undergrads
CCT performed experience, yes:no	17:30
Video game experience, yes:no	20:27

provides 6 degrees of freedom position/orientation input and 3 degrees of freedom force feedback.

As illustrated in Figure 1A and B, the participants were seated in front of the monitor that displayed a virtual patient lying in a supine position. All instruments required in the surgery were integrated into the simulator. These instruments can be easily switched from one another and operated freely using the haptic device held by the user's dominant hand. The virtual hand for palpation is controlled by the haptic device operated by the user's nondominant hand. The anatomical models (e.g., thyroid and cricoid cartilages, cricothyroid membrane, trachea, *etc.*) of the patient, as shown in Figure 1C, were derived from high-resolution computed tomography data. The simulation was continuously updated in real-time based on the user's interactions with the two haptic devices.

Simulation tasks

On the basis of a hierarchical task analysis of CCT discussed in the study by Demirel et al,[17] four key tasks were identified and implemented in the simulator: (1) landmarks identification, (2) skin incision, (3) cricothyroid membrane incision, and (4) intubation. The description of each task is illustrated in Figure 2.

In the simulator, the patient's skin, cartilages, and membrane were modeled with different material properties using real-time physics-based modeling techniques.[23] When interacting with different regions of the patient's neck, the haptic devices could provide the user with a distinct sensation of palpation. When the virtual scalpel was pressed on the skin with sufficient force, a cut was initiated.

Automated performance assessment

The user's performance was assessed based on the completeness and accuracy of each task and quantified automatically by the simulator using an overall score, referred to as performance score in this article. Four performance assessment metrics were implemented in the simulator, one for each task (see Fig. 2). For a detailed explanation of each metrics, we refer the user to our prior study.[17] The overall performance score is the sum of all four task scores, with a range from 0 to 40. The scoring for the performance assessment metrics of each task is detailed in Table A in the Appendix.

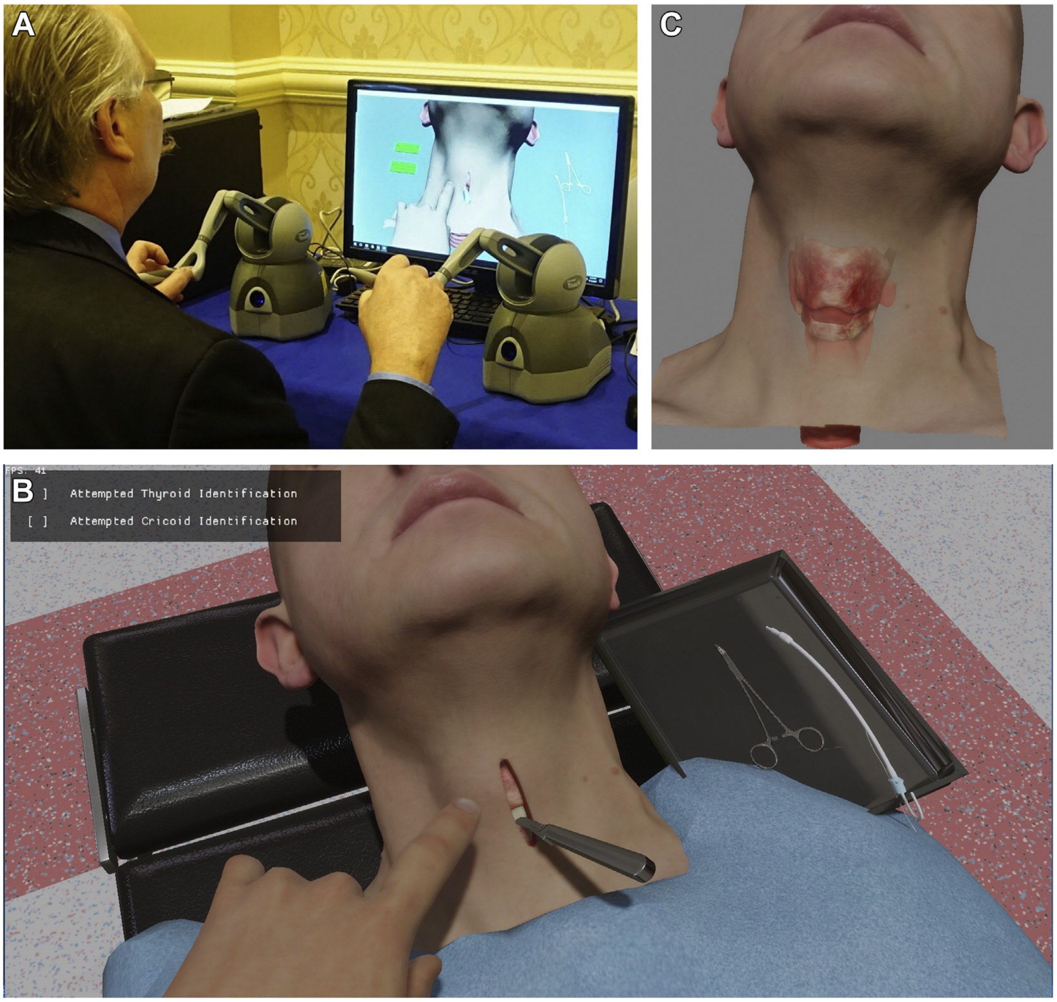

Fig. 1 – VAST-CCT simulator: (A) Haptic-enabled VR CCT simulator, (B) the virtual patient and surgical instruments, and (C) the virtual cricothyroid anatomy are derived from high-resolution CT data and simulated with realistic material properties. (Color version of figure is available online.)

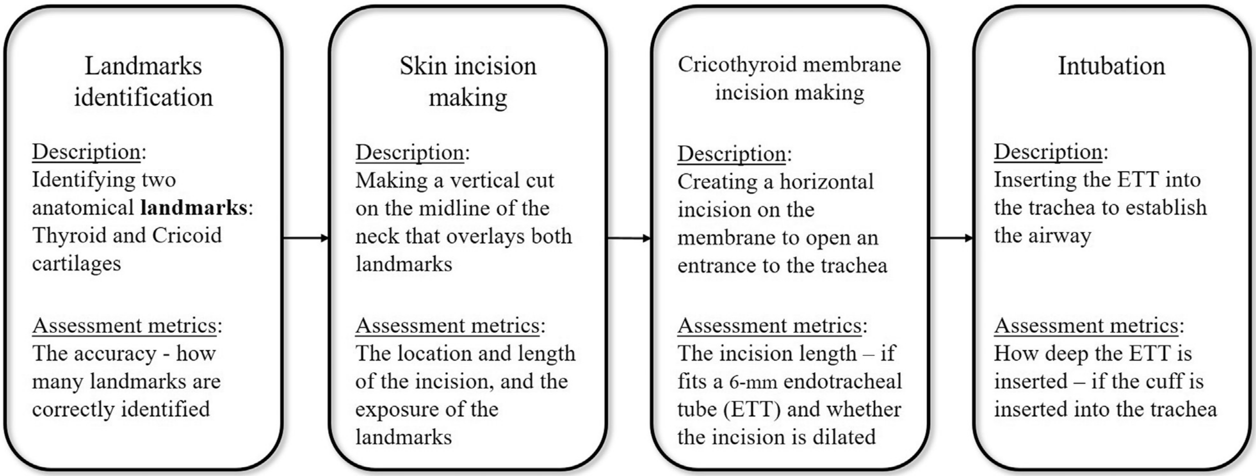

Fig. 2 – Simulation tasks and performance assessment metrics.

Procedure

Before conducting the experiment, participants were asked to fill out a questionnaire about their demographics, medical experience, and the total number of CCTs they had performed as shown in Table 1. The number of cases performed was noted to be 0-20 with a median of 0. The participants were then presented with an instruction demo describing the CCT tasks and how to operate the simulator, accompanied by a verbal explanation. Each participant could perform up to three practice trials to become familiar with the setup of the simulator such as how to operate the haptic device to control the surgical instruments or change the instruments in VR. Once the practice phase was over, each participant performed only one trial on the simulator as the formal test, and a performance score was recorded. Participants were then asked to fill out a questionnaire consisting of nine questions by rating from 1 (not realistic/useful) to 5 (very realistic/useful) regarding the face validation of the CCT simulator.

Establishment of groups with different CCT experience

To determine the minimum number of actual CCTs that significantly discriminates performance on a CCT simulator, we explored the impact of three different AETs, that is, the minimum number of actual CCT cases performed (1, 5, and 10), on simulator scores. The selection of AET was based on the distribution of the actual CCT experience among the participants. The number of CCT cases that the participants had performed was noted to be 0-20 with a median of 0. Because more than half (30 of 47) of the participants had no CCT experience, 1 was chosen as the minimum number of cases. Among the 17 participants who had performed CCT before, the number of CCT cases performed ranged from 1 to 20 with a median of 5; we then chose the median number, that is, 5, as the second experimental threshold. Similarly, among the participants who had performed CCT a minimum of 5 times, the number of CCTs ranged from 5 to 20 with a median of 10, and the third experimental threshold was therefore chosen as 10.

For each AET, participants who had performed greater or equal to the number of cases signified in the threshold were assigned to the more-experienced (ME) group; otherwise, they were assigned to the less-experienced (LE) group. Among the 47 participants, the participant numbers for ME and LE groups established in accordance with each AET are different as shown in Table 2.

Establishment of groups with different medical experience

To investigate the impact of surgeons' medical experience (i.e., attendings, fellows, *etc.*) on their performance of VR CCT simulator, we established two groups based on whether a participant is an attending or not and then compared their simulator performance using the same data. Similarly, we also assigned fellows and senior residents (PGY-5) to the same group of attendings and compared their simulator performance to the rest of the participants. As shown in Table 1, among the 47 participants, 17 were attendings, 2 were fellows, and 12 were residents (3 of which were PGY-5).

Statistical analysis

For each AET, that is, 1, 5, or 10, the performance scores between the ME and LE groups were compared. Our purpose was

Table 2 – Of the 47 participants, the number of participants assigned to more-experienced and less-experienced groups for each actual experience threshold.

Actual experience thresholds (AET)	# More-experienced (ME)	# Less-experienced (LE)
1	17	30
5	9	38
10	6	41

to estimate the minimum number of actual CCTs performed that significantly discriminate simulator performance.

All data (performance scores) were examined for normality using the Anderson-Darling test. The standard F test was then used to assess the equality of variances between the ME and LE data for each threshold. As the performance scores for each group were normally distributed and had the same variance, a two-sample *t*-test (with an equal variance assumption) was used to compare the performance scores between the ME and LE groups. Our null hypothesis was that the performance of ME and LE groups is the same. Because three comparisons between ME and LE groups were conducted with respect to the three AETs (i.e., 1, 5, 10), each comparison was tested at a significance level of 0.0167 (i.e., 0.05/3) based on Bonferroni correction.

By following this same procedure, we also compared the simulator performance between groups with different medical experience such as attending and nonattending groups.

The face validation questions were analyzed by calculating the mean, standard deviation (SD), and percent of responses that were greater than 3 on a 5-point scale.

Results

The optimal AET distinguishing simulator performance

Performance comparisons between ME and LE groups established based on all three AETs are summarized in Figure 3, where performance scores of ME and LE groups are plotted with respect to each threshold. The results demonstrated that no significant difference was found between these two groups when the minimum number of CCTs was 1 ($P = 0.066$); the ME group performed significantly better than the LE group when they had performed at least 5 CCTs ($P = 0.014$) or 10 CCTs ($P = 0.005$). We, therefore, considered 5 to be the minimum actual CCT experience distinguishing simulator performance.

Furthermore, when raising the minimum number of cases, the mean performance score increases (i.e., 22.53 $\pm$ 6.98, 25.33 $\pm$ 7.94, 27.83 $\pm$ 7.08) for the ME group but almost same (i.e., 18.30 $\pm$ 7.62, 18.53 $\pm$ 7.01, 18.66 $\pm$ 7.01) for the LE group, and the difference (gap) between the two groups enlarges. In the meantime, the same trend can be found in the subscores of the four tasks that make up the total performance score. We have included the mean scores of the four tasks for all the participants across the three experience thresholds in Fig. A in the Appendix. Moreover, we also found that the ME group generally takes much less time than the LE group to complete the entire procedure. With the increase in the minimum number of cases, the median completion time (in seconds, non-normal distribution) for both the ME group (i.e., 123, 116, 117) and the LE group (i.e., 214.5, 209.5, 205) does not change much.

Performance comparison between groups with different medical experience

As illustrated in Figure 4, no significant difference ($P = 0.506$) was found between attendings and nonattendings, although the attending group did obtain a slightly higher (mean: 20.82 $\pm$ 7.63) score than the nonattending group (mean: 19.27 $\pm$ 7.66). Similarly, no significant difference ($P = 0.769$) was found when comparing the group of attendings, fellows, and senior residents (mean: 20.18 $\pm$ 7.93) to the other participants (mean: 19.52 $\pm$ 7.46).

Face validation

Post-task questionnaire results are shown in Table 3 on a 5-point scale. For all the participants, eight of the nine questions were rated a mean of 3 or greater (89%). The highest rating was

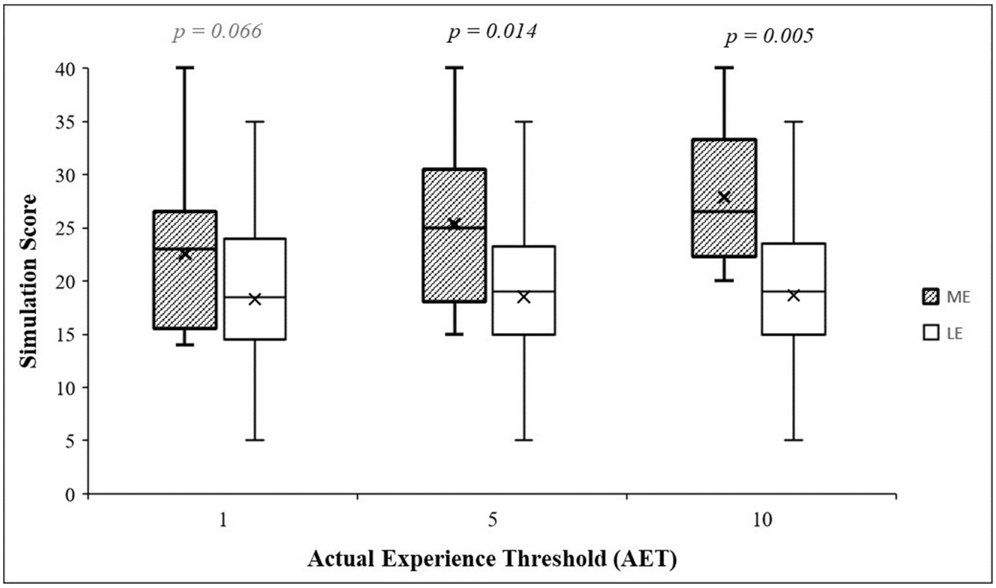

Fig. 3 – VR simulator performance scores comparison between more-experienced (ME) and less-experienced (LE) groups that were established based on three actual experience thresholds (i.e., the minimum number of actual CCTs performed): 1, 5, and 10.

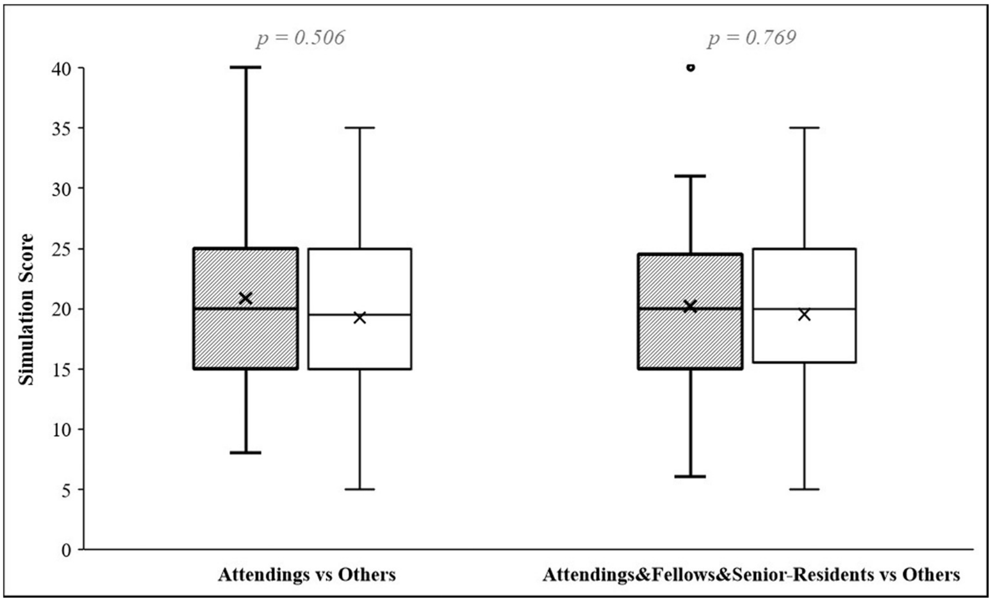

Fig. 4 – VR simulator performance comparison between groups that were established based on participants' medical experience.

assigned to the overall usefulness of the simulator in learning fundamental CCT technical skills (mean: 3.91). 89% of the participants rated the usefulness of the force feedback (sensation of feeling the tools on the target and in the task space) in the simulator in helping their performance at least 3 of 5, and 91% of the participants rated the degree of overall realism (looks and feels) of the simulator at least 3 of 5. The question that rated the lowest was the degree of realism of the instrument handling in the simulator (2.99).

Discussion

Selection of candidate actual experience thresholds

As explained in the Methods section, the three AETs, 1, 5, and 10, were selected based on the distribution of the prior CCT experience among the 47 participants. To the best of our knowledge, there is no literature investigating the impact of the actual CCT experience on VR simulation performance. Therefore, we decided to derive the potential thresholds from actual data. We considered that the data that were collected from a national surgical conference could help us understand the actual CCT experience from a general surgical population and so as to investigate the impact of actual CCT experience on the simulation performance.

Performance comparison between groups with different experience

Another common way of creating groups with different experience is based on the surgeon's medical experience. Attendings, fellows, and senior residents were frequently identified as more experienced in many studies.[9] To examine whether this classification is appropriate for CCT, we further compared the performance between groups with different medical experience using the same data. Based on the result illustrated in Figure 4, no significant difference was found between groups established based on either of the categories. Our interpretation of this result is that CCT is a rare surgical procedure even for attendings, especially for those who are not from the disciplines offering more opportunities to perform CCT. Simply identifying a surgeon's experience of surgery based on their medical experience may create groups showing no difference in simulator scores. This finding thus justifies our hypothesis of exploring the impact of the number of clinical cases on the simulator performance.

Technical challenges of VR CCT simulation

Based on the feedback from the participants, the overall usefulness of the simulator was considered moderately realistic (mean rating 3.42 of 5), and 93% of the participants rated the usefulness of the simulator in learning the fundamental CCT technical skills over a 3 of 5, with a mean of 3.91. However, the realism of the instrument handling received the lowest rating (mean rating 2.99 of 5). Several participants commented that it was not realistic to use a hand-held haptic stylus to palpate the virtual patient's skin for the landmark identification, where surgeons usually use their fingers to palpate during the real surgery; the force feedback received during palpation was helpful but too weak. This suggested that the interface for CCT landmark identification needs to be improved and more realistic force feedback rendering is demanded. Prior studies[21,22] on VR surgical simulations have also found that the handling of instruments and quality of force feedback received lower rates. This indicates the importance of instrument handling and force feedback in the design of VR surgical trainers.

Table 3 – Post-task questionnaire.

Questions: Rating from 1 (not realistic/useful) to 5 (very realistic/useful)	Mean	SD	Rating ≥ 3
Degree of realism of the target objects (how realistic they look) in the VR environment	3.56	0.84	89%
Degree of realism of the instrument handling (how realistic it feels) in VAST-CCT	2.99	1.00	61%
Degree of overall realism of the VAST-CCT simulation (how it looks AND feels)	3.42	0.84	91%
Quality of the force feedback (sensation of feeling the tools on the target and in the task space) in the VAST-CCT.	3.51	1.02	80%
Degree of usefulness of the force feedback (sensation of feeling the tools on the target and in the task space) in the VAST-CCT in helping your performance.	3.71	0.90	89%
Usefulness of the VAST-CCT simulation in learning hand-eye coordination skills	3.93	1.00	86%
Usefulness of the VAST-CCT simulation in learning ambidexterity skills	3.80	1.07	86%
Degree of overall usefulness of the VAST-CCT in learning the fundamental cricothyrotomy technical skills	3.91	0.83	93%
Assessment of how trustworthy the VAST-CCT is to quantify accurate measures of performance	3.49	0.90	86%

Limitations and future work

This study had a few limitations that should be acknowledged.

First, we only evaluated surgeons' CCT performance on a VR simulator that presents a straightforward situation. However, CCT is not without complications.[16,24] Performance on a patient with normal anatomy is often considered to be less challenging than on patients with a short, fat neck, or patients who cannot have their neck placed in hyperextension because of limited mobility or other safety concerns. Other important variables such as the difficulty level of actual CCTs performed by each participating surgeon and their impact on the simulator performance need to be analyzed in future research.

Second, this study only considered the impact of surgeons' actual CCT experience on the simulator performance. In fact, other confounding factors, including the experience with other CCT simulations such as mannequins used in Advanced Trauma Life Support course should be considered when comparing groups, as it may influence VR simulation scores. Moreover, despite the opportunity for performing CCT is infrequent, there are some other procedures that contribute to the skills necessary for successful CCT performance such as thyroidectomy and tracheostomy. The influence of surgeons' experience of those procedures could also be taken into account in future studies.

Third, although there were 47 participants in this study, only a small number of them (17) had actual CCT experience. Although this limited the ability to evaluate the system in terms of skill differentiation, the clinicians who do not have much chance to perform CCT are the potential target user group for training. Future work is needed to determine whether the VAST-CCT simulator can be used to distinguish between experts and novice surgeons such as those who are specialized in trauma or acute care. CCT is an important and life-saving skill where there is an infrequent opportunity for practice. VR simulation that provides repetitive, patient-safety practice and automated performance assessment could augment training for such a critical skill. This study explored the impact of different numbers of actual CCTs performed on VR simulator scores. We found that participating surgeons who had experience with a minimum of 5 actual CCTs performed significantly better on the simulator than those who had performed fewer cases, based on group differences. Based on the feedback from the participants, the VR simulator was considered moderately realistic and useful by the participants to learn the fundamental CCT technical skills. These indicate the potential validity of the simulator in supporting the training on such low-frequency but high-acuity procedure. However, due to the limitations of instrument handling and force feedback as mentioned by the participants, the current system can supplement rather than replace the current training methods. However, with further improvements and validation, the VAST-CCT simulator can supplement the current training methods as a cost-effective manner of practice. In the next stage of the study, the training effectiveness of the CCT simulator could be assessed by comparing participants' performance after training on different simulation systems such as mannequins and their ability to transfer their skill between simulations and to retain the skill over time.

Acknowledgment

The research reported in this article was supported by the NIH/NIBIB under Award Number R01EB005807, R01EB010037, R01EB009362, R01EB014305, R01EB025241; NIH/NHLBI under Award Number R01HL119248; NIH/NCI under Award Number R01CA197491; and NIH under Award Number R44OD018334.

Authors' contributions: All the authors have seen and approved this manuscript and contributed equally to the submission. D.Q. contributed to data acquisition, data analysis, simulation development, and was primary author of the article. E.P. contributed to the study conception and design, data analysis, and article drafting. U.K. contributed to data analysis and article drafting. N.M. contributed to article drafting and simulation development. M.R.A. contributed to data acquisition and simulation development. M.H. contributed to study design and data acquisition. R.L. contributed to data

acquisition and article drafting. D.B.J. contributed to study design and article drafting. M.T. contributed to data acquisition and simulation development. D.D. and T.H. contributed to task analysis and simulation development. S.D. and N.S. contributed to study design, article drafting, simulation development, and final approval before submission.

Disclosure

Drs. Di Qi, Emil Petrusa, Uwe Kruger, Mohamad Rassoul Abu-Nuwar, Mohamad Haque, Melih Turkseven, Doga Demirel, Tansel Halic, Suvranu De, and Nicholas Milef have no conflicts of interest or financial ties to disclose. Dr. Noelle Saillant has no relevant conflicts related to this manuscript and is on the Haemonetics TEG 6s advisory board. Dr. Robert Lim has no relevant conflicts related to this manuscript and reports personal fees from UpToDate, Inc. Dr. Daniel B. Jones has no relevant conflicts related to this manuscript and is on the advisory board of Allurion Technologies Inc.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jss.2020.03.021.

REFERENCES

1. Alaraj A, Lemole MG, Finkle JH, et al. Virtual reality training in neurosurgery: review of current status and future applications. *Surg Neurol Int*. 2011;2:52.
2. Ruthenbeck G, Reynolds K. Virtual reality for medical training: the state-of-the-art. *J Simul*. 2015;914:16–26.
3. Moglia A, Ferrari V, Morelli L, Ferrari M, Mosca F, Cuschieri A. A systematic review of virtual reality simulators for Robot-assisted surgery. *Eur Urol*. 2016;69:1065–1080.
4. Khor WS, Baker B, Amin K, Chan A, Patel K, Wong J. Augmented and virtual reality in surgery-the digital surgical environment: applications, limitations and legal pitfalls. *Ann Transl Med*. 2016;4:454.
5. Ward RC, Muckle TJ, Kremer MJ, Krogh MA. Computer-based case simulations for assessment in health care: a literature review of validity evidence. *Eval Health Prof*. 2017, 016327871771860.
6. Wynn G, Lykoudis P, Berlingieri P. Development and implementation of a virtual reality laparoscopic colorectal training curriculum. *Am J Surg*. 2017;216:610–617.
7. Harrington CM, Kavanagh DO, Quinlan JF, et al. Development and evaluation of a trauma decision-making simulator in Oculus virtual reality. *Am J Surg*. 2018;215:42–47.
8. McDougall EM. Validation of surgical simulators. *J Endourol*. 2007;21:244–247.
9. Van Nortwick SS, Lendvay TS, Jensen AR, Wright AS, Horvath KD, Kim S. Methodologies for establishing validity in surgical simulation studies. *Surgery*. 2010;147:622–630.
10. Gallagher AG, O'Sullivan GC. *Fundamentals of surgical simulation: principles and practice*. Springer-Verlag London: Springer; 2012.
11. Arikatla VS, Ahn W, Sankaranarayanan G, De S. Towards virtual FLS: development of a peg transfer simulator. *Int J Med Robot*. 2014;10:344–355.
12. Chellali A, Zhang L, Sankaranarayanan G, et al. Validation of the VBLaST peg transfer task: a first step toward an alternate training standard. *Surg Endosc Other Interv Tech*. 2014;28:2856–2862.
13. Suebnukarn S, Chaisombat M, Kongpunwijit TS, Rhienmora PS. Construct validity and expert benchmarking of the haptic virtual reality dental simulator. *J Dent Educ*. 2014;78:1442–1450.
14. Varoquier M, Hoffmann CP, Perrenot C, Tran N, Parietti-Winkler C. Construct, face, and content validation on voxel-man® simulator for otologic surgical training. *Int J Otolaryngol*. 2017;2017:1–8.
15. Wang EE, Vozenilek JA, Flaherty J, Kharasch M, Aitchison P, Berg A. An innovative and inexpensive model for teaching cricothyrotomy. *Simul Healthc*. 2007;2:25–29.
16. Langvad S, Hyldmo PK, Nakstad AR, Vist GE, Sandberg M. Emergency cricothyrotomy - a systematic review. *Scand J Trauma Resusc Emerg Med*. 2013;21:43.
17. Demirel D, Butler KL, Halic T, et al. A hierarchical task analysis of cricothyroidotomy procedure for a virtual airway skills trainer simulator. *Am J Surg*. 2016;212:475–484.
18. Simulation in trauma education: beyond ATLS References. *Injury*. 2014;45:817–818.
19. Liu A, Bhasin Y, Bowyer M. A haptic-enabled simulator for cricothyroidotomy. *Stud Health Technol Inform*. 2005;111:308–313.
20. Demirel D, Yu A, Halic T, et al. Virtual airway skills trainer (VAST) simulator. *Stud Health Technol Inform*. 2016;220:91–97.
21. Fu Y, Cavuoto L, Qi D, et al. Validation of a virtual intracorporeal suturing simulator. *Surg Endosc*. 2018;33:2468–2472.
22. Sankaranarayanan G, Lin H, Arikatla VS, et al. Preliminary face and construct validation study of a virtual basic laparoscopic skill trainer. *J Laparoendosc Adv Surg Tech A*. 2010;20:153–157.
23. Muller M, Heidelberger B, Hennix M, Ratcliff J. Position based dynamics. *J Vis Commun Image Represent*. 2007;18:109–118.
24. Davis DP, Bramwell KJ, Vilke GM, Cardall TY, Yoshida E, Rosen P. Cricothyrotomy technique: standard versus the rapid four-step technique. *J Emerg Med*. 1999;17:17–21.