The effect of simulator fidelity on procedure skill training: a literature review

Alan Kawarai Lefor[1], Kanako Harada[1], Hiroshi Kawahira[2], Mamoru Mitsuishi[1]

[1]Department of Bioengineering, School of Engineering, The University of Tokyo, Tokyo, Japan
[2]Jichi Medical Simulation Center, Jichi Medical University, Tochigi, Japan

Correspondence: Alan Kawarai Lefor, Jichi Medical University, Yakushiji 3311-1, Shimotsuke, Tochigi, Japan
Email: alefor@jichi.ac.jp

Abstract

Objective: To evaluate the effect of simulator fidelity on procedure skill training through a review of existing studies.

Methods: MEDLINE, OVID and EMBASE databases were searched between January 1990 and January 2019. Search terms included "simulator fidelity and comparison" and "low fidelity" and "high fidelity" and "comparison" and "simulator". Author classification of low- and high-fidelity was used for non-laparoscopic procedures. Laparoscopic simulators are classified using a proposed schema. All included studies used a randomized methodology with two or more groups and were written in English. Data was abstracted to a standard data sheet and critically appraised from 17 eligible full papers.

Results: Of 17 studies, eight were for laparoscopic and nine for other skill training. Studies employed evaluation methodologies, including subjective and objective measures. The evaluation was conducted once in 13/17 studies and before-after in 4/17. Didactic training only or control groups were used in 5/17 studies, while 10/17 studies included two groups only. Skill acquisition and simulator fidelity were different for the level of training in 1/17 studies. Simulation training was followed by clinical evaluation or a live animal evaluation in 3/17 studies. Low-fidelity training was not inferior to training with a high-fidelity simulator in 15/17 studies.

Conclusions: Procedure skill after training with low fidelity simulators was not inferior to skill after training with high fidelity simulators in 15/17 studies. Some data suggest that the effectiveness of different fidelity simulators depends on the level of training of participants and requires further study.

Keywords: Simulation education, low-fidelity, high-fidelity, laparoscopic surgery, skill assessment

Introduction

The former paradigm for teaching medical procedures, "see one, do one, teach one", has no basis in educational theory and has been abandoned. Simulation education in medicine is being revised in an educational theory framework. Deliberate practice, cognitive task analysis and proficiency-based training are being used to design improved educational programs.[1] Laparoscopic surgery training in the late 1980s and early 1990s often proceeded without verification of skills or competency.[2] Since that time, training to perform procedures has undergone widespread change and simulation has become an important component. Selecting the optimal simulator for a specific procedure and trainee group is difficult.

Simulator fidelity is must be considered when developing simulation curricula. The definition of fidelity is a crucial issue and not consistently applied.[3,4] There are both objective (mathematical) and subjective definitions (based on a trainee's performance matrix) of fidelity.[5] Some authors have differentiated between psychological fidelity and engineering fidelity.[6] While some feel that fidelity relates to the replication of reality, a recent study suggests that an accurate representation of cues and stimuli is more important.[7] There is no standard definition of fidelity for laparoscopic simulation and no accepted classification.

Having established the importance of simulation education and the wide range of definitions of fidelity, it must be determined whether fidelity is related to performance. The relationship of simulator fidelity and educational outcomes has been evaluated for over 50 years with conflicting results regarding the connection between learning and fidelity.[4,8,9] Higher fidelity may not translate into more effective training, and lower fidelity simulation may improve training and education.[9] It assumed that the closer a simulator is to the "real

world", the better the transfer of skills to clinical care.[6] It is also assumed that more complex skills require more complex simulators.[6] These assumptions lead to the conclusion that skill transfer with high fidelity (HF) simulators is better than with low fidelity (LF) simulators.[10] Despite the assumptions which may appear to be self-evident, there is little evidence to support them, particularly in regard to procedure skills training.[6,9]

There are few reviews of fidelity directly comparing HF and LF simulation in medical education.[6,11,12] There is a review of only HF simulators[13] and another review of only LF simulators.[14] The lack of a standard definition or classification of simulator fidelity complicates such studies. Norman and colleagues performed a focused review of 24 studies comparing LF and HF simulators for examination skills, procedural skills and scenario management.[6] They included seven studies of procedural skill training, specifically reviewing the association of fidelity and skills transfer and concluded that there is no association. Nguyen and colleagues conducted the only systematic review to date that specifically evaluates simulator fidelity for laparoscopic skills, comparing laparoscopic video trainers with simple box trainers.[11] Based on a meta-analysis of five studies, they conclude that laparoscopic video trainers and simple box trainers are "equally proficient for the acquisition of laparoscopic skills". Munshi et al. conducted a general review of HF versus LF simulators in clinical education and included two studies of procedural skills and conclude that HF is not always superior to LF and that the ideal simulator fidelity depends on the task being simulated.[12] These results are similar to non-medically related simulation studies reported over 50 years ago.[9]

Another area of interest is the relationship of participant experience and simulator fidelity. Less experienced participants may gain significant educational benefit from lower fidelity simulators, while experienced participants may need higher fidelity simulators to realize an educational benefit. This idea was suggested by Alessi4 in studies from the aviation industry as well as in 1989 by Hays and Singer.[8] This relationship has not been well studied in medical education to date.

It is timely to review the available evidence to guide the future development and use of future simulators, particularly in robotic surgery training. We undertook this literature review to evaluate the effect of simulator fidelity on training for procedural skills. This review focuses on the relationship between skills transfer and simulator fidelity for teaching procedural skills and the relationship of skills transfer and fidelity to the level of experience of study participants.

Methods

A literature review methodology of existing studies was used. Only studies which conducted a direct comparison of HF and LF simulation for teaching a particular clinical procedural skill were included. There are no uniformly used criteria by

which to classify simulators as HF or LF. Each author classified the simulators using no specific criteria and judged as LF or HF relatively within each study.

Literature search strategy

A literature review was undertaken according to PRISMA guidelines[15] to examine the effect of simulator fidelity on training in procedure skill training. MEDLINE, OVID and EMBASE databases were searched for articles on simulators that considered simulator fidelity published between January 1990 and January 2019. Search terms included "simulator fidelity and comparison" and "low fidelity" and "high fidelity" and "comparison" and "simulator". Relevant articles from the search were identified by titles and abstracts. The full paper was then assessed. Reference lists from articles identified were also reviewed to identify additional studies not identified by the original search. The PRISMA flow chart is shown in Figure 1.

Study selection and eligibility criteria

All included studies evaluate results based on simulator fidelity as the major outcome. Included studies use randomization to assign participants to LF or HF simulators. Studies with a didactic-only group, or a control group, or with a cross-over design are also included. Single group studies were excluded.

Studies which did not examine simulation of procedural skills were excluded, such as studies which simulated scenario management or learning physical examination. Only studies in the English language were included. Some studies provided extensive training in the simulation while others did not. Some studies included a didactic component for all participants while others did not. There were a variety of outcomes assessment tools used in the studies identified.

Data extraction

The full text of the included studies was reviewed, and data extracted to a standardized data sheet (Excel, Microsoft, Redmond WA). The data extracted is shown in the headings of the Appendix. This instrument for critical appraisal was developed for the purpose of this study. Data abstracted included study design, nature of any interventions, study subjects, outcome measures (subjective, performance, clinical) and results (subjective, performance, clinical). The data abstracted were then reviewed.

Quality appraisal

Data were abstracted by two investigators and the data sheets compared. Any discrepancies regarding conclusions were resolved by discussion. Discussions were held to yield a single document (Appendix) which summarized the 17 studies reviewed and in particular regarding the results of the study with regard to the effect of simulator fidelity on performance.
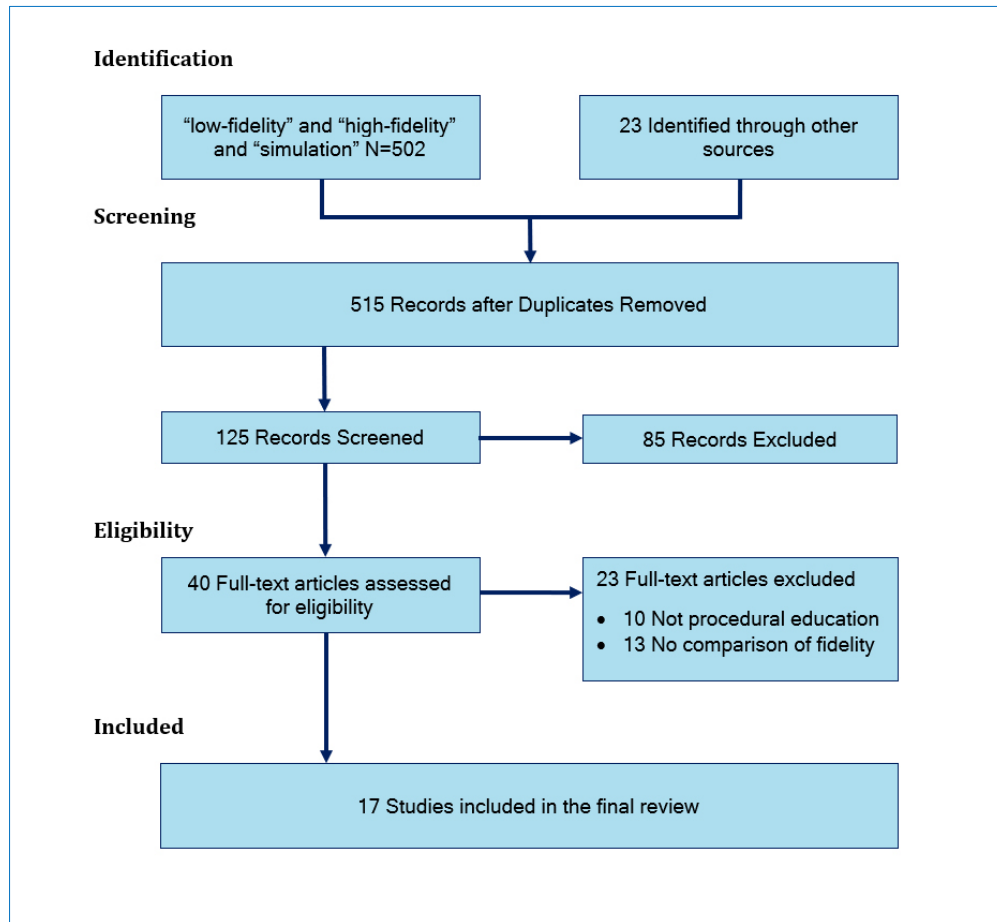
Figure 1. PRISMA chart showing the flow of studies in the review

Results

Study selection

The result of the search for studies comparing performance on LF and HF simulators is shown in Figure 1. Studies which did not include a comparison of performance on HF and LF simulators were excluded. Finally,[17] studies were reviewed in detail and included in this review.

The classification of fidelity used in these studies is complicated by the lack of a standard classification scheme. For example, in one study a video box was classified as LF[18] while in another study the video box was classified as HF.[19] Common categories of laparoscopic simulators are shown in Table 1, in increasing order of fidelity. We propose a classification scheme for laparoscopic simulator fidelity, to allow comparison of simulators labeled as Type 1 through Type 3 (Table 1). The simulators used in the studies reviewed were categorized by Type and then denoted as HF or LF in a consistent manner. All studies compared two types of simulators characterized as HF (higher type number in the classification in Table 1) and LF (lower type). For studies of non-laparoscopic simulator fidelity, we accepted the authors classification of HF and LF. This classification provides a standard description, including the simulation device itself and the constraints of the exercise.

Study characteristics

Studies were reviewed, and data recorded in a standard format. Aggregate review data is shown in the Appendix. All studies included a randomization scheme for participants, and three studies included crossover.[19,20,21] The simulation exercises in all studies were classified by simulator type, surgical constraints and task constraints. In 10 studies, there were two groups compared, with LF and HF simulators and five studies also had a didactic/control group.[22-26] One study had a third group with progressive training using simulators with three levels of fidelity.[3] One study compared three types of simulators.[27]

Procedural skills

Of 17 studies reviewed, eight were focused on laparoscopic surgery skills. There were no studies evaluating robotic skills. The remaining nine studies evaluated simulation training for procedural skills, including trans-bronchial needle aspiration.20, vascular anastomosis,[28] fiberoptic oral intubation,[29] cricothyroidotomy,[30] microvascular anastomosis,[24] endourologic skills (basket stone removal),[22] multiple emergency procedures,[23] intravenous catheter placement[3] and phlebotomy.[31]

Evaluations

Of 17 studies reviewed, 13 evaluated participants one time and four studies evaluated participants both before and after training. One study used participant surveys only for evaluation.[20] All other studies used a combination of global rating scores, time measurements, checklists and subjective evaluations by expert raters which included evaluation of recorded videos.

Table 1. Classification of laparoscopic surgery simulators

Simulator Type	Description
1	Simple box (e.g. plain cardboard box) with/without a mirror with a webcam or no video
2	Box trainer: box utilizing a camera/laparoscope display light source and instruments used for laparoscopic training (e.g. Fundamentals of Laparoscopic Surgery system)
2s	Box trainer (same as Type 2) with motion sensors to measure distance and direction moved in order to calculate the economy of movement
3	Virtual reality (VR) simulators: Display-based computer software and hardware similar to that used in laparoscopic surgery

Simulator Fidelity and laparoscopic procedure skill training

This review includes eight studies designed to compare HF and LF simulators for laparoscopic surgery skill training. Simulators were classified on the scheme proposed here, Type 1 through Type 3. The classification of laparoscopy skill simulators is shown in Table 1. Five of eight studies compared Type 1 and Type 2 simulators. Five of eight studies compared two groups. Three of the studies had three groups (control or another simulator)[3,27,26] and two studies had a crossover design.[19,21]

Effect of simulator fidelity

Of nine studies of non-laparoscopic procedure training, 7/9 showed that training with the HF simulator was not better than training with a LF simulator. Review of laparoscopic surgery skill training studies showed that 8/8 concluded that training with an HF simulator did not have outcomes superior to training with LF simulators. Five studies included HF and LF groups as well as a didactic or control group.[22-26] These five studies found similar outcomes with LF and HF simulators and that either LF or HF had better outcomes than a control group with no intervention or a group that received didactic training only.

The two studies that concluded that participants trained with HF simulators have better performance than those trained on LF simulators included a study of intravenous catheter insertion[3] and a study of vascular anastomosis skills.[28] In the vascular anastomosis study, time and global rating scores were similar for the LF and HF groups, but the HF group received a higher score on the overall performance in the animal surgery component of the study.

This review includes studies of a wide range of procedure skill training studies and shows that 15/17 studies have the similar conclusion, that training with an HF simulator does not have outcomes superior to training with an LF simulator.

Effect of level of training

Of 17 studies reviewed, 14/17 used study participants at the same level of training so that no conclusions can be made regarding a relationship between fidelity, level of training and effect on performance. One study had participants at three levels (medical students, residents and faculty) but did not compare the effect of level of training on performance.[21] One study included fourth year (USA) medical students and first-year residents and considered all participants in aggregate.[18]

One (1/17) study evaluated the performance of junior and senior surgery residents separately on HF and LF simulators.[28] The authors concluded that skill acquisition was significantly affected by both simulator fidelity and level of training as measured by checklist scores and final product analysis on the animal surgery model. The global rating scores for junior and senior residents were similar for residents at both levels trained on both LF and HF simulators. The authors found that both junior and senior residents had better skill transfer from the bench model to the live animal model after practicing on HF simulators compared to LF simulators.

Clinical outcomes evaluation

Of the 17 studies, 3/17 followed training with the evaluation of the skill on a patient (N=1) or a live animal surgery laboratory (N=2). The study with evaluation on a patient was for oral intubation in addition to laboratory performance on simulators, participants were evaluated by success rate, first-time success rate and time.[29] Animal surgery included microsurgery[24] and a vascular anastomosis.[28] The microsurgery study evaluated final performance based on patency of the vas deferens. The vascular anastomosis study evaluated the quality of the vascular anastomosis based on an expert rater. Three studies followed training with the evaluation of a procedure on ex-vivo tissue.[20,23,26]

Cost

The cost of the simulators used in the studies was stated in 4/17 studies. The LF simulator used for urology skills cost CDN $20, while the HF simulator cost CDN $3700.[22] The cardboard box used for laparoscopic skills training cost €0, while the HF simulator was quoted as €30,000.[32] Another laparoscopic skill study used three simulators including a simple webcam (USD $100), a mirror box (USD $300) and a video system (USD $2095).[27] A laparoscopic skill study used an HF video trainer at the cost of $31,435 compared with a commercially produced camera-less box trainer which cost $185.[33]

Discussion

This review of simulator fidelity and procedural skill training includes 17 studies of which eight were comparisons of simulator fidelity for teaching laparoscopic surgery skills. Overall, with some minor anomalies, the results of 15/17 studies reviewed are consistent in that while HF simulation and LF simulation are both effective for improving performance compared to no training or didactic training alone, HF simulators did not result in significantly improved performance compared to LF simulators. This was generally true across teaching a wide range of procedural skills. Two studies found improved performance with HF compared to LF simulation.[3,28]

The effect of simulator fidelity on performance has been studied for many years. Over 65 years ago, it was suggested that HF simulators do not result in improved skills transfer compared to LF simulators.[4,8] These studies were not based on medical education but included data from a range of other fields such as aviation and manufacturing. Simulation and considerations of fidelity in medical education should be considered separately for examination skills, scenario management and procedural skills.

One of the problems in this type of comparison is a lack of a standard classification scheme. The LF simulator in one study is the HF simulator in another. Therefore, we propose a system for classifying simulation exercises to classify the simulator type (Table 1). All of the studies in the Appendix were classified in this manner to facilitate comparisons.

Data regarding the impact of simulator fidelity for scenario training is different than that for procedure training. In a study of performance in Advanced Cardiac Life Support, investigators concluded that those trained on HF simulators performed better.[34] These results are not consistently reported, with a systematic review of scenario management simulation showing no difference in performance for participants trained on LF and HF simulators.[6] In another scenario-based training simulation, 102 medical students were trained using LF or HF simulators and investigators that the students trained on HF simulators were over-confident and had worse performance than those trained on LF simulators.[35]

Data for procedure skill training does not generally show a benefit from training on HF simulators. The availability of low-cost computing power has led to a range of HF simulators for laparoscopic training many of which are based on virtual reality (VR) technology. Despite their widespread use, VR trainers have been shown to be associated with improved skills transfer to the operating room only for basic skills and in laparoscopic cholecystectomy, and not for advanced procedures.[36] In a systematic review of eight trials, six of which compared VR training to no training, VR was associated with improved operating time and performance, but an effect on clinical outcomes and cost has not been shown.[37] In a meta-analysis of nine studies comparing VR trainers to box trainers, there was no significant difference in performance.[38]

Despite the fact that LF trainers can result in performance improvements similar to HF trainers, VR systems continue to be emphasized. A European consensus statement includes VR trainers as part of a proficiency-based training program.[39]

There have been several approaches to studying the effects of simulation training on participants at different experience levels. LF simulation in endovascular procedures was shown to improve motivation in novice trainees but does not necessarily improve practical skills.[40] In one study of psychological stress, junior residents had significantly increased stress during HF simulation compared to LF simulation, while senior surgeons showed no differences.[41] It was also suggested that HF simulators might result in poorer performance by novice trainees because they are over-stimulated by the HF environment.[9] A meta-analysis of VR trainers was not able to conclude about the effect of level of training on effectiveness of VR simulators compared to box trainers.[38] In one study reviewed here, medical students with no prior experience performed better with a LF simulator than a HF simulator.[22] Existing data are insufficient to draw a conclusion about level of experience and simulator fidelity. In a single study, investigators compared junior and senior resident performance separately and found improved skill transfer after using a HF simulator for vascular surgery.[28] In a systematic review, investigators compared the performance of laparoscopically naïve participants after training with a Type 1 LF simulator (box trainer) and found their performance better than those who had no training.[42] The effect of level of training on the benefits derived from simulators of different fidelity needs further study.

Simulators classified as HF are generally more expensive than LF simulators. Using the most appropriate simulator for the target audience is important for efficient use of resources. In a systematic review of low-cost simulators, investigators reviewed 73 unique simulators which cost <£1500, including 60 non-commercial and 13 commercially available simulators.[14] They reported that commercial simulators ranged from £60 to £1007 and non-commercial simulators ranged from £3 to £216. Of the 17 studies reviewed here, the cost was mentioned in just three studies. The LF endourology model cost CDN \$20 and the HF model cost CDN \$3700.[22] In a study of three laparoscopic simulators, the Type 2 HF video system cost USD \$2095, the Type 1 mirror box cost USD \$300 and the Type 1 webcam system was USD \$100.[23] In another study of laparoscopic trainers, the conventional Type 2 video system cost €30,000, while a no-cost Type 1 cardboard box provided equal training benefits. The general lack of cost information reporting has been noted in other reviews.[38] These results stress the importance of reporting costs in future studies of simulator training to allow comparisons of this important parameter.

The ultimate goal of simulation training is to improve clinical outcomes for patients. This remains difficult to prove. In this review, 3/17 studies followed simulation training with a patient intervention or a live animal laboratory. In

a study of VR training versus a control group with a total of 13 residents, trainees who completed VR training had improved results in their first ten laparoscopic cholecystectomy procedures (fewer errors, faster times) compared to the control group.[43] Future studies should include clinical exercises after simulation training when possible.

The results of this literature review have implications for procedure skill training in the future. It is evident from numerous studies that training with a simulator (HF or LF) results in improved performance compared to no training at all or didactic training alone.[22-26,42] All but two of the 17 studies reviewed did not find that HF simulators (also generally more expensive than LF simulators) result in the improved acquisition of procedural skills compared to LF simulators. This could significantly affect budgets for simulation education since it may not be necessary to always use a HF simulator.

Few studies have examined the effect of participant training level on simulation training. It may be necessary to use HF simulators to train people with more experience while LF simulators may be perfectly adequate for those with little or no experience. The paucity of data in the literature makes it difficult to draw conclusions and supports future studies of this issue which also has implications for the cost of training. More studies are needed with participants having varied levels of experience to determine who will benefit most from different types of simulators. The need to provide objective feedback to trainees is emphasized by a number of studies and should be part of future studies. Simulator fidelity is another parameter to be considered in the design of simulation curricula as well as future trials.

Training in robotic surgery is emphasized in contrast to what happened when laparoscopic surgery was introduced.[2] Simulators designed for robotic-assisted surgery training are generally high-fidelity, large and expensive because they closely resemble the actual daVinci robotic-assisted surgical system robot (Intuitive, Sunnyvale CA) which is the most commonly used system. A reasonably low-cost portable VR trainer for the daVinci system was recently described.[44] The Fundamentals of Robotic Surgery course has been developed and validated and is becoming more widespread.[45] The FIRST exercises have also been validated.[46] However, a meta-analysis of 107 studies concluded that there is no universally accepted method of robotic surgery skill assessment.[47] In a systematic review of robot-assisted surgery VR simulators, investigators concluded that it is not clear which exercises and metrics can distinguish levels of training in performance on the daVinci system.[48] Further study is essential as training in robot-assisted surgery becomes more widespread.

There are acknowledged limitations of this review. First, there is some heterogeneity of the included studies, particularly regarding methods of evaluation. The included studies cover a wide range of procedural skill training using various methods for evaluation of outcomes. Comparisons of studies are somewhat limited by lack of a standard scheme for classification of simulators, and comparisons in the included studies are based on relative fidelity. There is a wide range of simulators compared, which makes comparisons difficult. In addition, the training schemes used are widely variable. Future studies would benefit from a unified scoring system for the assessment of outcomes.

Conclusions

Despite some heterogeneity of the 17 included studies, there are several important conclusions from this review. Of the studies evaluated, 15/17 show that LF simulation results in similar outcomes to HF simulation. LF simulators may provide significant educational benefits for less experienced trainees while HF simulators may be of benefit to more experienced trainees. These results may lead to cost savings in medical education since simulators that are generally less expensive provide similar training to more expensive simulators. We suggest that future studies of laparoscopic simulation education adopt the classification scheme introduced here to facilitate future comparisons. A lack of uniform criteria for evaluating procedural skills remains a formidable barrier to determining whether or not simulation training results in improved clinical performance, which is the ultimate goal of simulation education and requires further study.

Conflicts of Interest

The authors declare that they have no conflict of interest.

References

1. Hanks JB. Simulation in surgical education: influences of and opportunities for the Southern Surgical Association. J Am Coll Surg. 2019;228(4):317–328.
2. Lefor AK. Robotic and laparoscopic surgery of the pancreas: a historical review. BMC Biomedical Engineering. 2019; 1(1):2-10.
3. Brydges R, Carnahan H, Rose D, Rose L, Dubrowski A. Coordinating progressive levels of simulation fidelity to maximize educational benefit. Acad Med. 2010;85(5):806–812.
4. Alessi SM. Fidelity in the design of instructional simulations. Journal of Computer Based Instruction 1988; 15 (2):40–47.
5. Liu D, Macchiarella N, Vincenzi DA. Simulation Fidelity in Human factors in simulation and training. New York: Taylor & Francis; 2001.
6. Norman G, Dore K, Grierson L. The minimal relationship between simulation fidelity and transfer of learning. Med Educ. 2012;46(7):636–647.
7. Tun JK, Alinier G, Tang J, Kneebone RL. Redefining simulation fidelity for healthcare education. Simulation & Gaming. 2015. 46 (2): 159-174.
8. Hays RT, Singer MJ. Simulation Fidelity in Training System Design, New York: Springer;1989.
9. Feinstein AH, Cannon HM. Fidelity, verifiability and validity of simulation: constructs for evaluation. Developments in Business and Experiential Learning. 2001; 28:57-75.
10. Doherty-Restrepo J, Tivener K. Current literature summary: review of high-fidelity simulation in professional education. Athletic Training Education Journal. 2014; 9(4):190-195.
11. Nguyen T, Braga LH, Hoogenes J, Matsumoto ED. Commercial video laparoscopic trainers versus less expensive, simple laparoscopic trainers: a systematic review and meta-analysis. J Urol. 2013;190(3):894-899.
12. Munshi F, Lababidi H, Alyousef S. Low versus high fidelity simulations in teaching and assessing clinical skills. Journal of Taibah University Medical Sciences. 2015;10(1):12-5.

13. Issenberg SB, McGaghie WC, Petrusa ER, Lee Gordon D, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach. 2005;27(1):10-28.

14. Li MM, George J. A systematic review of low-cost laparoscopic simulators. Surg Endosc. 2017;31(1):38-48.

15. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. 2015;4(1):1.

16. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg. 1997;84(2):273–278.

17. Vassiliou MC, Feldman LS, Andrew CG, Bergman S, Leffondré K, Stanbridge D, et al. A global assessment tool for evaluation of intraoperative laparoscopic skills. Am J Surg. 2005;190(1):107-113.

18. Diesen DL, Erhunmwunsee L, Bennett KM, Ben-David K, Yurcisin B, Ceppa EP, et al. Effectiveness of laparoscopic computer simulator versus usage of box trainer for endoscopic surgery training of novices. J Surg Educ. 2011;68(4):282-289.

19. Keyser EJ, Derossis AM, Antoniuk M, Sigman HH, Fried GM. A simplified simulator for the training and evaluation of laparoscopic skills. Surg Endosc. 2000;14(2):149-153.

20. Davoudi M, Wahidi MM, Zamanian Rohani N, Colt HG. Comparative effectiveness of low- and high-fidelity bronchoscopy simulation for training in conventional transbronchial needle aspiration and user preferences. Respiration. 2010;80(4):327-334.

21. Tan SC, Marlow N, Field J, Altree M, Babidge W, Hewett P, et al. A randomized crossover trial examining low- versus high-fidelity simulation in basic laparoscopic skills training. Surg Endosc. 2012;26(11):3207-3214.

22. Matsumoto ED, Hamstra SJ, Radomski SB, Cusimano MD. The effect of bench model fidelity on endourological skills: a randomized controlled study. J Urol. 2002;167(3):1243-1247.

23. Anastakis DJ, Regehr G, Reznick RK, Cusimano M, Murnaghan J, Brown M, et al. Assessment of technical skills transfer from the bench training model to the human model. Am J Surg. 1999;177(2):167-170.

24. Grober ED, Hamstra SJ, Wanzel KR, Reznick RK, Matsumoto ED, Sidhu RS, et al. The educational impact of bench model fidelity on the acquisition of technical skill: the use of clinically relevant outcome measures. Ann Surg. 2004;240(2):374-381.

25. Munz Y, Kumar BD, Moorthy K, Bann S, Darzi A. Laparoscopic virtual reality and box trainers: is one superior to the other? Surg Endosc. 2004;18(3):485-494.

26. Bruynzeel H, de Bruin AF, Bonjer HJ, Lange JF, Hop WC, Ayodeji ID, et al. Desktop simulator: key to universal training? Surg Endosc. 2007;21(9):1637-1640.

27. Chung SY, Landsittel D, Chon CH, Ng CS, Fuchs GJ. Laparoscopic skills training using a webcam trainer. J Urol. 2005;173(1):180-183.

28. Sidhu RS, Park J, Brydges R, MacRae HM, Dubrowski A. A Laboratory-based vascular anastomosis training: a randomized controlled trial evaluating the effects of bench model fidelity and level of training on skill acquisition. J Vasc Surg. 2007;45(2):343-349.

29. Chandra DB, Savoldelli GL, Joo HS, Weiss ID, Naik VN. Fiberoptic oral intubation: the effect of model fidelity on training for transfer to patient care. Anesthesiology. 2008;109(6):1007-1013.

30. Katayama A, Nakazawa H, Tokumine J, Lefor AK, Watanabe K, Asao T, et al. A high-fidelity simulator for needle cricothyroidotomy training is not associated with increased proficiency compared with conventional simulators: a randomized controlled study. Medicine (Baltimore). 2019;98(8): e14665.

31. Scerbo MW, Schmidt EA, Bliss JP. Comparison of a virtual reality simulator and simulated limbs for phlebotomy training. J Infus Nurs. 2006;29(4):214-224.

32. Chandrasekera SK, Donohue JF, Orley D, Barber NJ, Shah N, Bishai PM, et al. Basic laparoscopic surgical training: examination of a low-cost alternative. Eur Urol. 2006;50(6):1285-1291.

33. Sharpe BA, MacHaidze Z, Ogan K. Randomized comparison of standard laparoscopic trainer to novel, at-home, low-cost, camera-less laparoscopic trainer. Urology. 2005;66(1):50-54.

34. Conlon LW, Rodgers DL, Shofer FS, Lipschik GY Impact of levels of simulation fidelity on training of interns in ACLS. Hosp Pract (1995). 2014;42(4):135-141.

35. Massoth C, Röder H, Ohlenburg H, Hessler M, Zarbock A, Pöpping DM, et al. High-fidelity is not superior to low-fidelity simulation but leads to overconfidence in medical students. BMC Med Educ. 2019;19(1):29.

36. Yiannakopoulou E, Nikiteas N, Perrea D, Tsigris C. Virtual reality simulators and training in laparoscopic surgery. Int J Surg. 2015; 13:60-64.

37. Nagendran M, Gurusamy KS, Aggarwal R, Loizidou M, Davidson BR. Virtual reality training for surgical trainees in laparoscopic surgery. Cochrane Database Syst Rev. 2013;(8): CD006575.

38. Alaker M, Wynn GR, Arulampalam T. Virtual reality training in laparoscopic surgery: A systematic review & meta-analysis. Int J Surg. 2016; 29:85-94.

39. van Dongen KW, Ahlberg G, Bonavina L, Carter FJ, Grantcharov TP, Hyltander A, et al. European consensus on a competency-based virtual reality training program for basic endoscopic surgical psychomotor skills. Surg Endosc. 2011;25(1):166-171.

40. Aeckersberg G, Gkremoutis A, Schmitz-Rixen T, Kaiser E. The relevance of low-fidelity virtual reality simulators compared with other learning methods in basic endovascular skills training. J Vasc Surg. 2019;69(1):227-235.

41. Bakhsh A, Martin GFJ, Bicknell CD, Pettengell C, Riga C. An evaluation of the impact of high-fidelity endovascular simulation on surgeon stress and technical performance. J Surg Educ. 2019;76(3):864-871.

42. Nagendran M, Toon CD, Davidson BR, Gurusamy KS. Laparoscopic surgical box model training for surgical trainees with no prior laparoscopic experience. Cochrane Database Syst Rev. 2014. 7;(1):CD010479.

43. Ahlberg G, Enochsson L, Gallagher AG, Hedman L, Hogman C, McClusky DA, et al. Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. Am J Surg. 2007;193(6):797-804.

44. F Ferro M, Brunori D, Magistri F, Saiella L, Selvaggio M, Fontanelli GA. A portable da Vinci simulator in virtual reality. Third IEEE International Conference on Robotic Computing (IRC); 25-27 Feb 2019; Naples, Italy: IEEE; 2019.

45. Smith R, Patel V, Satava R. Fundamentals of robotic surgery: a course of basic robotic surgery skills based upon a 14-society consensus template of outcomes measures and curriculum development. Int J Med Robot. 2014;10(3):379-384.

46. Goh AC, Aghazadeh MA, Mercado MA, Hung AJ, Pan MM, Desai MM, et al. Multi-institutional validation of fundamental inanimate robotic skills tasks. J Urol. 2015;194(6):1751-1756.

47. Chen J, Cheng N, Cacciamani G, Oh P, Lin-Brande M, Remulla D, et al. Objective assessment of robotic surgical technical skill: a systematic review. J Urol. 2019;201(3):461-469.

48. 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(6):1065-1080.

Appendix

Seventeen studies of the effect of simulator fidelity on procedure skill training

Ref	Year	Procedural Skill	Study Design / Intervention / Exercise Classification	Study Subjects	Outcome Measures/Evaluations	Clinical Evaluation*	Results: Opinion, Performance, Clinical
				Non-Laparoscopic Skills			
20	2009	Trans-bronchial needle aspiration	-2 groups, Randomized Crossover: LF (rubber tube) and HF (VR system) -All participants trained on both LF and HF	44 practicing physicians	-Opinion: Participant Likert-scale surveys completed after training on both types (LF/HF) and then a survey after completing both -Performance: Not evaluated -Clinical: Not evaluated	No	LF preferred -Opinion: LF preferred 19/44 vs HF 11/44. LF was judged more realistic 23/44 vs 17/44 for HF -Performance: Not evaluated -Clinical: Not evaluated
3	2010	IV insertion	-Randomized to LF (Virtual IV, Laerdal), Mid-fidelity (plastic arm), HF (SimMan, Laerdal) or progressive practice. -After practice on LF or HF, or with progressive from LF to medium to HF, tested once on a standard simulator	45 medical students	-Opinion: Not evaluated -Performance: Blinded raters scored Integrated Procedural Performance Instrument Rating, Global Rating Scale, Checklist and Communications -Clinical: Not evaluated	No	Progressive from LF-Medium-HF is best, HF better than LF -Opinion: Not evaluated -Performance: Progressive from LF to medium fidelity to HF better than HF (effect size 0.78); HF better than LF (effect size 0.72). Progressive group had more practice time overall, but less on the HF -Clinical: Not evaluated
29	2008	Fiberoptic oral Intubation	-Randomized to LF (bronchoscope, simple bench model) or HF (VR, Accutouch Flex bronchoscopy simulator) simulators followed by -Clinical skill test on patient	28 respiratory therapists	-Opinion: Not evaluated -Performance: Objective checklist, time, success rate and GRS -Clinical: GRS, Success rate, time, first-attempt rate	Yes	LF=HF -Opinion: Not evaluated -Performance: GRS, checklist, LF=HF -Clinical: GRS, checklist scores same LF=HF, success rates same, both LF and HF showed equal improvement first to second attempts
22	2002	Endourology-stone removal	-Randomized 3 arm: didactic, LF (self-designed) and HF (VR Model, Limbs & Things, Bristol UK) -One evaluation after LF/HF or didactic. -LF model CDN $20, HF model CDN $3700	40 medical students	-Opinion: Not evaluated -Performance: GRS, checklist and time -Clinical: Not evaluated	No	LF=HF, HF and LF both superior to didactic -Opinion: Not evaluated -Performance: GRS LF 79% vs. HF 88% (p=0.08), Checklist LF 90% vs. HF 94% (p=0.17), overall pass rating p=0.95 -Clinical: Not evaluated
23	1999	Six Procedures: burr hole, chest tube, bowel anastomosis, wound closure, tendon repair, K-wire insertion	-Randomized 3 arms: didactic, HF (cadaver), or LF (dry model) -One-time evaluation on cadaver model of each group for all 6 procedures	23 residents	-Opinion: Not evaluated -Performance: GRS and checklist -Clinical: Not evaluated	No	LF=HF, HF and LF both superior to didactic -Opinion: Not evaluated -Performance: GRS LF 64%, HF 67%, p>.05; Checklist LF 68%, HF 69%, p>.05 -Clinical: Not evaluated
24	2004	Microsurgery	-Randomized to 3 groups for training: didactic, LF (silicone tube) and HF (rat vas deferens). -All participants evaluated on both LF and HF -Animal surgery by all participants. One-time evaluation	50 residents	-Opinion: Not evaluated -Performance: GRS, checklist, time -Clinical: patency of vas deferens anastomosis in a rat model	Yes	LF=HF, HF and LF both superior to didactic -Opinion: Not evaluated -Performance: GRS LF 64%, HF 70%, checklist LF 84%, HF 89%, p>.05 -Clinical: Anastomotic patency same LF=HF groups, both better than didactic

28	2007	Vascular Anasto-mosis	-Randomized to HF (cadaver) or LF (plastic model) 3hr simulator training either LF or HF. -Animal surgery after training -One-time evaluation after animal surgery. Analyzed junior and senior residents separately	27 Surgery residents	-Opinion: Not evaluated -Simulator Performance: Not evaluated -Clinical: Checklist, GRS, time, hand motion, final product	Yes	GRS LF=HF for Junior and Senior residents. HF final product better than LF for Junior and Senior residents. -Opinion: Not evaluated -Simulator Performance: Not evaluated -Clinical: Final product analysis: Junior LF (3.1), HF (4.0), p<.05, Senior LF (3.6), HF (5.0), p<.05. GRS scores LF=HF both junior and senior
30	2018	Cricothyroidotomy	-Randomized to HF (3D printed larynx) or LF (simple tube) -Lecture for all then training on LF or HF, 10 times then *ex-vivo* -porcine larynx	52 residents	-Opinion: similarity of simulator to porcine larynx -Performance: success rate, time using porcine larynx -Clinical: Not evaluated	No	LF=HF for opinion and performance -Opinion: Both groups reported similar appearance to porcine larynx (LF 3/5, HF 4/5, p=0.81). -Performance: HF and LF no difference. Success rate (LF 52%, HF 48%, p>.05), time (45 vs 49sec, p>.05) -Clinical: Not evaluated
31	2006	Phlebotomy	-Randomized to HF (VR) or LF (simulated arm) -Before/After evaluation	45 third year medical students	-Opinion: Not evaluated -Performance: Checklist, pretest, post-test. Compared before/after scores for each training method -Clinical: Not evaluated	No	LF scores higher than HF -Opinion: Not evaluated -Performance: Post-test scores higher for LF compared to HF -Clinical: Not evaluated

Laparoscopic Skills

24	2004	Laparoscopic skills-cutting and clip application	-Randomized to 3 groups: control, HF (Type 3, box, cutting), -LF (Type 2s, box, cutting). Before/ after evaluation	24 junior residents	-Opinion: Not evaluated -Performance: Motion data (both LF and HF), time on pre-test, then training (3 groups), then reassessment -Clinical: Not evaluated	No	LF=HF, HF and LF both superior to controls -Opinion: Not evaluated -Performance: Motion, time LF=HF, both LF and HF better than control -Clinical: Not evaluated
18	2011	Five laparoscopic skills: camera handling, knot tying, scope tracking, clipping, needle transfer	-Randomized to HF (Type 3, box, pegs/suture/cutting) simulator or LF (Type 2, box, pegs/suture/cutting). -6-months of training -Evaluation on porcine model (live tissue) at 0,2,6 months	11 PGY-1 residents and 12 fourth-year medical students	-Opinion: Not evaluated -Performance: checklist, video review, porcine surgery to test the skills -Clinical: Not evaluated	Yes	LF=HF, students and residents improved -Opinion: Not evaluated -Performance: Not evaluated on simulators -Clinical: Porcine model, both LF and HF groups showed improved scores in all 5 skills from 0 to 2 to 6 months, LF vs HF p>.05
19	1999	Seven basic laparoscopic skills	-Randomized to LF (Type 1, box, pegs/suture/cutting)) or HF (Type 2, box, pegs/suture/cutting) -Crossover -One-time evaluation on both LF and HF	22 surgery residents	-Opinion: Not evaluated -Performance: Precision and speed scores -Clinical: Not evaluated	No	LF (1,499 ± 237) = HF (1,209 ± 435, p>.05) -Opinion: Not evaluated -Performance: Scores correlated with level of training and the mirrored box scores differentiate junior and senior residents -Clinical: Not evaluated
33	2005	Four basic laparoscopic skills	-Pre-test on HF (Type 2, box, blocks/suture/cut) -Then randomized to 5 hours training on LF (Type 1, box, blocks/suture/cut) or HF (Type 2, box, blocks/suture/cut) -Post-test on HF simulator	22 Urology residents	-Opinion: asked if simple trainer was effective at post-test -Performance: GRS and time/efficiency -Clinical: Not evaluated	No	LF=HF for improvement of scores at post-test -Opinion: 91% felt that the simple trainer was effective -Performance: All participants improved at post-test compared to pre-test, p<.05. GRS for LF (43,.91) and HF (55.87, p>.05). Time and efficiency for all 4 skills improved by the same amount for LF and HF. -Clinical: Not evaluated
27	2005	Peg transfer and pattern cutting	-Randomized to 3 groups -Compared 3 simulators: LF (Type 1, box, peg/cutting compared to a commercial LF (Type 1, (box, peg/cutting) and HF (Type 2, box, peg/cutting)	42 laparoscopically naïve residents	-Opinion: Not evaluated -Performance: Time and performance scores -Clinical: Not evaluated	No	LF (webcam) same as HF. Either of these better than LF (2-mirror) -Opinion: Not evaluated -Performance: Time compared, and order of simulators did not make a difference

			-Six groups tested on all 3 simulators used in a different order				-Clinical: Not evaluated
26	2007	Laparoscopic suturing	-Randomized to 3 groups: control, LF (Type 1, box, knot) and HF (Type 2, box, knot) -Video review of all suturing -Final evaluation on cadaveric animal model	30 medical students	-Opinion: Not evaluated -Performance: time for suturing and review of videos -Clinical: Not evaluated	No	LF=HF, either is better than control -Opinion: Not evaluated -Performance: mean times: control (12min), LF (6.6min), HF (6.6min) -Clinical: Not evaluated
32	2006	Laparoscopic Skills: Sugar cube transfer, mint transfer and disk cut-out	-Randomized to 24h (8 sessions) training on either HF (Type 2, box, pegs/cutting) or LF (Type 1, box, pegs/cutting) -Evaluated after training one time on both simulators	36 medical students	One-time test on both simulators -Opinion: Not evaluated -Performance: Accuracy, time, depth perception assessed subjectively. -Clinical: Not evaluated	No	LF=HF for all tasks tested -Opinion: Not evaluated -Performance: Subjective scores for all tasks by blinded rater were similar for all 3 tasks for both groups (p>.05). Times were similar (p>.05). LF group had better accuracy score on disk cut-out than HF. LF trained participants were faster when tested on the HF simulator. -Clinical: Not evaluated
21	2012	Laparoscopic skills: skill training was different for LF and HF	-Randomized with crossover: LF (Type 2, box, pegs/cutting/suturing) and HF (Type 3, box, pegs/cutting/suturing) -After achieving minimal proficiency on LF or HF, crossed over to the other simulator and tested	228 medical students, residents and senior physicians	-Opinion: Not evaluated -Performance: Scores on FLS box and LapSim VR. Baseline scores compared with scores after training on the other simulator -Clinical: Not evaluated	No	LF-trained participants did better on HF than HF-trained participants did on LF -Opinion: Not evaluated -Performance: LapSim trained participants had a 20% task pass rate while FLS to LapSim trained participants had a 29% pass rate (p<.01). Score increase similar (p>.05) after training on either device. -Clinical: Not evaluated

LF: Low fidelity; HF: High fidelity; FLS: Fundamentals of Laparoscopic Surgery; LapVR: CAE Healthcare (Sarasota FL USA); LapSim: Surgical Science (Gotheborg, Sweden); VR: Virtual Reality
*Clinical evaluation indicates that skill was tested on a patient or live animal after training; GRS: Global Rating Scale