Training and transfer effect of FluoroSim, an augmented reality fluoroscopic simulator for dynamic hip screw guide-wire insertion: a single-blinded randomized controlled trial

Abstract

Background

FluoroSim, a novel fluoroscopic simulator, can be used to practise dynamic hip screw (DHS) guide-wire insertion in a high fidelity clinical scenario. We aim to demonstrate a training effect in **medical students** naïve to this operation and simulation.

Methods

45 **medical students** were recruited and randomized to either training (n=23) and control (n=22) cohorts. The training cohort had more exposure to FluoroSim (five attempts each week) compared to the control cohort (single attempt each week) over a two-week period and a one-week washout period in between. Five real-time objective performance metrics were recorded including (i) tip-apex distance (mm), (ii) **predicted** cut-out rate (%), (iii) total procedural time (s), (iv) total number of radiographs (n), and (v) total number of guide-wire retries (n).

Results

At baseline, there was no significant difference in the performance metrics which confirmed absence of a selection bias. Intra-group training effect demonstrated a significant improvement in all metrics for the training cohort only. A **significant difference between groups** was demonstrated as the training cohort significantly outperformed the control cohort in three metrics (procedural time [25%], number of radiographs [57%] and number of guide-wire retries [100%]; $p<0.001$). A learning curve showed an inversely proportional correlation between frequency of attempts and procedural time as well as the number of digital fluoroscopic radiographs used, indicating development of psychomotor skills. There was also an improved baseline of the learning curve after a one-week washout period suggesting skills retention.

Conclusion

Skills acquisition on FluoroSim was demonstrated with repeat exposure in a safe, radiation-free and a high-fidelity clinical simulation with actual theatre equipment. The task of DHS guide wire insertion requires cognitive and psychomotor skills which take a variable number of attempts to acquire as demonstrated on the learning curve. Further work is required to demonstrate that the skill tested by FluoroSim is the same skill required for intraoperative DHS guide-wire insertion. However, FluoroSim signifies an improvement in extra-clinical training opportunities for orthopedic trainees.

Level of evidence

Level 1

Introduction

Hip fractures: clinical burden

Hip fractures affect 70,000 people per annum in the UK,[1] costing the National Health Service an estimated 1.4 billion pounds[2] with one-year mortality rate of 33%.[3] Extracapsular hip fractures are commonly treated using open reduction and internal fixation with the dynamic hip screw (DHS).[4]

Technical skills in DHS surgery

The tip-apex distance (TAD) described by Baumgaertner is a measure of the position of the tip of the DHS guide wire/screw in relation to the apex of the femoral head.[5] Failure of the DHS implant is predicted by the TAD, deemed to be the only clinically validated outcome.[6] The likelihood of implant failure, known as cut-out, is greatly reduced at an optimal TAD below 25mm.[7] Surgical trainees need to be educated **about** the significance of TAD to prevent cut-out.

Current training of DHS surgery

DHS surgery is learnt in theatre under fluoroscopic guidance. Due to radiation risks to the trainee and trainer,[8-10] **fluoroscopy should not be practised during simulation**. These risks are amplified as junior trainees use more imaging than their senior colleagues.[11] Other digital fluoroscopy systems have been developed, but they are not utilized within surgical education.[12]

Training surgeons leads to a reduction in theatre efficiency and increases patient risk.[13,14] These issues are juxtaposed with a 50-80% reduction in training time for trainee orthopedic surgeons due to the European Working Time Directive.[15-17] A reduced amount of training time has been perceived by European and North American surgeons to have had a negative effect on their education.[15, 18]

Simulation: traditional training adjunct

Surgical simulation provides both a safe and constructive platform for trainees to develop their technical skills away from the theatre.[19] Patient safety may be prioritized; trainees may be able to curb their learning curve in a simulation setting, improving their technical skills away from the patient.

Current simulation options for DHS surgery include virtual reality (VR) and dry-bone models. **Workshop femurs, eponymously named as saw bones,** are the mainstay for fracture fixation simulation.[20] They allow the user to practise a task with actual surgical equipment therefore developing the **required psychomotor** skills.

VR simulation for fracture fixation has been validated. Bonedoc is a digital software that allows a user to go through the cognitive steps of the DHS operation with digital fluoroscopy, inputting motor data using a computer mouse.[21] TraumaVision (Swemac Simulation AB, Linkoping, Sweden) is a VR system, similar to Bonedoc, with the advantage of a haptics enabled stylus pen.[22] The motor movements of the operation are inputted into the software using the phantom stylus pen which responds with haptic feedback. All actions are done with the pen which is not comparable to actual surgical equipment. This lowers the simulation fidelity. This simulation system costs tens of thousands of USD excluding maintenance and system upgrades.

Aims and Objectives

This study presents the training effect observed with a novel fluoroscopic simulator, FluoroSim.[23] FluoroSim is a digital fluoroscopy system that can provide imaging and real-time objective performance metrics for workshop bone simulation without radiation. This study aims to demonstrate:

1. exposure to FluoroSim can improve the skill of inserting a DHS guide-wire (i.e. intra-group training effect)
2. participants with a greater exposure to FluoroSim outperform the control cohort measured (i.e. inter-group training effect)
3. skills learnt on FluoroSim are transferable to another validated DHS guide-wire simulator, namely TraumaVision[20,22] (i.e. transfer effect)

Null hypothesis

There was no difference in the performance metrics between the training and the control cohorts.

Methods

Set up

FluoroSim is a novel fluoroscopic simulation software which tracks a DHS guide-wire using two orthogonally placed cameras.[23] The FluoroSim software ran on iMac with OS X El Capitan 10.11.6 (Apple Inc., California, U.S.A), set up in the simulation lab at the Royal National Orthopedic Hospital (RNOH), Stanmore, UK.

The system required calibration with a **workshop femur** (3B Scientific, Hamburg, Germany). A marker was placed at the femoral head apex during the calibration process, representing the center of calibration. A 3D printed jig designed to hold polystyrene drilling blocks was produced. The calibration femur was replaced with the jig. The drilling jig was draped to represent a theatre scenario and all participants were told the approximate location of the greater trochanter to orientate them (figure 1). **All participants performed the procedure with the jig set up to represent the right hip, irrespective of hand dominance.**

Figure 1

A Stryker system 4 rotary drill (Stryker, Michigan, U.S.A.), a guide-wire and a 135-degree angle guide (Innovation Ortho Line Limited, London, England) were used for the simulation. In addition, TraumaVision, a haptics-enabled virtual reality DHS simulator was used to assess skills transfer.

The total cost of FluoroSim was below USD 1,000; largely for hardware whereas the software was developed using open source coding. This compares to the cost of TraumaVision which costs tens of thousands of USD.

Power calculation

Our previous study found that expert trauma and orthopedic surgeons (Attendings) achieved a significantly lower TAD of 22.7mm than novices (Residents).[24] We expect the training cohort to outperform the control group. Using this data, an *a priori* power calculation with power set to 80% and type one error set to 5% was calculated. This study needed 11 participants per cohort.

Objective Performance Metrics

The FluoroSim and TraumaVision software calculated real-time metrics including;

1. TAD (mm)
2. **Predicted cut-out rate (COR, %), calculated from Baumgaertner's curve (Baumgaertner *et al*. 1995)**
3. Total procedural time (s)
4. Total number of radiographs (n)
5. Total number of guide-wire retries (n)

Recruitment and randomization

Undergraduate student doctors from three London universities were recruited on a voluntary basis and gave informed consent to participate. 45 participants were randomized (using an electronic randomized number generator from Microsoft Excel [Microsoft Corporation, Redmond, Washington, USA]) into two cohorts: training (n=23) and control (n=22). A CONSORT diagram can be found in figure 2.

Figure 2

All participants watched a four-minute explanatory video to standardize their understanding, completed a demographic questionnaire and had a single attempt on TraumaVision. The training cohort then had five attempts on FluoroSim in week one. This was repeated in week two after a one week wash-out period. The control cohort completed a single attempt on FluoroSim each week. After each attempt, participants received one minute of feedback on their real-time objective performance metrics. Both cohorts then had a final attempt on TraumaVision at the end of the second week.

All participants were blinded to their allocation and tested in isolation. With each attempt, all five objective performance metrics were recorded from FluoroSim and TraumaVision.

Inclusion/exclusion criteria

Inclusion criteria: undergraduate student doctors naïve to DHS surgery and orthopedic simulation.

Exclusion criteria: previous DHS experience, previous fluoroscopy simulation experience, graduates, and participants who failed to attend both sessions.

Statistics

SPSS (version 24.0, IBM, Armonk, New York) was used for data analysis. Normality of continuous data was assessed with Shapiro-Wilk testing and histograms at $\alpha = 5\%$.

Demographic analysis: the age and the year of study were compared between the cohorts using the independent t-test and the Mann-Whitney U-test respectively. Gender was compared with the Chi squared test. Hand dominance and previous simulator experience was compared with the Fishers exact test.

Baseline outcome metric analysis: comparison between all five objective metrics achieved by each cohort on the first attempt utilized the Mann-Whitney U test.

Intra-group FluoroSim training effect: the first and last attempt were compared within each cohort using the Wilcoxon signed rank test. The difference between the two attempts was calculated, defined as the improvement. Further, the improvement was compared between the cohorts using a Mann-Whitney U test.

Inter-group FluoroSim training effect: the difference in final scores of the objective metrics between the last attempts of the training (tenth) and the control (second) cohorts was analyzed using the Mann Whitney U test.

Transfer effect: the first and last attempts on Trauma Vision were compared within each cohort using the Wilcoxon signed rank test. Further, the improvement was calculated (as above) and compared between the cohorts using the Mann-Whitney U test.
For all statistical tests completed, the type one error was set to 5% to determine statistical significance.

Ethics, funding and potential conflicts of interest
The study protocol was assessed by the Project Evaluation Panel at the Royal National Orthopedic Hospital. It was decided that ethical approval was not required as this was not a

clinical study. All participants gave informed consent prior to entering the study and were aware that they could drop out at any time during the study.

This project was funded by the Professor A.T. Fripp fund. BvD is a clinical fellow funded by NIHR. RAW received a bursary from Goldberg Schachmann and Freda Becker Memorial Fund. There were no conflicts of interest.

Results

Demographics and randomization

14/23 (61%) of the training cohort and 16/22 (73%) of the control cohort were male (p = 0.40). The training group had a significantly higher mean [range] age (21.7 [18-27]) compared to the control cohort (20.3, [18-24]; p = 0.04). Year of study (p = 0.33) and previous simulator experience (p = 0.10) between **both** cohorts **were insignificant**. **23/23 (100%) training participants and 21/22 (95%) control participants were right hand dominant (p = 0.49).**

There was no significant difference in baseline skill for all metrics on the FluoroSim between both cohorts (p-values between 0.16 to 0.38). This reflects the absence of selection bias.

Inter- and Intra-group training effect

A significant intra-group training effect was observed between the first and last attempts for all outcome metrics in the training cohort. No metrics within the control cohort reached statistical significance (Table 1 and Figure 3). When the improvements were compared between the cohorts, all comparisons reached significance (Table 1).

Table 1

A significant inter-group training effect was present for procedural time, number of radiographs and number of guide-wire retries when comparing the final attempts of the control and training cohorts (Table 2 and Figure 3).

Figure 3 a-e

Table 2

Learning curve for training cohort

A learning curve was plotted for the training cohort (Figure 4).

Figure 4

Transfer effect from FluoroSim to TraumaVision

The training cohort demonstrated a significant improvement on TraumaVision for procedural time, number of radiographs and number of guide-wire retries (Table 3). The control cohort showed a statistically significant improvement in procedural time only (Table 3). When the improvement for each metric was compared between the cohorts, no metrics reached significance.

Table 3

Null hypothesis analysis

The null hypothesis was rejected for intra-group training effect and the learning curve for the training cohort, whereas the null hypothesis was partially rejected for inter-group training effect with FluoroSim and transfer effect from FluoroSim to TraumaVision.

Discussion

Main findings

This study has demonstrated a significant intra-group training effect in the training cohort after ten attempts on FluoroSim for each performance metric. The improvement was significantly greater than that demonstrated by the control cohort for each metric. An inter-group training effect was present for the procedural time, number of radiographs taken and number of guide-wire retries only.

Additionally, the training cohort demonstrated a significant improvement on TraumaVision for procedural time, number of radiographs and number of guide-wire retries. For the control group, a significant improvement on TraumaVision was seen for procedural time only. The improvement observed by the training cohort on TraumaVision was not significantly greater than the improvement observed in the control cohort.

FluoroSim training effect

This study followed the methodology used, in part, to validate the VR simulator TraumaVision. Sugand *et al*. demonstrated that after ten attempts using TraumaVision the training cohort achieved a significantly better outcome in all metrics recorded, including TAD and COR, compared to **the** control **cohort**.[22] This current study saw no significant inter-group training effect with the two aforementioned metrics. We hypothesize that the instructional video and the initial simulator attempt on TraumaVision prior to the first FluoroSim attempt cemented the cognitive understanding of where the guide-wire should be positioned within the femoral head. Therefore, on the initial FluoroSim attempt, most participants understood cognitively how to achieve **an ideal** TAD (and thus COR).

The learning curve of the training group presented in figure 4 strengthens this hypothesis. Participants understood the concept of TAD early on, so by their second attempt they achieved an average TAD below the 25 mm threshold.[5] In addition to cognitive understanding, users had to develop novel psychomotor skills and utilize visuospatial awareness. With increasing number of attempts, an improvement (i.e. reduction) in the procedural time and the number of radiographs used was observed, suggesting a development of psychomotor skills. By the second attempt, all undergraduate participants demonstrated the same technical competencies when compared to the first attempt (dots at first attempt on figure 4) by expert surgeons (Attendings) from the former construct validation study.[24] **Since operative success is multivariate, novices demonstrated a shorter learning curve of guide wire insertion with assistance of FluoroSim in regards to TAD, procedural time of specifically inserting the guide wire and number of radiographs taken.**

Furthermore, the learning curve demonstrates a change in baseline ability after the one week wash out period. Between the first and sixth attempts, there is an improvement in the baseline in all metrics (statistically significant for procedural time $p < 0.001$, number of radiographs $p = 0.005$ and number of guide-wire retries $p = 0.011$), suggesting skills retention after a single week. However, comparing the fifth (last attempt of first week) and sixth (first attempt of second week) attempts with one week apart, a small amount of skills decay is observed (statistically significant for procedural time; $p = 0.004$). This highlights the importance of repeat exposure to a training tool when learning a psychomotor skill. Previous simulator validation studies have used psychomotor, perceptual and visuo-spatial tests to assess baseline motor ability.[25-27]

The FluoroSim training cohort showed a greater improvement on all TraumaVision metrics compared to the control cohort, however the improvement comparisons between the two cohorts were insignificant. A possible justification for this is that although the two systems test the same cognitive skill, the psychomotor skill of manipulating a haptics-enabled stylus pen (TraumaVision) is different to that of an actual drill (FluoroSim). Cross simulator studies have been done in laparoscopic and arthroscopic simulation demonstrating skills transfer,[28, 29] however both modalities (a VR laparoscopic simulator and a box-top laparoscopic simulator) used the same equipment to assess the same psychomotor skills.

Limitations

Developing an objective feedback system based on the users' performance within the simulation software, similar to that of the Knee Arthroscopy Surgical Trainer,[30] would remove any inherent feedback bias. The training cohort were limited to ten attempts over two weeks only on FluoroSim.

Future work

Further validation **work is required to demonstrate** concurrent validation and **skill** transfer as well as gaining approval from **official medical device regulations** before FluoroSim can be accepted for **theatre utilization by surgeons rather than medical students.**

One subsequent study will look at residents 'warming up' on the FluoroSim (training) prior to guide wire insertion in patients as opposed to those without any simulation exposure (control). FluoroSim can then be recorded intraoperatively for residents to assess strength of technical skills (e.g. TAD, number of radiographs, duration of fluoroscopy, number of guide wire attempts and total procedural time) with or without

practising on the simulator pre-operatively. TAD, the only clinically validated predictor of cut-out,[6] and total procedural time can be compared between both cohorts. Alongside the absence of radiation, the training cohort of residents may potentially improve patient safety with achieving a more accurate TAD, less dependence on imaging and reduced procedural time. Yet, the next step is to have FluoroSim approved for use in theatre after ethical and FDA approval once this simulation study is published as an essential preceding phase.

Conclusion

This study has demonstrated the merits of repeat exposure to FluoroSim leading to a training effect. The training cohort developed **improved** psychomotor skills **with a shorter learning curve required to competently** insert a DHS guide-wire using the FluoroSim with actual surgical **tools**. To further **demonstrate** that the skills being developed are those needed for the actual procedure, transfer validation from the laboratory to the operating theatre is required. Nevertheless, this technology is the first of its kind to improve training opportunities in orthopedic trauma simulation while protecting **surgeons** and patients from hazards of radiation.

References

1. Royal College of Physicians. National Hip Fracture Database annual report 2015. London: RCP; 2015.
2. National Hip Fracture Database. The National Hip Fracture Database National Report 2010: The National Hip Fracture Database; 2010.
3. Marsh D, Currie C, Brown P, et al. The care of patients with fragility fracture: *British Orthopaedic Association*; 2007.
4. Utrilla AL, Reig JS, Munoz FM, et al. Trochanteric Gamma nail and compression hip screw for trochanteric fractures - A randomized, prospective, comparative study in 210 elderly patients with a new design of the Gamma nail. *J Orthop Trauma.* 2005;19(4):229-33.
5. Baumgaertner MR, Curtin SL, Lindskog DM, et al. The value of the tip-apex distance in predicting failure of fixation of peritrochanteric fractures of the hip. *J Bone Joint Surg Am.* 1995;77A(7):1058-64.
6. Andruszkow H, Frink M, Fromke C, et al. Tip apex distance, hip screw placement, and neck shaft angle as potential risk factors for cut-out failure of hip screws after surgical treatment of intertrochanteric fractures. *Int Orthop.* 2012;36(11):2347-54.
7. Baumgaertner MR, Solberg BD. Awareness of tip-apex distance reduces failure of fixation of trochanteric fractures of the hip. *J Bone Joint Surg Br.* 1997;79B(6):969-71.
8. Baratz MD, Hu YY, Zurakowski D, et al. The primary determinants of radiation use during fixation of proximal femur fractures. *Injury.* 2014;45(10):1614-9.
9. Giannoudis PV, McGuigan J, Shaw DL. Ionising radiation during internal fixation of extracapsular neck of femur fractures. *Injury.* 1998;29(6):469-72.

10. Mastrangelo G, Fedeli U, Fadda E, et al. Increased cancer risk among surgeons in an orthopaedic hospital. *Occup. Med. (Lond.)* 2005;55(6):498-500.

11. Khan IA, Kamalasekaran S, Fazal MA. Risk of ionising radiation to trainee orthopaedic surgeons. *Acta Orthop Belg.* 2012;78(1):106-10.

12. Grutzner PA, Suhm N. Computer aided long bone fracture treatment. *Injury.* 2004;35:57-64.

13. Babineau TJ, Becker J, Gibbons G, et al. The "Cost" of operative training for surgical residents. *Arch Surg.* 2004;139(4):366-9.

14. Wilson T, Sahu A, Johnson DS, et al. The effect of trainee involvement on procedure and list times: A statistical analysis with discussion of current issues affecting orthopaedic training in UK. *Surg J R Coll Surg E.* 2010;8(1):15-9.

15. Egan C, Elliott R, Fleming P. European Working Time Directive and the use of simulators and models in Irish orthopaedics. *Irish J Med Sci.* 2012;181(1):143-6.

16. Temple J. Time for training: A review of the impact of the European Working Time Directive on the quality of training. London: Department of Health; 2010.

17. Chikwe J, de Souza AC, Pepper JR. No time to train the surgeons - More and more reforms result in less and less time for training. *Br Med J.* 2004;328(7437):418-9.

18. Zuckerman JD, Kubiak EN, Immerman I, et al. The early effects of code 405 work rules on attitudes of orthopaedic residents and attending surgeons. *J Bone Joint Surg Am.* 2005;87A(4):903-8.

19. Choy I, Okrainec A. Simulation in Surgery: Perfecting the Practice. *Surg Clin North Am.* 2010;90(3):457-73.

20. Akhtar K, Sugand K, Sperrin M, et al. Training safer orthopedic surgeons Construct validation of a virtual-reality simulator for hip fracture surgery. *Acta Orthop.* 2015;86(5):616-21.

21. Blyth P, Stott NS, Anderson IA. A simulation-based training system for hip fracture fixation for use within the hospital environment. *Injury.* 2007;38(10):1197-203.

22. Sugand K, Akhtar K, Khatri C, et al. Training effect of a virtual reality haptics-enabled dynamic hip screw simulator A randomized controlled trial. *Acta Orthop.* 2015;86(6):695-701.

23. van Duren B, Sugand K, Wescott R, et al. Augmented reality fluoroscopy simulation of the guide-wire insertion in DHS surgery: A proof of concept study. *Med Eng Phys.* 2018;55:52-9.

24. Sugand K, Wescott R, Carrington R, et al. Teaching basic trauma: validating FluoroSim, a digital fluoroscopic simulator for guide-wire insertion in hip surgery. *Acta Orthop.* 2018;89(x):xx.

25. Arora H, Uribe J, Ralph W, et al. Assessment of construct validity of the endoscopic sinus surgery simulator. *Arch Otolaryngol Head Neck Surg.* 2005;131(3):217-21.

26. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance - Results of a randomized, double-blinded study. *Ann Surg.* 2002;236(4):458-64.

27. Van Sickle KR, McClusky DA, Gallagher AG, et al. Construct validation of the ProMIS simulator using a novel laparoscopic suturing task. *Surg Endosc.* 2005;19(9):1227-31.

28. Gallagher AG, Seymour NE, Jordan-Black JA, et al. Prospective, Randomized Assessment of Transfer of Training (ToT) and Transfer Effectiveness Ratio (TER) of Virtual Reality Simulation Training for Laparoscopic Skill Acquisition. *Ann Surg.* 2013;257(6):1025-31.

29. Akhtar K, Sugand K, Wijendra A, et al. The Transferability of Generic Minimally Invasive Surgical Skills: Is There Crossover of Core Skills Between Laparoscopy and Arthroscopy? *J Surg Educ.* 2016;73(2):329-38.

30. Cannon WD, Garrett WE, Hunter RE, et al. Improving Residency Training in Arthroscopic Knee Surgery with Use of a Virtual-Reality Simulator A Randomized Blinded Study. *J Bone Joint Surg Am.* 2014;96A(21):1798-806.

Figure legend

Figure 1: FluoroSim set up. (A) 1. Antero-posterior (AP) camera view, 2. Simulated digital fluoroscopy on monitor, 3. Colored balls as markers for guide wire, 4. Actual drill used in theatre, 5. Cross table lateral camera view, 6. Draped polystyrene foam block or 3D dry bone acting as a simulated femur, 7. 135-degree angle guide for guide wire insertion. (B) Orthogonal views provided by both cameras offering both AP and lateral fluoroscopic views.

Figure 2: CONSORT diagram

Figure 3 A-E: a series of box plots of the inter- and intra-group comparisons for each metric with a percentage improvement and significance value. The objective metric is presented on the y-axis with the cohort on the x-axis. The central line of each box shows the median value, with the IQR being represented by the boundaries of the box. The range without outliers is shown by the whiskers.

Figure 4: Line graph representing the learning curve of the training group for each objective metric per attempt. Median values are plotted. The graph legend states where values have been mathematically adjusted to fit on the same graph. The dots mark the values achieved by the expert surgeon cohort in a separate study.[24]

Table 1: Intra-group comparison. The median (inter quartile range) improvement is presented for each objective metric achieved on FluoroSim with a significance value. Further, the significance for the comparison of improvement between the cohorts is presented.

Table 2: Inter-group comparison. The median (inter quartile range) for each objective metric from their final attempt on FluoroSim is presented. Each metric was compared between cohorts to give a significance value.

Table 3: Change seen in the five objective metrics recorded from TraumaVision, after a different amount of exposure to FluoroSim. The median (inter quartile range) improvement for each objective metric has been recorded with a significance value. Further, the significance for the comparison of improvement between the cohorts is presented.