Augmented reality fluoroscopy simulation of the guide-wire insertion in DHS surgery: A proof of concept study

van Duren BH, BEng DPhil MBCHB CEng MIMechE MRCS
Academic Clinical fellow Trauma & Orthopaedics
Institute of Orthopaedics & Musculoskeletal Sciences
University college London
London

Sugand K PgCHBE BSc MBBS MRCS
MSk Lab, Imperial College London
Charing Cross Hospital
London

Wescott R
Intercalated BSc Student
University college London
London

Carrington R FRCS
Consultant Orthopaedic Surgeon
Royal National Orthopaedic Hospital
Stanmore

Hart A FRCS
Professor of Orthopaedics
Institute of Orthopaedics & Musculoskeletal Sciences
University college London
London

Abstract

Background: Hip fractures contribute to a significant clinical burden globally with over 1.6 million cases per annum and up to 30% mortality rate within the first year. Insertion of a dynamic hip screw (DHS) is a frequently performed procedure to treat extracapsular neck of femur fractures. Poorly performed DHS fixation of extracapsular neck of femur fractures can result in poor mobilisation, chronic pain, and increased cut-out rate requiring revision surgery. A realistic, affordable, and portable fluoroscopic simulation system can improve performance metrics in trainees, including the tip-apex distance (the only clinically validated outcome), and improve outcomes.

Method: We developed a digital fluoroscopic imaging simulator using orthogonal cameras to track coloured markers attached to the guide wire which created a virtual overlay on fluoroscopic images of the hip. To test the accuracy with which the augmented reality system could track a guidewire, a standard workshop femur was used to calibrate the system with a positional marker fixed to indicate the apex; this allowed for comparison between guide-wire tip-apex distance (TAD) calculated by the system to be compared to that physically measured. Tests were undertaken to determine: 1. how well the apex could be targeted 2. the accuracy of the calculated TAD 3. The number of iterations through the algorithm giving the optimal accuracy-time relationship.

Results: The calculated TAD was found to have an average root mean square error of 4.2mm. The accuracy of the algorithm was shown to increase with the number iterations up to 20 beyond which the error asymptotically converged to an error of 2mm.

Conclusion: This work demonstrates a novel augmented reality simulation of guide-wire insertion in DHS surgery. To our knowledge this has not been previously achieved. In contrast to virtual reality, augmented reality is able to simulate fluoroscopy while allowing the trainee to interact with real instrumentation and performing the procedure on workshop bone models.
Introduction

Hip fractures contribute to a significant clinical burden globally with over 1.6 million cases per annum(1). Domestically, 70,000 patients fracture their hip in the UK each year, resulting in £2 billion of care(2). These fractures carry a high morbidity and mortality, with up to a 30% mortality rate within 1 year(3–6). Additionally, after hip fracture, patients are five times more likely to be institutionalised at one year than their age-matched controls(7). Extracapsular fractures account for a significant subsection of neck of femur fractures.

Extracapsular neck of femur fractures can be treated with a fixed angle sliding screw device, better known as a sliding, compression or dynamic hip screw (DHS). However, mechanical failure rates of up to 20% have been reported(8–11). The tip-apex distance (TAD: the sum of the distance between the tip of the lag screw to the apex of the femoral head on both the antero-posterior [AP] and cross table lateral [CTL] radiographs) is considered to be the strongest predictor for screw cut-out, necessitating further surgery and increasing morbidity or mortality(12). Studies have found that the optimal TAD needs to be less than 20mm for reducing the risk of cut-out and mechanical failure(12). To achieve this optimal TAD during surgery, intraoperative fluoroscopy is used to position a guide wire which forms the trajectory along which the implant is placed.

It is becoming increasingly difficult for trainee surgeons to gain exposure to these operations early on in their career(13,14) resulting in less practice and ultimately poorer outcomes owing to poor technique, poor implant placement, and increased procedural time. Poorly performed procedures can result in poor mobilisation increasing the risks associated with co-morbidities, chronic pain, and higher cut-out rates requiring revision surgery. Simulation is a well-recognised teaching adjunct to train high-risk tasks within a safe and controlled environment to prevent harm to patients(15–20). Simulation in orthopaedic surgery can be divided into three broad categories: 1) low-fidelity saw bones (dry-lab) 2) virtual reality (VR) and 3) cadaveric (wet-lab) simulation(21–25).

In spite of DHS surgery being one of the more common lower limb trauma procedures worldwide, there are very few resources available to simulate DHS surgery. In the aforementioned methods of simulating DHS surgery, the placement of the guide-wire to achieve an optimal TAD of less than 20mm, arguably the most crucial part of the procedure, is unrealistic. Current methods may teach the principle of the procedure, however the practical tasks involved are far removed from reality due to either the lack of fluoroscopy or actual tools which reduces the fidelity and realism of the simulation. Fluoroscopy is usually not permitted in simulation due to the exposure of radiation. A realistic, affordable, and easily accessible simulation of DHS surgery would provide a useful step in developing and perfecting the practical implementation to the operating theatre that will inevitably enhance patient safety.

Training will be improved by developing an affordable and high-fidelity modality to simulate the fluoroscopic guidance used within orthopaedic theatres. In this paper we present a novel augmented reality fluoroscopic simulator to simulate the insertion of the guide-wire during DHS surgery. To our knowledge this has not been previously achieved. The proposed device make use of visual tracking using two video cameras in conjunction with image processing algorithms, to overlay the guide-wire position on corresponding fluoroscopic images. Additionally, we present results obtained while testing the accuracy of the proposed fluoroscopic simulator.
Methods

Design
In order to achieve a realistic simulation of the fluoroscopy used to guide placement of guide-wire placement in theatre, a method to simulate the fluoroscopic imaging was developed.

A prototype of this technology was developed and is described here to illustrate the methodology. To track the instrument (specifically the guide-wire), two orthogonally placed cameras are used to detect coloured markers attached to the guide-wire. Two Logitech C920 cameras (Logitech, Romanel-sur-Morges, Switzerland) were used; one placed in a position to capture the cross table lateral (CTL) and the other placed to capture the anterior-posterior (AP) view, as would be the case when using a single fluoroscopic C-arm. The cameras were placed at fixed positions ([0,500,0] & [500,0,0] for AP and CTL cameras respectively) from the centre of calibration (see figure 1a).

Two coloured (green and yellow) markers were attached to the guide wire (figure 1a & 3). The CTL and AP camera images of the femoral head of a workshop bone (3B Scientific, Hamburg, Germany) were mapped onto previously obtained genuine fluoroscopic CTL and AP images using an affine transformation matrix. Three points on each camera image (AP: tip, lesser trochanter, greater trochanter; CTL: tip, inferior neck, superior neck) and the corresponding points on the fluoroscopic images were selected using a graphical user interface (GUI). Figure 2a-d illustrates the corresponding points between video and fluoroscopic overlay images used as inputs. The corresponding points were used to calculate an affine transformation matrix which then allowed for the mapping of points from the camera images to corresponding points on the fluoroscopic images in real time.

The position of the guide wire in the camera images was determined based on tracking the attached coloured markers. To obtain the marker positions the recorded RGB (red, green, blue) colour space images were converted to HSV (hue, saturation, value) colour space images and a threshold filter (minimum and maximum thresholds for hue, saturation, and value for ambient light conditions were predetermined by the user) were used to isolate the coloured pixels for each marker individually (26). The images were processed prior to segmentation using a Gaussian filter (5 pixel radius). A best fit circle was then fitted to the identified pixels and the centre of the circle was used as the marker position (see figure 1). To minimise positional variation due to ambient light conditions the position of the marker was averaged over multiple images; a variable specified by the user. Averaging over a larger number of images resulted in a more stable marker position but a slower computation time.

Knowing the position of the guide-wire tip relative to the markers meant that the guide-wire position could be mapped onto the corresponding fluoroscopic image as an overlay. The use of two markers, at predetermined positions from both each other and relative to the tip, made it possible to calculate the wire trajectory taking into account the effect of out of plane motion on the two-dimensional image representation (27). Application of the calculated affine transformation matrix to the trajectory allowed for an image of the guide-wire to be overlaid on the fluoroscopy image, to simulate a fluoroscopic environment without being exposed to harmful radiation.

Python (version 2.7) and OpenCV (version 3.0.0) were used to develop the image processing algorithms used to track the coloured marker spheres as described. The system was capable of providing both static ‘snapshots’ as well as live screening as is the case in a real theatre environment. The code developed recorded a number of parameters as outputs to measure the performance of the user/trainee: the time taken, the number of images taken, the AP, CTL, and absolute distance from the tip of the wire to the apex and a calculated TAD.

Testing
Three tests were performed to assess the accuracy of the system: 1. Absolute accuracy under laboratory conditions 2. Accuracy convergence with number of images 3. In use accuracy when the simulation was used by naïve subjects.

Absolute accuracy
To test the accuracy, or how well the system represents reality, a specialised “marker jig” and a “drilling block” were built (figure 3). The system was set up with a right sided proximal femur and the pointer of the maker jig was placed at the apex of the femoral head (figure 3a). The system was then calibrated to determine the affine transformation matrix. Once calibrated the femur was removed and replaced by the drilling block (figure 3b) and this was subsequently covered to obscure the pointer marking the apex which remained in situ (figure 3c).

Using this setup, the guide wire was inserted at 135 degrees using a 135 degree guide (135 degrees being the angle between the femur and femoral neck) under simulated fluoroscopic guidance, aiming for 18 separate points in relation to the marker, to give both a range of distances and positions from the defined “apex”. The position of the tip of the guide-wire, when placed, only deviated from the marker in a single plane. This positioning was recorded in cartesian coordinates (x, y, z which corresponded to the x-axis being perpendicular to the neck, the y-axis along the neck and, the
z-axis being anterior-posterior); the deviation in the x axis was recorded as (±, 0, 0), in the y axis as (0, ±, 0) and in the z axis as (0, 0, ±). Once the guide-wire was in position, the cover was removed and the distance between the tip of the guide-wire and the apex marker was measured using a digital vernier caliper. This distance was then compared to the absolute distance calculated by the fluoroscopy simulator.

In total 270 separate measurements were recorded; 90 measurements on three separate occasions. Three separate test session were 1 week apart to determine whether there was any significant variation in accuracy due to calibration. Measurements were recorded between 0-30mm from the apex in all planes to determine if accuracy was dependent on both the distance from the apex and the plane in which the tip deviated.

Convergence
In addition to testing simple accuracy, the effect of the number of images used in averaging the marker position (see methods) to minimise positional variation due to ambient light conditions was determined. With the tip of the guide-wire placed at the apex (as marked by the marker jig) 10 measurements were taken using the simulation and recorded. The variation in measurement values was taken to be the spread of the measurements (largest - smallest). This was repeated for using 5 image intervals up to 40 images.

In use accuracy
Finally, as part of a separate validation study, we obtained measurements from 406 simulator attempts, completed by 67 medical students using the same set up as described above. All students were naive to the DHS procedure, but were instructed to aim for the apex of the femoral head on each attempt. The absolute difference between the measured distance and the simulator calculated distance was recorded for accuracy purposes.

Analysis
Basic descriptive statistical methods were used to compare the simulator determined TAD to that measured using a vernier caliper. Both the absolute errors and root mean square (RMS) errors are presented. Average error and standard deviation (SD) are presented as well as the median error and inter quartile range (IQR) to quantify the effect of outliers. To investigate the effect of the distance of the guidewire tip to femoral apex on accuracy three groups (1. <10mm 2. 10mm<20mm 3. >20mm from the apex) were compared. Kruskal-Wallis testing compared all three groups followed by Mann-Whitney U testing corrected for multiple comparisons. To determine the significance of the difference in accuracy between the x, y and z axes, the same method was used. To determine the significance between the three separately calibrated sessions the Kruskal-Wallis test was used. The difference between the absolute accuracy and the in use accuracy was analysed using the Mann-Whitney U test.
Results

Absolute accuracy
The average error for the 270 accuracy measurements was 0.4mm [SD: 5.5mm], the median was 1.1mm (Range: -15.75mm – 13.5mm). The median RMS error was 3.35mm [IQR 1.15mm – 6.53mm] and the average RMS was 4.2mm. The RMS error plotted against the actual distance of the tip of the guide wire from the apex is plotted in figures 4&5. The median RMS error showed a graded increase the further away the tip of the guide wire was from the marker (see table 1). The difference in distribution of absolute error was significant between all three groups with p < 0.001.

The RMS error showed a variation depending on the plane in which the tip of the guide wire deviated from the marker. Deviation in the x axis (median = 1.37mm, IQR [0.48mm, 2.73mm]) had a lower distribution of RMS error than deviation in both the y axis (median = 6.39mm, IQR [2.11mm, 7.67mm]) and the z axis (median = 5.10mm, IQR [2.57mm, 7.17mm]; table 2).

The RMS error was also compared for each different occasion of testing, as the calibration process is subjective. There was no significant difference in the distribution of RMS error for each of the three calibration attempts H(2) = 0.624, p = 0.732.

Convergence
The variation in values measured by the algorithm was shown to decrease with the number of images processed up to 20 images beyond which the range converged to approximately a 2mm range. There was linear increase in processing time corresponding to an increase in number of images processed to obtain a result corresponding to 1 second per 5 images.

In use accuracy
A comparison of the RMS error from the accuracy data set and the student data set showed that the RMS error obtained from the student data set (median = 4.79mm, IQR [2.28mm, 8.99mm]) was larger than that measured using the accuracy experiments as well as having an increased range of values [U = 42,447.50, p < 0.001].
Discussion

This study has demonstrated that the use of orthogonal cameras to track coloured markers attached to a guide-wire to overlay a fluoroscopic hip image is a viable means of creating a unique realistic, affordable, and easily accessible simulation of guide-wire placement. This methodology can provide a useful adjunct in training surgeons to perform DHS surgery and practice technical skills in a high-fidelity environment.

Currently available commercial VR simulators can emulate the surgical decision making process of the procedure as well as practising the guide wire insertion for optimal TAD(21–23). However, these simulators can cost tens of thousands of pounds excluding constant maintenance and upgrading costs. VR simulation lacks a high-fidelity environment since all steps (using the drill, reamer, screw driver etc.) are being simulated using a haptic-enabled phantom stylus pen rather than handling actual tools. Due to funding and resource limitations, most centres rely on practising technical skills on low-fidelity saw bones by over drilling through the femoral head to observe the proximity to optimal TAD. However, this is an inaccurate method of measuring the actual TAD since the femoral head is a three-dimensional sphere and an exit point of the guide wire represents a two-dimension measurement only. With the limitation of using fluoroscopy in simulation, trainees cannot practise fluoroscopic-guided guide wire insertion due to the ethics of unnecessary radiation exposure. Furthermore, this technique has also been implemented at national orthopaedic resident selection. This currently lacks content and face validity as well as being unrealistic. The use of the fluoroscopic simulation of guide-wire insertion described here overcomes this problem.

The root mean square error determined using the experimental setup showed that on average the system error was 4.2mm. Further analysis of the results showed that 92.5% of the calculated measurements were within 10mm of the actual measurement and 62.5% within 5mm of actual measurement. Given that the aim is to place the guide wire within a TAD of 20mm the errors described may be misleading, particularly in cases where the guidewire is further away from the apex. However within the context of simulating the insertion of the guidewire the use of the simulated fluoroscopy system provided a realistic simulation of the techniques required to achieve an optimal TAD. Future work to refine the accuracy of the system would aid this further.

A number of factors were noted to contribute to the observed error. The images obtained from the cameras were not corrected for distortion (e.g. pinhole distortion & perspective distortion); this was done to maintain the simplicity of the system. The disparity between calculated and measured values increased the further the guide wire tip was away from the apex. It is likely that this was due to multiple contributing factors. The points used to determine the affine transformation matrix were centred around the femoral head which only represents a small section of the entire image. This resulted in the transformations being less accurate further away from this section. Additionally in future additional accuracy could be obtained through 1. the use of more than three corresponding points to calculate the affine transformation matrix and 2. investigating the most reliable calibration points as errors also arise from a variation in anatomical interpretation when selecting points. Out of plane perspective was corrected for but also became more inaccurate the further out of plane the guide-wire was rotated. In future this could be corrected by calibrating the cameras in stereo. However, this would require a more complex calibration sequence making the system more difficult to use.

The effect of variations in ambient light were noted to influence the reliability of the calculated marker positions; the apparent colour of an object depends upon the illumination conditions, the viewing geometry and the camera parameters, all of which can vary over time. This was overcome to a large extent by averaging the calculated marker position taken from multiple sequential images. Using an increasing number of images improved colour constancy and reduced the variation in measurements. However as shown in figure 6 this effect was less pronounced when more than 20 images were used. This problem could be overcome using infrared markers and cameras, but again, this would add considerably to the expense and complexity.

In addition to the effect of light and image distortion on the accuracy of the system the measurements obtained from the simulator attempts by novices using the system showed an increase error and variation in results. This was observed to be due to the bending of the guide-wire when embedded in bone; the novices had a tendency to move the drill holding the guide-wire while the tip was fixed in the sawbone. The resulting curvature in the wire meant that the algorithm that calculated a straight line projection became more inaccurate. To overcome this users required instruction to maintain the linearity of the drill to the guide-wire at all times. In reality, the femur can move during the simulation if not secured properly however this did not occur during testing. A solution to this would be to track the femur itself in addition to the guidewire.
Conclusion

The system as a proof of concept demonstrated that the use of coloured markers to track the position of the guide-wire and overlay this on a fluoroscopic image calibrated to correspond with a sawbone provides an augmented reality fluoroscopic simulation of the insertion of a guide-wire in DHS surgery. The equipment used is readily available commercially and inexpensive, and the software was open source and freely available. To our knowledge augmented reality simulation of guide-wire insertion under fluoroscopic guidance has not previously been achieved. In contrast to virtual reality, augmented reality is able to simulate fluoroscopy while allowing the trainee to interact with real instrumentation while performing the procedure on workshop bone models. It combines widely accepted simulation techniques with a implementation of superimposed modernised technology.

The accuracy of the system described can be improved by using camera calibration algorithms and stereo camera three-dimensional space calibration. We consciously chose to develop this system without camera or environment calibration in the initial instance as it is the intention that the code would be readily available to students and trainee surgeons (as well as anybody else) who wished to use/recreate the system. As such the aim was to make it useable for a variation in camera types (depending on what is available/affordable) and because cameras make use of varying specifications (aperture, focal lengths) this would require these to be known as well as introducing added complexity.

Ultimately this technology provides a relatively inexpensive means of augmenting sawbone simulation teaching to include arguably one of the most important steps of a CHS/DHS procedure. This technology has the potential to be further developed to augment the reaming and screw insertion steps as well as being applicable to many other procedures in trauma and orthopaedics that rely on fluoroscopic guidance.
### Tables:

<table>
<thead>
<tr>
<th>Guide-wire tip distance (mm)</th>
<th>Median (mm)</th>
<th>IQR (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 10</td>
<td>1.53</td>
<td>2.95</td>
</tr>
<tr>
<td>10 &lt; x ≤ 20</td>
<td>4.97</td>
<td>5.09</td>
</tr>
<tr>
<td>&gt; 20</td>
<td>7.23</td>
<td>6.69</td>
</tr>
</tbody>
</table>

Table 1: Table summarizing the median RMS error with a graded increase in the distance of the tip of the guide-wire from the apex position marker.

<table>
<thead>
<tr>
<th>Axis comparison</th>
<th>Mann-Whitney U test statistic</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>X and Y axis</td>
<td>U = 1908.00</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>X and Z axis</td>
<td>U = 1272.00</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>Y and Z axis</td>
<td>U = 3846.00</td>
<td>p = 0.559</td>
</tr>
</tbody>
</table>

Table 2: Table showing the RMS error variation with regard to the plane in which the tip of the guide-wire deviated from the apex marker.
Figures:

Figure 1: A) The top image shows the experimental setup of FluoroSim. 1 and 5 represent the AP and CTL cameras respectively. 3 demonstrates the two coloured markers attached to the guide-wire. 4 is the orthopaedic drill and 7 is the 135-degree angle guide. 6 is the drilling block and 2 shows a digital radiograph produced by the system. B) The lower image shows a screen shot of the developed software displaying AP and CTL video images with tracked markers above corresponding to fluoroscopic images with an overlay of the guidewire position below.
Figure 2: Illustration of the corresponding points, for both AP & CTL cameras, used during calibration to calculate the affine transformation matrices; points of the femoral head are calibrated and superimposed onto the standard fluoroscopic image template. A and B demonstrate calibration in the AP plane and C and D demonstrate calibration in the CTL plane.
Figure 3: A) Initially the software was calibrated using a right sided femur with the plastic marker jig sitting at the apex of the femoral head. The software is then calibrated so the marker jig becomes the centre of calibration. B) The right sided femur is replaced with the drilling block but the marker jig which remains the centre of calibration for measurement purposes. C) The drilling block and marker jig were covered for testing.
Figure 4: A graph of the tip-apex distance calculated by the simulator plotted against the actual measured tip-apex distance (using vernier caliper). Perfect agreement (i.e. 0 error) is represented by the dashed line. The solid line represents the best fit line to the data with an $R^2$ value of 0.55.

Figure 5: A graph of the root mean square (RMS) error plotted against the actual measured tip-apex distance (using vernier caliper). The solid line shows the best fit line demonstrating the increase in error with increase in distance from the apex. The dashed line represents a RMS error of 5mm.
Declarations:

Competing interests:
None declared

Funding:
This project was funded by a grant from the Professor A.T. Fripp fund, Royal National Orthopaedic Hospital, Stanmore, London
B.H. van Duren is an NIHR funded Clinical Fellow in Trauma & Orthopaedics
R. Wescott received funding assistance from the Goldberg Schachmann & Freda Becker Memorial Fund

Ethical approval:
Not required
References:


27. Sutton MA, Yan JH, Tiwari V, Schreier HW, Orteu JJ. The effect of out-of-plane motion on 2D and 3D digital