Follow us @BoneJointOpen

Freely available online OPEN ACCESS

BJO

TRAUMA

Hip fracture surgery performed by cadaveric simulation-trained versus standard-trained orthopaedic trainees: a preliminary multicentre randomized controlled trial

H. K. James,
G. T. R. Pattison,
J. Griffin,
J. D. Fisher,
D. R. Griffin

From Warwick Clinical Trials Unit, University of Warwick, UK

Aims

To evaluate if, for orthopaedic trainees, additional cadaveric simulation training or standard training alone yields superior radiological and clinical outcomes in patients undergoing dynamic hip screw (DHS) fixation or hemiarthroplasty for hip fracture.

Methods

This was a preliminary, pragmatic, multicentre, parallel group randomized controlled trial in nine secondary and tertiary NHS hospitals in England. Researchers were blinded to group allocation. Overall, 40 trainees in the West Midlands were eligible: 33 agreed to take part and were randomized, five withdrew after randomization, 13 were allocated cadaveric training, and 15 were allocated standard training. The intervention was an additional two-day cadaveric simulation course. The control group received standard on-the-job training. Primary outcome was implant position on the postoperative radiograph: tip-apex distance (mm) (DHS) and leg length discrepancy (mm) (hemiarthroplasty). Secondary clinical outcomes were procedure time, length of hospital stay, acute postoperative complication rate, and 12-month mortality. Procedure-specific secondary outcomes were intraoperative radiation dose (for DHS) and postoperative blood transfusion requirement (hemiarthroplasty).

Results

Eight female (29%) and 20 male trainees (71%), mean age 29.4 years, performed 317 DHS operations and 243 hemiarthroplasties during ten months of follow-up. Primary analysis was a random effect model with surgeon-level fixed effects of patient condition, patient age, and surgeon experience, with a random intercept for surgeon. Under the intention-to-treat principle, for hemiarthroplasty there was better implant position in favour of cadaveric training, measured by leg length discrepancy ≤ 10 mm (odds ratio (OR) 4.08 (95% confidence interval (CI) 1.17 to 14.22); p = 0.027). There were significantly fewer postoperative blood transfusions required in patients undergoing hemiarthroplasty by cadaveric-trained compared to standard-trained surgeons (OR 6.00 (95% CI 1.83 to 19.69); p = 0.003). For DHS, there was no significant between-group difference in implant position as measured by tip-apex distance ≤ 25 mm (OR 6.47 (95% CI 0.97 to 43.05); p = 0.053). No between-group differences were observed for any secondary clinical outcomes.

Conclusion

Trainees randomized to additional cadaveric training performed hip fracture fixation with better implant positioning and fewer postoperative blood transfusions in hemiarthroplasty. This effect, which was previously unknown, may be a consequence of the intervention. Further study is required.

Correspondence should be sent to Hannah K James; email: hsmith22@doctors.org.uk

doi: 10.1302/2633-1462.48.BJO-2022-0143.R1

Bone Jt Open 2023;4-8:602–611.

Cite this article: *Bone Jt Open* 2023;4-8:602–611.

Keywords: Simulation, Education, Training, Hip fracture

Introduction

Surgeon-related factors affecting the outcome of hip fracture surgery are a neglected yet important research area. There is evidence that the technical skill[1] and experience level[2] of the surgeon is inversely associated with postoperative outcomes, and that a measurable learning curve exists in acquiring procedural competence in orthopaedic surgery.[3,4] Hip fracture operations are usually performed by junior trainees. The associated learning curve, which has been estimated to be around eight to ten procedures to achieve performance plateau,[4] takes place on live patients.

There is interest in the role of simulation for rapid surgical skill acquisition in a controlled environment remote from patients, where competency can be objectively assessed before trainees operate on real patients. This may shorten the learning curve and safeguard patients from potential avoidable harms.[5] In the post-COVID-19 training climate, simulation has been cited as a central part of the training recovery strategy.[6]

Cadaveric simulation, the use of deceased human bodies to practice surgery, allows trainees to practice entire operations in an ultra-realistic 'high-fidelity' environment. There is an abundance of recent evidence that cadaveric simulation may induce short-term behavioural change in the simulation laboratory.[7] There is, however, a lack of evidence of skill retention over time, transfer of skills following training into the operating theatre, or benefit to patients.[7] Our aim was to determine whether it was feasible to conduct a study to test if additional cadaveric simulation training or standard on-the-job surgical training alone leads to the best patient outcomes for dynamic hip screw (DHS) and hemiarthroplasty performed by junior trainees.

Methods

Study design and participants. This was a pragmatic, multicentre, assessor-blind randomized controlled trial in nine NHS hospitals in the UK. The study was carried out according to the published protocol.[8] Participants were recruited from three orthopaedic training programmes in the West Midlands. Postgraduate year (PGY) 3 to 5 trainees were eligible for recruitment, i.e. Core Trainees year one or two (CT1 to 2), Speciality Trainees year one or two (ST1 or 2), and Speciality Trainees year 3 (ST3). In the UK, surgical training takes at least ten years following graduation. Exclusion criteria were consent refusal or unavailability for the course. Support for the study was agreed prospectively with training programme directors. Eligible trainees were identified by liaison with programme administrators and invited to participate by email. Recruitment was undertaken in June 2014 and the intervention delivered in September 2014. All participants gave written, informed consent. The study was approved by NHS Research Ethics (15/WM/0464) and Confidentiality Advisory Group (16/CAG/0125).

Randomization and masking. Randomization lists were computer-generated, using a simple blocking scheme (of size 4) prepared by a statistician with no further involvement in the trial. It was not possible to mask participants to their allocation. The researchers who collected outcome data and analyzed results were masked to allocation by concealment of training allocation.

Training interventions. The cadaveric-trained group received an intensive two-day cadaveric simulation training course where four procedures were taught: DHS insertion, hip hemiarthroplasty, ankle fracture fixation, and lower limb fasciotomy. The course was held in the second month of the surgical training year. These procedures were selected as they are mapped to the UK curriculum for progression to PGY 6.[9]

The course ran for two full days with nine expert faculty teaching on eight fresh-frozen cadavers, which were purchased under license from a specialist supplier.[10] The participant:faculty and participant:cadaver ratios were 2:1. The physical, psychological, and environmental fidelity of the training was maximized by using full surgical dress for participants, surgical drapes for the cadaveric 'patients', and comprehensive surgical instrument trays, implants and cement. Image intensification and radiographers were available, and scrub nurses were assigned to each station. Medical students acted as anaesthetists. Eight simulated operating theatres were set up as two parallel circuits of four. Left-sided procedures were performed on day one and right-sided procedures on day two. All participants performed the four procedures, once as first surgeon and once as assistant, during the course. Immediate structured feedback was given to the participants after each procedure through the completion of procedure-based assessments, which is the current gold-standard technical skill assessment tool in UK surgical training.[11] Maximum use was made of the cadaveric surgical environment, with no pressure of time and no patient safety requirement to ensure that novice surgeons did not operate beyond their current level of competence. After the course, the cadaveric-trained group returned to their respective hospitals and continued to receive standard training.

Standard training was delivered in the working environment with trainees receiving training in the management of these conditions from their educational supervisors, when suitable patients with these conditions presented to the training hospitals. Trainees allocated to the 'standard training' group received the above-standard training from the start of the training year. This group then received the cadaveric course towards the end of the training year, a condition mandated in the ethical approval granted for the study. It was considered unfair to exclude half of the trainees from the educational

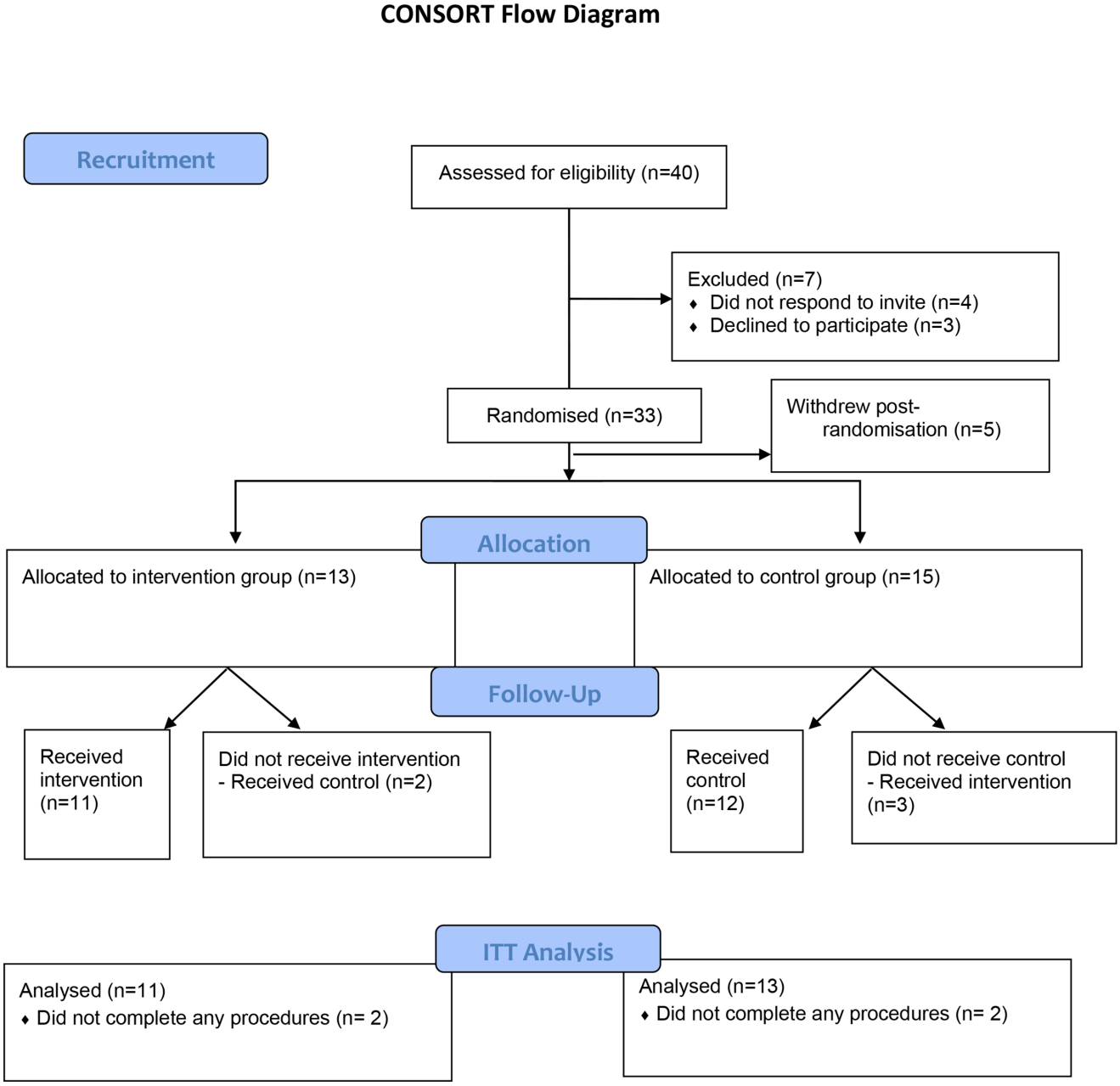

Fig. 1

CONSORT flow diagram. ITT, intention-to-treat.

experience of a cadaveric course (although the effect of the training course was not known at the time of approval). The effect of the catch-up course on these trainees was not studied. Standard training also includes fortnightly didactic teaching sessions on curriculum-matched topics, which was received by both groups in the trial.

Outcomes. Each procedure had a distinct set of pre-specified outcomes. The primary outcome measure for DHS procedures was the position of the implant on the first post-implantation radiograph, measured as tip-apex distance (TAD) in millimetres, and leg length discrepancy (LLD) in millimetres for hemiarthroplasty procedures.

TAD > 25 mm in DHS is known to predict risk of operation failure by device cut-out (risk ratio 12.7),[12] and a high anterior lag screw position predicts cut-out risk independently of TAD.[13] In hemiarthroplasty, a LLD > 10 mm is known to increase the risk of postoperative pain.[14] Varus malposition of the femoral stem > 5° from neutral in the AP view is associated with failure rates of up to 46% in total hip arthroplasty (THA).[15] There are no equivalent survival studies in hemiarthroplasty, although it is believed that varus malpositioned stems are more likely to dislocate than those in neutral or valgus.[16] These radiological measurements have shown face and construct

Table I. Participant baseline demographic information. Numbers reflect groups as randomized.

Variable	Intervention (n = 13)	Control (n = 15)
Mean age, yrs (SD)	28.0 (1.7)	30.7 (4.9)
Female, n (%)	5 (38)	3 (20)
Mean completed months of T&O experience (SD)	13.5 (12.3)	15.2 (13.9)
Mean number of previous procedures (SD)		
DHS	14.9 (10.8)	8.9 (8.0)
Hemiarthroplasty	9.9 (9.9)	6.7 (9.9)
Mean number of procedures during study (SD)		
DHS	15.2 (10.5)	10.7 (7.5)
Hemiarthroplasty	12.1 (12.0)	7.3 (5.0)

DHS, dynamic hip screw; SD, standard deviation; T&O, Trauma and Orthopaedics.

validity and reliability for measuring technical skill in a relevant population for this trial in previous work.[17]

The procedures of interest performed by trainees during the ten-month follow-up period were studied. Only procedures coded as 'supervised-trainer scrubbed' and 'supervised-trainer unscrubbed' were included in the analysis.[18] The radiographs of these procedures were accessed, and the measurements obtained by a single researcher (HKJ).

All measurements were taken by the same member of the research team (HKJ – orthopaedic registrar), using the standard toolkit available in the picture archiving and communication system (PACS). A radiology manual was produced at the start of the study detailing measurement techniques for each parameter. Most radiographs for DHS were from intraoperative image intensification (II). For these, manual scaling was performed using a known, fixed parameter (screw thread diameter in mm), as typically II images are not automatically scalable in PACS.

For hemiarthroplasty, LLD was measured in PACS based on the anteroposterior radiograph using the teardrop method. A pelvic reference line was drawn between the inferior borders of the acetabular tear drops, and a femoral reference line was drawn between the two lesser trochanters. The perpendicular distance between the two lines was measured and then compared for both sides.

Femoral stem alignment was measured as the angle subtended by the longitudinal axis of the femoral shaft and the longitudinal axis of the stem. Femoral offset was measured as the distance between the centre of the head and the longitudinal axis of the femur. Quality of cementation was judged using the Barrack grading system, where 'poor' cementation was defined as Barrack C and D.

We did not measure rotation or seek to standardize or calibrate any film, as this pragmatic trial used radiographs collected as part of routine clinical care, and therefore the quality was beyond our control.

Secondary outcome measures were procedure time, intraoperative radiation dose in $mGym^2$ to patients (DHS), blood transfusion requirement (hemiarthroplasty), acute postoperative complication rate, length of hospital stay, and 12-month mortality.

Complications were extracted from discharge information and hospital episode data, and defined as occurring during the immediate postoperative hospital stay. They were categorized as medical and surgical complications: medical complications were hospital-acquired pneumonia, renal complications, cardiac complications, and inpatient death; surgical complications were wound complications, deep infection, cut-out (DHS), and dislocation (hemiarthroplasty). Subgroup analyses by complication type were not undertaken because of small numbers.

Statistical analysis. This is the first randomized trial to objectively measure transfer of open surgical skills from the simulation laboratory to the real-world operating theatre using patient-centred outcome measures. It was not known a priori what the minimal clinically important difference (MCID) was for the primary outcome measures.

Given the highly exploratory nature of the trial, and the fact that the surgical training centre could accommodate 16 delegates, a pragmatic maximum sample size was set at 16 in each arm of the trial. No formal power calculations were undertaken, and no interim analyses were planned. The plan was approved by the acting data monitoring committee.

We investigated differences in the primary outcome measure – implant position on the first post-implantation radiograph – between the two training groups on an intention-to-treat (i.e. train) (ITT) basis.

Multivariable multilevel mixed effects models were used to assess the effects of the training intervention on implant position, allowing for potential within-surgeon correlation between repeated observations, treating surgeon clustering as a random effect. The regression models adjusted for important fixed effects: patient condition (American Society of Anesthesiologists grade), patient age (years), and surgeon experience at baseline (number of prior cases).

Primary inferences were drawn from the intention-to-train analysis without imputation for missing data.

Table II. Summary statistics of primary and secondary outcomes at follow-up, by randomized group for both procedures.

Outcome	Count	Control	Intervention	Difference (95% CI)
Hemiarthroplasty				
Radiological				
LLD, n (%)				
≤ 10 mm	200	75 (88)	125 (97)	
> 10 mm	14	10 (12)	4 (3)	9% (1% to 16%)
Femoral stem alignment, n (%)				
Acceptable alignment	183	65 (78)	118 (93)	
Nonacceptable alignment	27	18 (22)	9 (7)	15% (5% to 25%)
Femoral offset, n (%)				
< 15 mm compared to contralateral side	188	71 (83)	117 (85)	
≥ 15 mm compared to contralateral side	36	15 (17)	21 (15)	2% (-7% to 12%)
Quality of cementation, n (%)				
Good	132	59 (69)	73 (58)	
Poor	80	27 (31)	53 (42)	-11% (-24% to 2%)
Non-radiological, n (%)				
Procedure time (mins)	217	91.7	77.4	14.3 (7.3 to 21.4)
Blood transfusion within 48 hours, n (%)				
No	202	66 (85)	136 (97)	
Yes	16	12 (15)	4 (3)	13% (4% to 21%)
Any complication, n (%)				
No	186	55 (76)	131 (87)	
Yes	35	17 (24)	18 (13)	7% (-3% to 17%)
Length of stay (days)	217	20.1	19.5	0.6 (-3.6 to 4.8)
Survival to 12 mths postoperatively, n (%)				
Yes	170	61 (76)	109 (72)	
No	60	16 (24)	44 (28)	-4% (-16% to 7%)
DHS				
Radiological				
Tip-apex distance, n (%)				
≤ 25 mm	208	111 (85)	146 (95)	
> 25 mm	75	19 (15)	7 (5)	10% (3% to 24%)
Position of lag screw in femoral head, n (%)				
Central ± inferior	217	91 (69)	126 (81)	
Superior and/or anterior	71	41 (31)	30 (19)	12% (2% to 22%)
Plate flush to lateral cortex, n (%)				
Yes	181	82 (62)	99 (62)	
No	111	51 (38)	60 (38)	1% (-11% to 12%)
8 cortex screw hold, n (%)				
Yes	284	128 (98)	156 (98)	
No	6	3 (2)	3 (2)	0% (-3% to 4%)
Non-radiological				
Mean procedure time, mins	274	67.2	59.4	7.8 (3.5 to 12.1)
Mean intraoperative radiation dose, Gym^2	147	0.141	0.225	-0.084 (-0.162 to -0.005)
Any complication, n (%)				
No	233	95 (77)	138 (85)	
Yes	54	29 (23)	25 (15)	8% (-1% to 17%)
12-month mortality, n (%)				
No	267	112 (91)	155 (94)	
Yes	21	11 (9)	10 (6)	3% (-6% to 12%)

CI, confidence interval; DHS, dynamic hip screw; LLD, leg length discrepancy.

Additionally, we performed a simple descriptive analysis of between-group comparisons for procedure time, complications, length of stay, mortality, and blood transfusion. Secondary analysis of the primary outcomes under a 'per-protocol' approach was undertaken, to contextualize the results of the ITT analysis and to aid understanding of the impact of receiving training on the primary outcome measures.

Table III. Results of linear mixed models on hemiarthroplasty procedures, under intention to treat and per-protocol approach.

Outcome	Count	Multivariable random effect, OR (95% CI)*			
		Intention to treat	p-value	Per protocol	p-value
Radiological†					
LLD					
≤ 10 mm	200	1.00		1.00	
> 10 mm	14	4.08 (1.17 to 14.22)	0.027	6.19 (1.80 to 21.31)	0.004
Femoral stem alignment					
Acceptable alignment‡	183			1.00	
Unacceptable alignment	27	4.73 (0.90 to 24.94)	0.067	10.36 (2.73 to 39.34)	0.001
Femoral offset					
< 15 mm compared to contralateral side	188	1.00		1.00	
≥ 15 mm compared to contralateral side	36	1.09 (0.51 to 2.33)	0.833	1.30 (0.61 to 2.79)	0.496
Quality of cementation					
Good	132	1.00		1.00	
Poor	80	0.65 (0.31 to 1.39)	0.268	1.23 (0.57 to 2.67)	0.604
Non-radiological					
Procedure time (mins)§	217	3.39 (-10.74 to 17.51)	0.638	-1.06 (-15.34 to 13.22)	0.884
Blood transfusion within 48 hours					
No	202	1.00		1.00	
Yes	16	6.00 (1.83 to 19.69)	0.003	4.43 (1.52 to 12.95)	0.007
Complications					
Any complication					
No	186	1.00		1.00	
Yes	35	1.63 (0.51 to 5.22)	0.412	2.43 (0.79 to 7.44)	0.120
Length of stay (days)§	217	0.37 (-4.43 to 5.18)	0.879	3.92 (-0.70 to 8.53)	0.096
Survival to 12 mths postoperatively					
Yes	170	1.00		1.00	
No	60	1.19 (0.54 to 2.64)	0.669	0.75 (0.33 to 1.73)	0.507

*All multivariable models are adjusted for surgeon-level fixed effect covariates of patient condition, patient age, and surgeon experience, with a random intercept for surgeon.
†Binary outcomes were analyzed logistic regression models. Continuous outcomes analyzed with linear regression models.
‡Acceptable alignment = ≤ 5° from neutral.
§Mean difference (95% CI)
CI, confidence interval; LLD, leg length discrepancy; OR, odds ratio.

We present training effect estimates from all models (primary analysis) with 95% confidence intervals (CIs). All hypothesis testing was at the 5% level with no adjustments for multiple testing. All analyses were undertaken using statistical software STATA v. 16 (UK).

Results

A total of 40 trainees due to rotate into orthopaedic training posts (PGY 3 to 5 inclusive) in the West Midlands were screened for eligibility and invited to participate (Figure 1). Of these, 33 agreed to participate and were randomized. Of the 12 who were not randomized, eight declined and four did not respond to the invitation. Five trainees withdrew after randomization, leaving 28 participants randomly allocated to receive cadaveric training (n = 13) or standard residency training (n = 15). In total, 11 of 13 participants in the cadaveric group received the training as randomized; two did not, as they were unable to attend the cadaveric training course at short notice and so were switched to the standard training group. Of the 15 participants in the standard training group, 12 received standard training as randomized, and three did not. These three participants were unable to attend the post-trial course offering, provision of which was a condition of ethical approval. A pragmatic decision was therefore taken to switch these three trainees to the cadaveric-trained group so that they could receive a course to meet the obligations of equity-of-access requirement. Therefore, 14 participants received cadaveric training and 14 received standard training. Overall, 24 participants completed ten months' follow-up (Figure 1).

Table I summarizes the baseline characteristics of the participants by randomized group, demonstrating that the two groups were broadly similar. Participants in the intervention group were a mean of 2.7 years younger and had undertaken a mean of 1.7 months less speciality training than the control group. They had performed 6.0 and 3.2 more DHS and hemiarthroplasty procedures at baseline compared to the control group participants. This factor was noted prior to statistical analysis and was adjusted for in the regression model.

Table IV. Results of linear mixed models on dynamic hip screw procedures, under intention to treat and per-protocol approach.

Outcome	Count	Multivariable random effect,* OR (95% CI)			
		Intention to treat	p-value†	Per protocol	p-value†
Radiological					
Tip-apex distance					
≤ 25 mm	257	1.00		1.00	
> 25 mm	26	6.47 (0.97 to 43.05)	0.053	17.93 (3.54 to 90.71)	< 0.001
Position of lag screw in femoral head					
Central inferior†	217	1.00		1.00	
Superior anterior†	71	1.97 (0.91 to 4.30)	0.087	2.98 (1.57 to 5.64)	0.001
Plate flush to lateral cortex					
Yes	181	1.00		1.00	
No	111	1.03 (0.53 to 2.03)	0.923	1.67 (0.91 to 3.06)	0.097
8 cortex screw hold					
Yes	284	1.00		1.00	
No	6	1.03 (0.11 to 9.66)	0.982	1.37 (0.14 to 13.72)	0.787
Non-radiological					
Procedure time (mins)‡		4.82 (-3.71 to 13.35)	0.268	1.15 (-6.57 to 8.87)	0.770
Intraoperative radiation dose (Gym^2)‡		-0.03 (-0.20 to 0.15)	0.775	0.19 (0.05 to 0.33)	0.007
Complications					
Any complication					
No	233	1.00		1.00	
Yes	54	1.77 (0.92 to 3.40)	0.090	1.44 (0.73 to 2.83)	0.294
12-month mortality					
No	267	1.00		1.00	
Yes	21	1.80 (0.70 to 4.61)	0.222	0.67 (0.25 to 1.83)	0.436

*All multivariable models are adjusted for surgeon-level fixed effect covariates of patient condition, patient age, and surgeon experience, with a random intercept for surgeon.
†Binary outcomes were analyzed logistic regression models. Continuous outcomes analyzed with linear regression models.
‡Mean difference (95% CI)
CI, confidence interval; OR, odds ratio.

Primary outcomes. A total of 243 hemiarthroplasty operations were performed by participants during ten months' follow-up, 147 by participants allocated to the intervention and 96 by those allocated to the control group. Anteroposterior radiographs were available for 138 and 86 cases, respectively. A comparison of unacceptable leg length discrepancy (LLD) (> 10 mm) between groups showed there were significantly more unacceptable LLDs in the control group compared to the intervention group. Table II shows that four of 129 (3%) of the hemiarthroplasty patients whose operations were performed by the intervention group had an unacceptable LLD, compared to ten of 85 (12%) in the control group (odds ratio (OR) 4.08 (95% CI 1.17 to 14.22); p = 0.027). There were more varus malaligned implants where alignment was > 5° from neutral in the control group (18/83; 22%) compared to the intervention group (9/127; 7%) (OR 4.73 (95% CI 0.90 to 24.94); p = 0.067) (Table III).

A total of 317 DHS operations were performed by the study participants during ten months of follow-up: 177 by the intervention group and 140 by the control group. Two-view radiographs were available for 174 and 114 cases, respectively. A comparison of 'poor' TAD (> 25 mm) between the groups shows that there were more poorly positioned implants in the control group 15% (19/130) compared to 5% (7/153) in the intervention group (OR 6.47 (95% CI 0.97 to 43.05); p = 0.053). There were more high ± anteriorly placed lag screws in the control group (31%; 41/132) compared to the intervention group (19%; 30/156) but this was not significant (OR 1.97 (95% CI 0.91 to 4.30); p = 0.087) (Table IV).

In the additional analysis under the per-protocol (as-trained) approach, four of 142 (3%) of the hemiarthroplasty patients whose operations were performed by the intervention group had an unacceptable LLD, compared to ten of 72 (14%) in the control group (OR 6.19 (95% CI 1.80 to 21.31); p = 0.004); a larger effect than the equivalent ITT analysis. Likewise, in the DHS setting there were more poorly positioned implants in the control group, 'poor' TAD (> 25 mm); 21% (23/111) compared to 2% (3/172) in the intervention group (OR 17.93 (95% CI 3.54 to 90.71); p < 0.001), again larger than the equivalent ITT analysis (Table IV).

Secondary outcomes. There was no between-group difference seen in complication rate for DHS (p = 0.090) or hemiarthroplasty (p = 0.412), and there was no difference seen for 12-month mortality for DHS (p = 0.222) or hemiarthroplasty (p = 0.669).

The patients of intervention group surgeons stayed an average of 0.6 days less in hospital for hemiarthroplasty compared to patients of control group surgeons (mean intervention group inpatient stay 19.5 days (SD 15.7) vs control group 20.1 days (SD 14.4)). This mean difference from the adjusted model was not statistically significant (p = 0.879).

Mean procedure time for DHS was 7.8 minutes faster in the intervention group (59.4 minutes (SD 18.4)) compared to the control group (67.2 minutes (SD 16.9)) but this was not found to be statistically significant in the adjusted model (p = 0.268). Mean procedure time for hemiarthroplasty was also 14.3 minutes faster in the intervention group (77.4 minutes (SD 27.6)) compared to the control group (91.7 minutes (SD 21.7)) but this was not statistically significant (p = 0.638). There was no significant between-group difference seen in intraoperative radiation dose to patients during DHS (p = 0.775)

The patients of control group surgeons undergoing hemiarthroplasty were significantly more likely to require blood transfusion in the first 48 postoperative hours compared to the patients of intervention group surgeons (12/78 patients (15%) in control group vs 4 of 140 patients (3%) in cadaveric group (OR 6.00 (95% CI 1.83 to 19.69); p = 0.003).

Discussion

We were surprised to find that there were significant differences in the outcomes of the two groups. As this was a preliminary study, we were expecting to learn the practicalities and pitfalls of conducting a real-world study, and possibly be able to estimate the likely training effect of the intervention. Both of these factors influence the planning of further trials and are reflected on below.

In this pragmatic trial, we found that implant position on the first post-implantation radiograph was improved for both DHS and hemiarthroplasty, and that hemiarthroplasty operations performed by cadaveric-trained surgeons had a lower risk of requiring postoperative blood transfusion. If this effect is real, this would be the first randomized study to provide evidence that a training intervention improves patient outcomes in hip fracture operations performed by trainees.

For DHS, there were fewer malpositioned implants for the cases performed by cadaveric-trained surgeons, as measured by TAD and lag screw position in the femoral head. There is a known failure rate for badly positioned implants,[12] so although these were not collected for the patients in our study population, they can be accurately estimated. Device cut-out is a catastrophic complication requiring reoperation, with significant associated risks (including death) to the patient and cost to the health service.[19]

For hemiarthroplasty, there were significantly fewer unacceptable LLDs and fewer varus malpositioned implants in the patients of cadaveric-trained surgeons. LLD > 10 mm is known to be a predictor of postoperative complications in total hip arthroplasty (THA), increasing the risk of pain, premature wear, and failure. Varus malalignment > 5° from neutral in THA increases the risk of dislocation, although there is a complex interaction of factors in determining this risk.[20] Superior implant position should reduce the risk of revision surgery due to recurrent dislocation or failure, which carries risks to the patient and costs to the health service. There are systematic differences in the patient population between the elective arthroplasty and acute trauma settings, hence the generalizability of the THA survival evidence to hemiarthroplasty is unknown.

The improved implant positions seen for both DHS and hemiarthroplasty may be explained by the fact that during the cadaveric training, trainees had the opportunity to repeatedly adjust or redo parts of the procedure until they mastered the technique. They had real-time, personalized feedback from the supervising faculty who could allow them to struggle and learn from mistakes to a far greater extent than would be safe in the real operating theatre. In the cadaveric simulation laboratory, the trainee experiences and learns how to solve problems themselves, without real-life time constraints of the operating theatre or patient safety concerns.[5]

The lower blood transfusion requirement seen in the patients of the cadaveric-trained surgeons may be explained by these operations being performed more expeditiously by the cadaveric-trained surgeons with better implant positioning and superior soft-tissue handling technique. Less soft-tissue trauma and muscle bruising from repeated instrument readjustments, combined with better haemostasis, might mean these patients bleed less intraoperatively, and have a lower postoperative transfusion requirement.

The cost per trainee for the cadaveric training was approximately £1,200. If potentially better quality surgery is being performed, with fewer complications, there are likely to be economic benefits, as well as the health outcome-related benefits to the patient. This is an area worthy of further study.

This study is highly unusual in surgical education research, being randomized and multicentre, increasing both the internal and external validity of the findings. The trial is pragmatic, exploring the effectiveness of cadaveric simulation within the everyday reality of the surgical training environment. The number of participating centres (n = 9) from three large surgical programmes increases the generalizability of the findings to different educational environments.

Limitations of this trial include the fact that participants could not be masked to allocation. Some participants crossed over into the opposite trial arm because of various logistical reasons concerning the participants.

This is a reflection of the challenges of conducting real-world educational research. We were surprised that some trainees would not consent to enrolling in the trial, and further work needs to be done to understand why. We underestimated the problems with releasing the trainees from their place of work for the cadaveric course. We engaged the programme directors in the conduct of the trial, but not the clinical leads of the orthopaedic departments. Further trials will require ensuring these practical difficulties are predicted, and overcome, prior to the trial starting.

As this was a pragmatic trial, we relied on data which were collected in routine training and clinical practice. This includes the trainees self-assessing if the operation was one that they performed, under supervision. For the purposes of this study, we assumed that trainees performed all or a large part of the procedures they classified as such; this self-grading has the potential for bias.

The radiological measurements were made by one of the research team (HKJ) and were not remeasured. The reproducibility of this method has been separately studied and found to be satisfactory. [17]

Of note, the trainees in the intervention group had performed more hip fracture fixations and hemiarthroplasties prior to the course than the control group. This is an important confounder to consider, and randomization should have dealt with it, though it may have been the pragmatic crossovers who introduced an effect. However, there were no measurable differences between those who remained in their allocated groups and those who crossed over, and so the crossover effect is hard to quantify and may not have been significant.

The prior experience of trainees was accounted for in the regression analysis, which gives us some confidence that there is still a measurable training effect, even taking prior experience into account. Whether the increased numbers of DHS and hemiarthroplasty procedures performed by the training group, after the intervention, was a result of the intervention, or prior experience, or both, is unknown. There is likely to be a multiplicative effect, as increased experience usually instils confidence and increases competence and training opportunity.

A statistically significant improvement is seen after undertaking an intention-to-treat (i.e. train) analysis. This, with the general trend of the outcomes tending to favour cadaveric training, encourages us to think that the training effect is probably real, although the magnitude is currently unknown. This should be the subject of further investigation.

Our recommendations for conducting a further randomized educational trial are that 1) the use of randomized trials for complex surgical educational interventions, with the measurement of real-world patient outcomes, is feasible; 2) the trial should be coordinated by a Clinical Trials Unit who have the expertise required to commission, coordinate, and follow up the patient-level effects of complex education interventions; 3) the trial should be coordinated with both programme directors and clinical directors of orthopaedic departments to ensure there are no practical reasons why trainees cannot attend the intervention; and 4) even junior orthopaedic trainees have a significant range of operative experience prior to any educational intervention. This fourth potential confounder should be considered when designing a trial.

The use of simulation for training surgeons is an exciting frontier in surgical education. In a training environment that is increasingly competency-based, and a clinical environment that is increasingly risk-averse, it can efficiently train junior trainees in a safe and targeted way. It is not known, of the wide range of simulation technologies available, which is the most effective. A further non-inferiority randomized trial is needed to compare the effectiveness of low-cost simulation training using plastic bones and medium-cost using virtual reality, with cadaveric simulation training.

In summary, hip fracture is a significant public health burden and most of these major operations will be performed by junior trainees. Cadaveric simulation training may lead to clinically meaningful improvement in implant position for DHS and hemiarthroplasty, and reduce the risk of hemiarthroplasty patients requiring a postoperative blood transfusion. This area requires further study.

Take home message

- Cadaveric simulation training may lead to clinically meaningful improvement in implant position for dynamic hip screw and hemiarthroplasty, and may reduce the risk of blood transfusion requirement in patients undergoing hemiarthroplasty.
- This work indicates that how we train residents has a measurable impact on patients.
- The size and clinical meaning of this impact, and best application of simulation training for hip fracture surgery, requires further research work.

References

1. **Hogg ME**, **Zenati M**, **Novak S**, **et al**. Grading of surgeon technical performance predicts postoperative pancreatic fistula for pancreaticoduodenectomy independent of patient-related variables. *Ann Surg*. 2016;264(3):482–491.
2. **Schmidt CM**, **Turrini O**, **Parikh P**, **et al**. Effect of hospital volume, surgeon experience, and surgeon volume on patient outcomes after pancreaticoduodenectomy: A single-institution experience. *Arch Surg*. 2010;145(7):634–640.
3. **Lee Y-K**, **Biau DJ**, **Yoon B-H**, **Kim T-Y**, **Ha Y-C**, **Koo K-H**. Learning curve of acetabular cup positioning in total hip arthroplasty using a cumulative summation test for learning curve (LC-CUSUM). *J Arthroplasty*. 2014;29(3):586–589.
4. **Gustafsson A**, **Pedersen P**, **Rømer TB**, **Viberg B**, **Palm H**, **Konge L**. Hip-fracture osteosynthesis training: exploring learning curves and setting proficiency standards. *Acta Orthop*. 2019;90(4):348–353.
5. **James HK**, **Pattison GTR**, **Griffin DR**, **Fisher JD**. How does cadaveric simulation influence learning in orthopedic residents? *J Surg Educ*. 2020;77(3):671–682.
6. **No authors listed**. Maximising training: making the most of every training opportunity. Joint Committee on Surgical Training. 2021. https://www.jcst.org/key-documents (date last accessed 31 May 2023).
7. **James HK**, **Chapman AW**, **Pattison GTR**, **Griffin DR**, **Fisher JD**. Systematic review of the current status of cadaveric simulation for surgical training. *Br J Surg*. 2019;106(13):1726–1734.

8. **James HK**, **Pattison GTR**, **Fisher JD**, **Griffin D**. Cadaveric simulation versus standard training for postgraduate TRAUMA and orthopaedic surgical trainees: protocol for the CAD:TRAUMA study multicentre randomised controlled educational trial. *BMJ Open*. 2020;10(9):e037319.
9. **No authors listed**. Specialist Training in Trauma and Orthopaedics - 2015 Curriculum (2018 Update). In: *Training Standards Committee of the British Orthopaedic Association*. Vol 232. London: General Medical Council, 2018.
10. **No authors listed**. Science Care. https://www.sciencecare.com (date last accessed 5 June 2023).
11. **No authors listed**. Intercollegiate Surgical Curriculum Programme. Available from. https://www.iscp.ac.uk (date last accessed 5 June 2023).
12. **Baumgaertner MR**, **Curtin SL**, **Lindskog DM**, **Keggi JM**. The value of the tip-apex distance in predicting failure of fixation of peritrochanteric fractures of the hip. *J Bone Joint Surg Am*. 1995;77-A(7):1058–1064.
13. **Hsu C-E**, **Huang K-C**, **Lin T-C**, **Tong K-M**, **Lee M-H**, **Chiu Y-C**. Integrated risk scoring model for predicting dynamic hip screw treatment outcome of intertrochanteric fracture. *Injury*. 2016;47(11):2501–2506.
14. **Vanrusselt J**, **Vansevenant M**, **Vanderschueren G**, **Vanhoenacker F**. Postoperative radiograph of the hip arthroplasty: what the radiologist should know. *Insights Imaging*. 2015;6(6):591–600.
15. **Munuera L**, **Garcia-Cimbrelo E**. The femoral component in low-friction arthroplasty after ten years. *Clin Orthop Relat Res*. 1992;1992(279):163–175.
16. **Devitt A**, **O'Sullivan T**, **Quinlan W**. 16- to 25-year follow-up study of cemented arthroplasty of the hip in patients aged 50 years or younger. *J Arthroplasty*. 1997;12(5):479–489.
17. **James HK**, **Pattison GTR**, **Griffin J**, **Fisher JD**, **Griffin DR**. Assessment of technical skill in hip fracture surgery using the postoperative radiograph: pilot development and validation of a final product analysis core outcome set. *Bone Jt Open*. 2020;1(9):594–604.
18. **No authors listed**. eLogbook - the Pan-Surgical Electronic Logbook for the United Kingdom & Ireland. https://www.elogbook.org (date last accessed 31 May 2023).
19. **Said GZ**, **Farouk O**, **El-Sayed A**, **Said HG**. Salvage of failed dynamic hip screw fixation of intertrochanteric fractures. *Injury*. 2006;37(2):194–202.
20. **Kunutsor SK**, **Barrett MC**, **Beswick AD**, **et al**. Risk factors for dislocation after primary total hip replacement: meta-analysis of 125 studies involving approximately five million hip replacements. *Lancet Rheumatol*. 2019;1(2):e111–e121.

Author information:
- H. K. James, PhD, MRCS, Orthopaedic Registrar
- D. R. Griffin, MA, MPhil (Cantab), FRCS (Tr&Orth), Professor of Trauma & Orthopaedic Surgery

Department of Trauma & Orthopaedic Surgery, University Hospitals Coventry and Warwickshire, Coventry, UK; Warwick Clinical Trials Unit, Warwick Medical School, Coventry, UK.

- G. T. R. Pattison, MSc(Med Ed), FRCSEd(Tr&Orth), Consultant Orthopaedic Surgeon, Department of Trauma & Orthopaedic Surgery, University Hospitals Coventry and Warwickshire, Coventry, UK.
- J. Griffin, MSc, BSc, Research Associate in Statistics
- J. D. Fisher, PhD, Senior Research Fellow

Warwick Clinical Trials Unit, Warwick Medical School, Coventry, UK.

Author contributions:
- H. K. James: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing.
- G. T. R. Pattison: Conceptualization, Funding acquisition, Investigation, Methodology, Supervision, Validation, Writing – review & editing.
- J. Griffin: Data curation, Formal analysis, Software, Validation, Writing – review & editing.
- J. D. Fisher: Investigation, Methodology, Supervision, Writing – review & editing.
- D. R. Griffin: Conceptualization, Funding acquisition, Investigation, Methodology, Supervision, Writing – review & editing.

Funding statement:
- The authors disclose receipt of the following financial or material support for the research, authorship, and/or publication of this article: Versus Arthritis Educational Research Fellowship (Grant number 20485), Health Education West Midlands Faculty Development Grant and CVC Foundation.

ICMJE COI statement:
- H. K. James reports a Versus Arthritis Educational Fellowship Award for their PhD fellowship salary, and institutional grants from Health Education West Midlands and CVC Foundation for this study. D. R. Griffin reports a PhD studentship paid to the University of Warwick from Versus Arthritis for this study, and institutional research grants from NIHR and royalties from BrainLab, unrelated to this study.

Data sharing:
- The datasets generated and analyzed in the current study are not publicly available due to data protection regulations. Access to data is limited to the researchers who have obtained permission for data processing. Further inquiries can be made to the corresponding author.

Acknowledgements:
- The study team wish to thank senior medical statistician Dr Nick Parsons for preparing the randomization list.

Trial registration number:
- Current Controlled Trials ISRCTN20431944; trial registry URL: https://www.isrctn.com

Open access funding:
- The authors confirm that the open access fee for this article was funded by Versus Arthritis.