

Surgical Crisis Management Skills Training and Assessment

A Stimulation-Based Approach to Enhancing Operating Room Performance

Krishna Moorthy, MD, FRCS, Yaron Munz, MD, Damien Forrest, PhD,
Vikas Pandey, BSc, MBBS, FRCS, Shabnam Undre, FRCS, Charles Vincent, PhD,
and Ara Darzi, MD, FRCS, FACS

Background: Intraoperative surgical crisis management is learned in an unstructured manner. In aviation, simulation training allows aircrews to coordinate and standardize recovery strategies. Our aim was to develop a surgical crisis simulation and evaluate its feasibility, realism, and validity of the measures used to assess performance. **Methods:** Surgical trainees were exposed to a bleeding crisis in a simulated operating theater. Assessment of performance consisted of a trainee's technical ability to control the bleeding and of their team/human factors skills. This assessment was performed in a blinded manner by 2 surgeons and one human factors expert. Other measures consisted of time measures such as time to diagnose the bleeding (TD), inform team members (TT), achieve control (TC), and close the laceration (TL). Blood loss was used as a surrogate outcome measures.

Results: There were considerable variations within both senior (n = 10) and junior (n = 10) trainees for technical and team skills. However, while the senior trainees scored higher than the juniors for technical skills ($P = 0.001$), there were no differences in human factors skills. There were also significant differences between the 2 groups for TD ($P = 0.01$), TC ($P = 0.001$), and TL (0.001). The blood loss was higher in the junior group.

Conclusions: We have described the development of a novel simulated setting for the training of crisis management skills and the variability in performance both in between and within the 2 groups.

(*Ann Surg* 2006;244: 139–147)

From the Department of Surgical Oncology and Technology, Imperial College, London.

Supported in part by the BUPA Foundation.

Members of the Imperial College-St. Mary's Hospital Simulation Group are as follows: Shirley Martin, Lorraine Poore, Lee Edwards, John Abbot, Janet Henry, Benny Lo, and Prof. G. Z. Yang.

Reprints: Krishna Moorthy, MD, 28 Carless Avenue, Birmingham, B17 9EQ, UK. E-mail: k.moorthy@imperial.ac.uk.

Copyright © 2006 by Lippincott Williams & Wilkins

ISSN: 0003-4932/06/24401-0139

DOI: 10.1097/01.sla.0000217618.30744.61

The incidence of iatrogenic trauma to organs or major vessels during surgery depends to a large extent on the complexity of the procedure, the proximity of these structures to the operative field, and the expertise of the surgeon. While surgical training places a strong emphasis on the importance of avoiding such complications, trainees are seldom taught how to manage these crisis events when they occur. As successful management depends on the crisis management skills of the surgeon and the surgeon's experience, outcomes can often be varied. In addition, the management of these situations is based on anecdotal experience, from watching a mentor manage a similar situation, or from one's own personal experience where the surgeon may even have learned by making mistakes. Such learning is thus random and is based on the need to have faced similar situations previously.

Simulations offer the advantage of allowing a trainee to learn the consequences of a mistake and acquire the knowledge to manage a crisis without any harm to a patient or a plane load of passengers, with the benefit of performance feedback.¹ In aviation, simulator based training can be used to train flight crews to manage crisis scenarios making best use of the available resources. This is one of the key aspects of Crew Resource Management (CRM) training.² Repeated practice equips the pilot and the flight crew with the knowledge and the skills to manage these situations when they occur in real life. Anesthesia CRM (ACRM)³ is based on a similar principle and allows anesthetists to learn the skills of dealing with both common and rare crisis scenarios. The emphasis of both CRM and ACRM training is not only on technical skills but also on team skills such as communication, vigilance, and decision making.⁴ The aim is to train the individual in team skills as well as train teams to work together in the most effective manner.

CRM training, in its present form, has error management as its primary focus.⁵ Helmreich et al have described error management as an error troika.⁵ This consists of *error avoidance* where safety principles are used to prevent the occurrence of errors, *error trapping* where errors are trapped before they occur often by vigilance and team monitoring, and *error mitigation* where errors are rapidly diagnosed and effectively managed before they give rise to drastic consequences. CRM training in its present genera-

tion (sixth) “stresses the fact that effective error management is the hallmark of effective crew performance and that well managed errors are the indicators of effective performance.”⁵

From a surgical context, de Leval et al have suggested that error recovery strategies are just as important as error prevention measures.⁶ Their observations of major and minor errors during a pediatric cardiac procedure indicate that “the surgeon’s diagnostic skill, knowledge of the various surgical strategies to correct a problem, and communication with the rest of the team are important prerequisites of error compensation.” They described error compensation as a form of error recovery whereby “a strategy to remedy the situation is implemented before negative consequences ensue.” Just as in aviation, error management is a crucial aspect of surgical performance. Surgeons and OR teams that are superior in error recovery are more likely to have better patient outcomes.

The development of anesthesia simulators has made ACRM courses feasible. Since the development of the first anesthetic simulator, Sim One,⁷ further technological advances have resulted in the development of higher fidelity simulators. There are now simulators that provide physiologic and pharmacologic responses to drugs and preprogrammed scenarios.⁸ In comparison, it is difficult to simulate surgical procedures and more importantly simulate surgical crises. VR simulators are presently task trainers and many synthetic models lack tissue realism. Animal models may provide the opportunity to develop crisis situations such as vascular trauma, but there would be considerable ethical problems and their use is banned in countries like the United Kingdom. However, with increasing interest in the use of synthetic models for the training and assessment of technical skills,⁹ some groups across the world have taken the initiative to work closely with their industry partners to enhance the realism of such models. This study describes the development of a surgical crisis simulation using a synthetic model and the assessment of its management from a technical and nontechnical perspective. The aims of this study were to:

1. Establish the face validity (realism) of a crisis simulation for surgical trainees
2. Establish the construct validity (difference between groups) and reliability of the measures used in the assessment of performance and
3. Assess participants’ experience of the simulation.

MATERIALS AND METHODS

Subjects

There were 20 surgeons who were divided into 2 groups:

Group 1 (senior trainees): performed >50 saphenofemoral high tie procedures (n = 10)

Group 2 (junior trainees): performed 20 to 50 procedures (n = 10)

Trainees who had performed less than 20 procedures were excluded from the study as it was assumed that they would not possess the skills or the knowledge to handle the bleeding crisis. The groups reflect the average number of

procedures performed by 3 groups of surgical trainees in the United Kingdom, namely, basic surgical trainees (BSTs), junior higher surgical trainees (HSTs), and senior HSTs.

Study Design

The 2 groups of trainees were assessed during the performance of a saphenofemoral junction (SFJ) high-tie procedure on a synthetic bench model, which has been validated by a previous study (Limbs and Things, Bristol, UK).¹⁰ The 2 groups of trainees were compared for their technical and nontechnical skills using various measures described later.

Simulated Operating Theater (SOT)

The Physical Environment

A simulated operating theater (SOT) has been described in detail in our earlier article.¹¹ It was developed with the aim of replicating a real operating theater as closely as possible with an adjacent room serving as a control room. A novel technology called the Clinical Data Recorder (CDR) was developed to evaluate the technical skills of the surgeon as well as the fine nuances of communication and interaction between the personnel. The researchers and trainers view the proceedings in the SOT on a monitor placed in the control room. Images from ceiling-mounted cameras in the SOT are initially fed into a video-mixer before being viewed on the monitor. The images on the monitor thus depend on the number and the order of the cameras selected. For the purposes of this project, we used only 3 cameras, the input from which was fed into the CDR. The fourth image is the feed from the anesthetic monitor, which is streamed into the CDR along with the images from the cameras. This allows the assessors to view the patient’s parameters in synchronization with the occurrences in the SOT. The researcher/trainers are able to view the simulation in real time as well as record the simulation onto a DVD disc for future review and evaluation.

The Equipment

In addition to all the standard operating theater equipment, one of the primary features of the SOT is an anesthetic simulator (SimMan, Laerdal, UK). This is a simulator of moderate fidelity, which allows manipulation of the mannequin’s hemodynamic parameters through software installed on a notebook computer located in the control room. The simulator’s hemodynamic data are displayed on a monitor fixed to the anesthetic machine. The hemodynamic measures were adjusted by the research coordinators depending on the extent of blood loss according to a predetermined protocol.

The Procedure

The synthetic model of the SFJ was attached to the anesthetic simulator (mannequin). This is a silicon-based model with fairly accurate simulation of the SFJ with a layer of simulated skin overlying a layer of superficial fascia. It consists of a “saphenous vein” with 4 tributaries connected to a “femoral” vein. These are set in a cast of silicon that resembles fat. The vein has an elasticity that feels comparable to that of a real vein and one the main advantages is that sutures to close a laceration and ties onto tributaries do not cut through the silicon. The model was then draped with

surgical drapes and held with surgical towel clips. The trainees were requested to perform the simulated procedure in the SOT.

The Simulation

A standardized theater team consisting of an anesthesiologist, an operating department assistant (ODA) who is primarily an assistant to the anesthesiologist, a scrub nurse, a “circulating” nurse (assistant to the scrub nurse), and an assistant, first entered the SOT and assumed their positions. The surgeons entered the simulation area through the control room adjacent to the operating theater. They were briefed about the scenario and made familiar with the theater environment through the one-way glass. They were then asked to sign a consent form for participating in the study with a clause requesting their confidentiality. They then entered the operating theater having gowned and gloved.

Crisis Scenario

A 5-mm laceration was made in the medial aspect of the “femoral vein” in the model, which was then secured over the right “groin” of the mannequin (anesthetic simulator) (Fig. 1). The “femoral vein” was then connected to a tube, which ran along the walls and the floor of the SOT and was concealed from the surgeon’s view. This tubing was then connected to an intravenous line, which was in turn connected to a saline bag filled with simulated blood (Annexe Art, UK).

At a standardized point during the procedure, the simulation controller started the bleeding by opening up the IV tubing connected to saline bag. The pressure on the bag was strictly controlled and maintained at 20 mm Hg with the aid of a level one rapid transfuser (Level 1 Technologies, UK). The bleeding was maintained till the following factors gave the impression that the laceration had been effectively closed:

1. The slowing down of the simulated blood appearing through the surgical wound
2. The assessment of control by viewing the procedure on the monitor placed in the control room and concurrence with

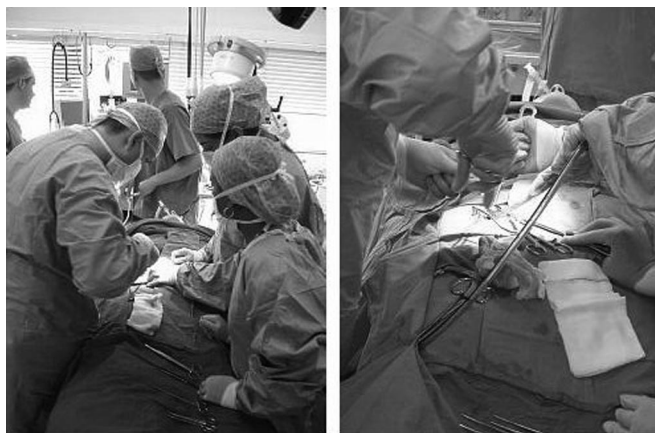


FIGURE 1. Simulated operating theater with team controlling bleeding from the femoral vein.

3. The simulation was stopped under the following circumstances:
 1. Bleeding effectively controlled
 2. Trainee called a senior surgeon/vascular surgeon due to inability to control the bleeding or refusal to do so due to lack of confidence.

Feedback

The participants received an objective, video-based, criteria-referenced feedback of their performance either soon after the simulation or within 2 weeks. The technical feedback was provided by the research fellow with reference to the opinion of 5 experts (vascular surgeons) who were consulted for the development of the assessment measures. The experts were asked to highlight the various measures that they believed that a trainee should take for the successful management of a femoral vein laceration. A human factors researcher provided the nontechnical feedback.

Assessment of Performance

In addition to assessing performance during the simulations for the purpose of the feedback, all the sessions were recorded onto DVD and played back to the assessors. The data from these assessments were then used as the study data.

Technical Ability to Control the Bleeding/Repair the Laceration

A global rating scale, based on the global rating scale developed by Reznick et al¹² for the assessment of generic technical skills, was developed for the assessment of the ability of the participants to control the bleeding and repair the laceration (Table 1). The scale was developed after consulting the 5 experts who were asked to identify all the measures required for the successful management of a femoral vein laceration. Two blinded surgical observers (surgical fellows) performed this assessment. The score was expressed as a percentage.

Nontechnical Skills

The NOTECHS rating scale developed for assessment of nontechnical skills in aviation¹³ was modified for the assessment of the nontechnical skills of the surgeons (Table 2). The scale is a minor modification of the one published in our earlier study.¹¹ The scale consisted of 4 categories for the assessment of nontechnical skills, and there was an additional category for the assessment of decision-making/crisis handling (DM). Each of these categories consisted of 3 to 5 elements. The elements were rated on a 6-point scale. The score for each category and the total score were then expressed as a percentage. A human factors researcher and the surgical fellow who was trained by the former performed the assessment. The training and calibration consisted of the first 5 simulations being observed together. The other simulations were all assessed independently to facilitate evaluation of interrater reliability.

TABLE 1. Rating Scale for Technical Ability to Control the Bleeding

	1	2	3	4	5
Identification of etiology	1 Does not or cannot identify source of bleeding	2	3 Recognizes severity but has some difficulty in localizing source	4	5 Early and accurate diagnosis of severity; uses suction; performs an organized check efficiently
Exposure	1 Inappropriate exposure; inadequate use of retractors and/or assistants	2	3 Adequate incision, however repeatedly adjusting retractors	4	5 Appropriate incision (extended if necessary); efficient use of retractors and assistants
Control of bleeding	1 Doesn't control bleeding or uses hazardous maneuvers to control bleeding	2	3 Adequate measures demonstrated but sometimes loses control	4	5 Proximal and distal control with finger pressure or vascular clamps
Clear vision of laceration	1 Does not adequately expose laceration prior to closure; vision obscured by blood/instruments	2	3 Maintains clear vision with some difficulties	4	5 Maintains dry field and clearly identifies laceration prior to closure
Technique of closure	1 Poor technique for closure resulting poor technical result	2	3 Adequate venotomy closure technique	4	5 Continuous sutures starting at apex, adequate and well-spaced bites
Knowledge on instruments	1 Uses the wrong instruments or inappropriate use of instruments	2	3 Familiar with all instruments and uses them appropriately	4	5 Demonstrates clear knowledge of all the correct instruments, eg, vascular clamps, needle holder and sutures
Overall performance	1 Unsure how to deal with this lesion, hesitant and does not mobilize assistance; likely to result in poor outcome	2	3 Adequate identification and management of hemorrhage with some difficulties encountered	4	5 Efficiently and confidently identifies and deals with hemorrhage without complication

Each component is scored on a 5-point scale. The two extremes and the middle are explained by clear descriptions.

Communication Count and "Utterance" Frequency (UF)

A communication utterance frequency was carried out. This method is similar to an approach used in an earlier study observing communication in the operating theater.^{11,14} The utterance frequency is defined as the number of episodes of communication per minute.

Objective Time-Measures of Crisis Management

These were time taken to diagnose the cause for the bleeding as the femoral vein laceration (TD), time taken to informing the team (TT), time taken to calling the consultant (attending)/ help (TH), time to use appropriate measures to control bleeding (TC) successfully and time to close the laceration (TL).

TABLE 2. Nontechnical Skills Assessment

Category	Element
Communication and interaction	(a) Instructions to assistant/scrub nurse; clear and polite (b) Awaits acknowledgment from the assistant/scrub nurse (c) Assistance sought from team members
Vigilance/situation awareness	(a) Monitored patient's parameters throughout the procedure (b) Awareness of anesthetist (c) Actively initiates communication with anesthetist during crisis periods
Team skills	(a) Maintains a positive rapport with the whole team (b) Open to opinions from other team members (c) Acknowledges the contribution made by other team members (d) Supportive of other team members
Leadership and management skills	(a) Adherence to best practice during the procedure, eg, does not permit corner cutting by self or team (b) Time management eg appropriate time allocation without being too slow or rushing team members (c) Resource utilization, ie, appropriate task-load distribution and delegation of responsibilities (d) Authority/assertiveness
Decision-making crisis	(a) Prompt identification of the problem (b) Informed team members; promptly, clearly, and to all team members (c) Outlines strategy/institutes a plan, ie, asks scrub nurse for suction, instruments, suture material (d) Anticipates potential problems and prepares a contingency plan, eg, asks anesthetist to order blood, calls for help (e) Option generation; takes the help of the team (seeks team opinion)

Time taken to calling for help was converted to a score of 1 or 0. From the discussions with the 5 experts it emerged that all junior trainees should call the consultant immediately on detection of the femoral vein laceration. On the other hand, they thought that senior trainees could proceed to close the laceration but should call for help if the blood loss was greater than 500 mL, if the trainee had attempted closure once without success, and if there was any concern about the “patient’s” condition. Thus, for the junior trainees, a score of 1 was given only if help was requested (calling a consultant) on diagnosis of the laceration. If this was not done or done late, they scored 0. Senior trainees who did not call the consultant when necessary (as mentioned above) scored 0.

Surrogate Outcome Measures

In the absence of any definite outcome measures, the total blood loss was used as a surrogate outcome measure. This was calculated by measuring the blood in the suction apparatus and by weighing the swabs.

Assessment of Participants’ Perception

A questionnaire was designed to ask the participants about their simulation experience. The participants were asked to respond to 11 statements (Table 3) with the intent of evaluating the face validity (realism) of the model, the simulation environment and the bleeding, their perception of the value of the simulation as a training/assessment exercise, and their perceived benefit of the feedback. They were asked to mark their answers on a 6-point Likert scale. Scores between 1 and 3 were considered to be disagreements, and those between 4 and 6 to be agreements. The scores of the junior and senior trainees were also analyzed separately.

Data Analysis

As the data were not found to have a normal distribution for most measures, nonparametric tests have been used for analysis. Fisher exact test was used to analyze the differences between the questionnaire results of the 2 groups of trainees (seniors and juniors) and analysis of the difference between the 2 groups for calling the consultant. Mann-Whitney *U* test was used to test the differences between the 2 groups for the various measures.

Interrater reliability between the observers was analyzed using the Intraclass efficient. Cronbach’s alpha was also used to assess the internal consistency of the elements in the nontechnical skills rating scale.

RESULTS

Table 3 gives details of the median scores for the questionnaire categorized as face validity (realism), usefulness for training, usefulness for assessment, and the benefit of feedback. There were consistently high levels of agreement for all the questions. A majority of the participants found the model, the SOT environment, and the bleeding scenario to be realistic; 95% of the participants found that simulation suitable for the training of technical skills and 85% of them found the simulation suitable for team skills training. A majority of the participants considered the crisis simulation suitable for the assessment of their technical and team skills. All the participants found the feedback useful. When the scores for the individual questions were compared between the 2 groups, there were no significant differences for any of the questions except for the fact that junior trainees found the technical skills feedback more useful than the senior trainees ($P = 0.04$).

Performance of the Subjects and Variability Between and Within the Groups

There was a significant difference between the 2 groups for their technical ability in controlling the bleeding. The senior trainees scored significantly higher than the juniors (median \pm interquartile range, 68.5 ± 13 vs. 51.8 ± 14.9 ; $P = 0.001$). From Figure 2A, it can be seen that there was considerable variability within both the groups. Table 4 compares the nontechnical skills of the 2 groups. The only significant difference between the groups was for UF ($P = 0.02$). Although there was a difference for leadership, this was not significant ($P = 0.07$). Examining Figure 2B reveals that there was a large variability within the groups for nontechnical skills. The variability within group 1 was particularly high for decision-making and UF (Fig. 2C, D). There was a significant difference between the 2 groups for TD ($P = 0.01$), TC ($P = 0.001$), TL ($P = 0.001$), and TH ($P = 0.007$). The difference was not significant for TT ($P =$

TABLE 3. The Median Values of the Statements Scored on 6-Point Likert Scales

Category	Statement	Median Score (Maximum Score = 6)	Agreements (%)
Face validity	The model is a realistic representation of the real procedure	4.5 (1)	95
	The SOT is a realistic representation of a real operating theatre	5 (0)	100
	The bleeding is a realistic simulation	4 (1)	90
	I would behave in the same way even in real life	5 (1)	100
Use for training	The crisis simulation is a good training opportunity for training technical skills	5 (1)	95
	The crisis simulation is a good training opportunity for training team skills	5 (1)	85
Use for assessment	The simulation is a good method to assess my technical skills	4.5 (1)	85
	The simulation is a good method to assess my team skills	5 (0.75)	95
Repeat the simulation	I would benefit by repeating the simulation again	5 (1.5)	90
Feedback	I found that technical skills feedback useful	5 (1.5)	100
	I found the team skills feedback useful	6 (1)	100

The values in parentheses are the interquartile range. The last column describes the percentage of subjects who agreed (score 4–6 on the 6-point scale) with the statement.

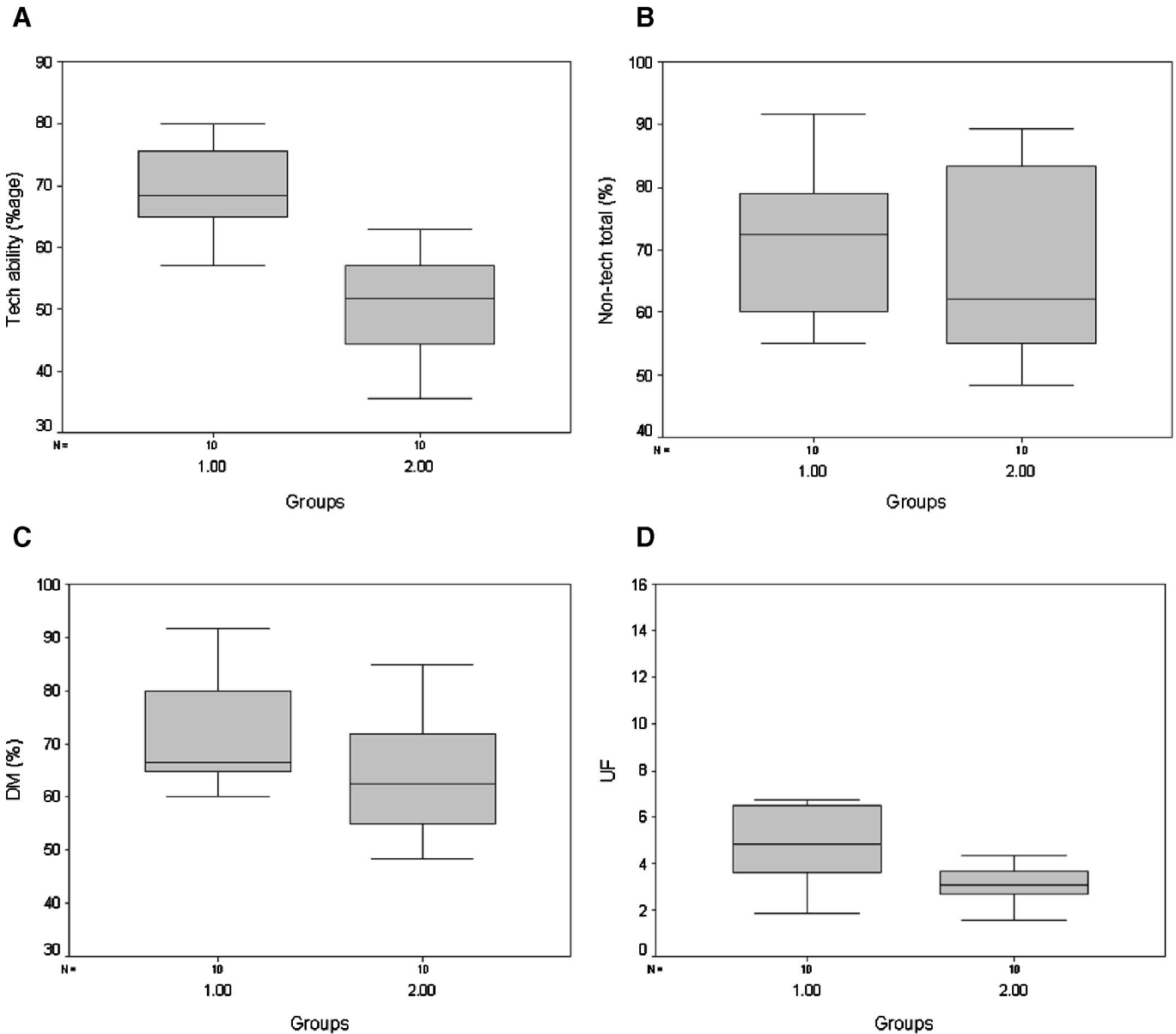


FIGURE 2. Box-plots of technical skills, nontechnical total score, decision-making (DM) score, and utterance frequency (UF) score. %, percentage score.

TABLE 4. The Median Values of the Scores for Nontechnical Skills, DM, and UF

	Group 1	Group 2	P
Nontechnical score (total) (%)	72.4 (22.0)	62.2 (30)	0.30
Communication (%)	60.4 (31.2)	60.4 (34.3)	0.82
Vigilance (%)	80.0 (32.6)	55.0 (43.0)	0.24
Team skills (%)	62.5 (18.2)	64.5 (33.3)	0.90
Leadership (%)	82.3 (9.3)	72.9 (13.0)	0.07
DM score (%)	66.6 (17.1)	62.5 (18.3)	0.24
UF	4.8 (3.2)	3.1 (1.2)	0.02

Values in parentheses are the interquartile range. The last column is the level of significance using Mann-Whitney U test to compare the 2 groups.

0.25). From Figure 3, it can be seen that there was considerable variability within group 2 for all these measures. There was considerable variability within group 1 for TT.

There was a significantly greater blood loss associated with group 2 as compared with group 1 (median ± interquartile range, 600 mL ± 350 vs. 300 ± 125 mL; *P* = 0.02). There was also considerable variability especially within group 2.

Reliability of the Assessment Measures

The internal consistency between the 5 categories of the nontechnical skills assessment score was 0.87. This means the categories were related and were measuring the same general construct, nontechnical skills.

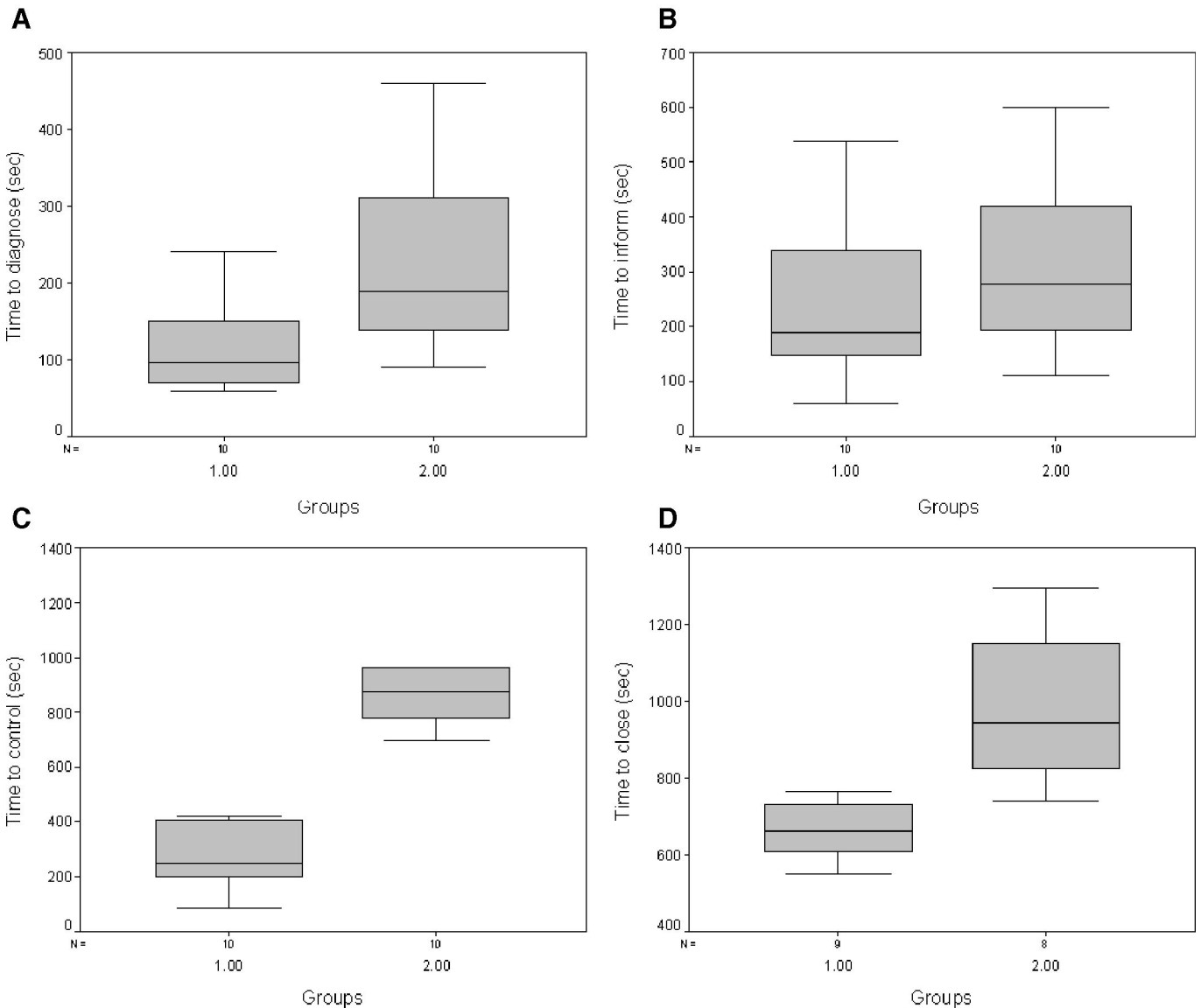


FIGURE 3. Box-plots of crisis management time measures.

High levels of agreement between the different observers were achieved. The intraclass coefficient (interrater reliability) between the 2 surgeons observing technical performance was 0.83 for technical ability to control the bleeding. The interrater reliability for the nontechnical skills was 0.87 for the first 5 simulations that were assessed together for calibration purposes and 0.84 for the others.

Additional Observations

In addition to the various measures used for the assessment of performance, the observers also made some additional observations that were considered to denote unsafe or safe practices. These observations were considered useful in helping understand the variability in performance of the subjects in between and within the groups.

Examples of Unsafe Practice

One junior trainee applied a traumatic clamp (nonvascular) on the femoral vein which led to a further tear in the femoral vein and another junior trainee applied traumatic artery forceps to the edges of the laceration to control the bleeding. Both trainees called for help only after executing these unsafe steps.

Four junior trainees showed evidence of tunnel vision. They were so focused on closing the laceration that they did not pay any attention to the patient's condition or on communication with anesthesiologist. In all 4 instances, there was a blood loss in excess of 500 mL.

One junior trainee did not call for help despite being unable to control bleeding with 2 sutures, which cut through the vein wall and compounded the situation.

Examples of Safe Practice

One senior and one junior trainee acknowledged their limitation at closing the laceration and called for help but only after using effective measures to control the bleeding. Incidentally, both of them scored high for the few parameters that could be assessed for technical ability to control the bleeding.

Monitoring

Only 6 subjects asked about the extent of blood loss once. There were 4 instances where the surgeons assumed that the blood loss was much lesser than what had been measured without obtaining confirmation.

There was a great variability in the number of times the surgeons asked about the patient's condition. This varied from 0 to 7 times during the bleeding. Median and interquartile ranges for the senior trainees was 2 (3) and for juniors was 2 (2.5).

DISCUSSION

This study has demonstrated the face validity or realism of a novel training and assessment environment for surgical trainees, which integrates technical skills with other skills crucial for effective performance during a surgical crisis simulation. All the participants found the SOT environment to be realistic, and more than 90% found the model and the bleeding scenario to be realistic as well.

This study has also established the construct validity of the simulated environment and the bleeding crisis by demonstrating a difference between senior and junior trainees. In addition to the fact that there was a difference between the 2 groups for the technical ability to control the bleeding, the performance of the senior trainees was superior even in terms of the crisis management time measures. They were faster at diagnosing the problem, using effective control measures, and closing the laceration. This is similar to the findings from simulations in anesthesia, which found that time to diagnose a problem and institute remedial measures was faster among senior anesthesiologists as compared with juniors.^{15,16}

The performance of the senior trainees was also superior in terms of patient safety compared with the juniors. Junior trainees were more likely to perform actions detrimental to patient safety such as the blind and inappropriate use of traumatic clamps, and they were also more likely to focus on closing the laceration at the cost of blood loss. They were also more likely to fail to recognize their limitations and call for help. Other than using the simulation just as an opportunity to teach trainees the skills to control major venous hemorrhage, it can even be used to teach them consequences of their unsafe or inappropriate actions. Indeed, some consider the main advantage of simulations to be the ability to make mistakes and learn from them¹⁷ without actually causing any harm to patients.

The results of the study also reflect that observed variations in performance both between and within the groups can be used for setting standards. The findings of variability in performance among both groups of trainees is similar to the findings of simulation-based performance among anes-

thetists.¹⁶ The true value of simulations probably lies in the ability to use them to identify those trainees that need further training to bring their performance to expected standards. Other high reliability organizations such as aviation and the military practice crisis handling "beyond the attainment of a criterion level of performance" to a stage of "overlearning" where responses become automated.¹⁸ Such a strategy may be useful in surgical training as well to ensure a consistently high level of performance among surgeons with respect to their technical and nontechnical skills under stressful conditions.

We have used blood loss as a surrogate outcome measure of performance. A few previous studies have shown that surrogate outcome measures reflect technical expertise. Datta et al used leak rates after a bench model vascular anastomosis and showed a strong correlation between skill and outcome.¹⁹ Similarly, Szalay et al showed the correlation between OSATS and outcome analysis by using a 4-component, 5-point scale to assess the final product.²⁰ Our study has shown that there was a significant difference between the 2 groups. However, of greater importance is the variability within the junior group.

This study has also indirectly demonstrated the benefit of the simulation by assessing the perceptions of the participants. A majority of the trainees considered the simulation to be useful for the training and assessment of their technical and nontechnical skills and found the feedback following the simulation beneficial. Participants' assessment of simulation-based training is considered to be a powerful method of assessing the value of such training in aviation and anesthesia.^{21,22} This is especially important as studies to demonstrate a transfer of skills from such complex simulations to real procedures is difficult and will take a great amount of time and resources.

In addition to assessing technical performance, this study has also attempted to place some emphasis on nontechnical skills such as communication, team coordination, and decision-making. The assessment of nontechnical skills during anesthesia simulations has recently become a major focus. Gaba et al⁴ modified the LOSA checklist developed by NASA and the University of Texas Human Factors Project for the assessment of flight crews. They used 10 crisis management behavioral markers: orientation to case, inquiry/assertion, communication, feedback, leadership, group climate, anticipation/planning, workload distribution, vigilance, and reevaluation. Fletcher et al developed and validated an Anesthesia Non-Technical Skills (ANTS).²³ Our study has certain limitations in the assessment of nontechnical skills that will need to be addressed by further research. The absence of construct validity of the nontechnical skills score could be because the nontechnical assessment scale was a modification of a tool used for the assessment of skills in aviation, with little relevance to surgery. However, further research will be required to evaluate the content validity of the assessment scale, probably by using a task analysis approach similar to Fletcher et al or using a Delphi-type questionnaire. However, the absence of a difference between the 2 groups could even be a result of the wide variability within the groups. This is probably because these skills have never been the focus of surgical training, and their development depends on a number of complex factors such as mentoring, culture, personality, and exposure to positive role models. It is also important to appre-

ciate that there is very little evidence on the construct validity of nontechnical skills, even in aviation and anesthesia.^{4,13} Further research will also be required to assess which elements on the nontechnical scale measure the same feature and thus reduce the number of elements by using a method such as factor analysis or principal component analysis.²⁴

The differences between the 2 groups for most measures are most likely to be secondary to the increased experience of the senior group. It is possible that this experiential learning could to a certain extent be acquired in the simulated environment with a potentially better response in a real setting. In other words, it is likely that skills learned in the simulated setting could be transferred to real situations. This will have to be explored by further research.

This is the first study to have extended the concept of crisis simulation training to surgery by describing the development of a simulated environment for the training and assessment of crisis management skills among surgeons. There is a potential to use such an environment to teach safety principles such as the recognition of one's limitations and the importance of monitoring blood loss and the patient's condition during the management of a surgical crisis. The study has also drawn focus on the need to make performance during crisis situations criteria and protocol based. This would enable performance between surgeons to become consistent and establish a benchmark that trainees have to reach, aid team coordination and decision-making, and potentially improve patient outcome.

ACKNOWLEDGMENTS

The authors thank all the surgical trainees who participated in the study; Mr. S. Bann, Mr. V. Datta, Dr. S. Mackay, and Dr. R. Kneebone with help in the development of the Simulated Operating Theater; and Mr. Nick Gerolemou of Limbs and Things, Bristol, U.K. for his help in the development and modification of the SFJ model.

REFERENCES

- Helmreich RL. On error management: lessons from aviation. *BMJ*. 2000;320:781-785.
- Orlady HW, Foushee HC, eds. *Cockpit Resource Management: Background and Overview*. CA: Moffet Field, 2003.
- Howard SK, Gaba DM, Fish KJ, et al. Anesthesia crisis resource management training: teaching anesthesiologists to handle critical incidents. *Aviat Space Environ Med*. 1992;63:763-770.
- Gaba DM, Howard SK, Flanagan B, et al. Assessment of clinical performance during simulated crises using both technical and behavioral ratings. *Anesthesiology*. 1998;89:8-18.
- Helmreich RL, Merritt AC, Wilhelm JA. The evolution of Crew Resource Management training in commercial aviation. *Int J Aviat Psychol*. 1999;9:19-32.
- de Leval MR, Carthey J, Wright DJ, et al. Human factors and cardiac surgery: a multicenter study. *J Thorac Cardiovasc Surg*. 2000;119(4 Pt 1):661-672.
- Abrahamson S, Denson JS, Wolf RM. Effectiveness of a simulator in training anesthesiology residents. *J Med Educ*. 1969;44:515-519.
- Murray DJ. Clinical simulation: technical novelty or innovation in education. *Anesthesiology*. 1998;89:1-2.
- Macintyre IM, Munro A. Simulation in surgical training. *BMJ*. 1990;300:1088-1089.
- Datta V, Bann S, Beard J, et al. Comparison of bench test evaluations of surgical skill with live operating performance assessments. *Br J Surg*. 2002;89(suppl 1):92.
- Moorthy K, Munz Y, Adams S, et al. A human factors analysis of technical and team skills among surgical trainees during procedural simulations in a Simulated Operating Theatre (SOT). *Ann Surg*. 2005;242:631-639.
- Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. *Br J Surg*. 1997;84:273-278.
- Avermate van JAG, Kruijssen EAC. NOTECHS: the evaluation of non-technical skills of multi-pilot air crew in relation to the JAR-FCL requirements. Research report for the European Commission. DGVII NLR-CR-98443. 1998.
- Calland JF, Guerlain S, Adams RB, et al. A systems approach to surgical safety. *Surg Endosc*. 2002;16:1005-1014.
- Byrne AJ, Jones JG. Responses to simulated anaesthetic emergencies by anaesthetists with different durations of clinical experience. *Br J Anaesth*. 1997;78:553-556.
- Gaba DM, DeAnda A. The response of anesthesia trainees to simulated critical incidents. *Anesth Analg*. 1989;68:444-451.
- Jha AK, Duncan BW, Bates WB. Simulator-based training and patient safety. <http://www.ahcpr.gov/clinic.ptsafety/chap45.htm>. 2003.
- Workload measurement. In: Farmer E, van Rooij J, Riemersma J, et al, eds. *Handbook of Simulator-Based Training*. Aldershot, UK: Ashgate, 1999:261-268.
- Datta V, Mandalia M, Mackay S, et al. Relationship between skill and outcome in the laboratory-based model. *Surgery*. 2002;131:318-323.
- Szalay D, MacRae H, Regehr G, et al. Using operative outcome to assess technical skill. *Am J Surg*. 2000;180:234-237.
- Helmreich RL, Foushee HC. Why cockpit resource management? Empirical and theoretical bases of human factors in training and aviation. In: Wiener E, Kanki BG, Helmreich RL, eds. *Cockpit Resource Management*. San Diego: Academic Press, 1993:3-45.
- Holzman RS, Cooper JB, Gaba DM, et al. Anesthesia crisis resource management: real-life simulation training in operating room crises. *J Clin Anesth*. 1995;7:675-687.
- Fletcher G, Flin R, McGeorge P, et al. Anaesthetists' Non-Technical Skills (ANTS): evaluation of a behavioural marker system. *Br J Anaesth*. 2003;90:580-588.
- Le May S, Dupuis G, Harel F, et al. Clinimetric scale to measure surgeons' satisfaction with anesthesia services. *Can J Anaesth*. 2000;47:398-405.