Live tissue versus simulation training for emergency procedures: Is simulation ready to replace live tissue?

Stephen L. Barnes, MD,[a] Alex Bukoski, DVM, PhD,[b] Jeffrey D. Kerby, MD, PhD,[c] Luis Llerena, MD,[d] John H. Armstrong, MD,[d] *and* **Catherine Strayhorn, BS,[e] the University of Missouri Combat Casualty Training Consortium[a],** *Columbia, MO, Birmingham, AL, and Tampa and Sarasota, FL*

Background. *Training of emergency procedures is challenging and application is not routine in all health care settings. The debate over simulation as an alternative to live tissue training continues with legislation before Congress to banish live tissue training in the Department of Defense. Little evidence exists to objectify best practice. We sought to evaluate live tissue and simulation-based training practices in 12 life-saving emergency procedures.*
Methods. *In the study, 742 subjects were randomized to live tissue or simulation-training. Assessments of self-efficacy, cognitive knowledge, and psychomotor performance were completed pre- and post-training. Affective response to training was assessed through electrodermal activity. Subject matter experts gap analysis of live tissue versus simulation completed the data set.*
Results. *Subjects demonstrated pre- to post-training gains in self-efficacy, cognitive knowledge, psychomotor performance, and affective response regardless of training modality ($P < .01$ each). With the exception of fluid resuscitation in the psychomotor performance domain, no statistically significant differences were observed based on training modality in the overall group. Risk estimates on the least pretest performance subgroup favored simulation in 7 procedures. Affective response was greatest in live tissue training ($P < .01$) and varied by species and model. Subject matter experts noted significant value in live tissue in 7 procedures. Gap analysis noted shortcomings in all models and synergy between models.*
Conclusion. *Although simulation has made significant gains, no single modality can be identified definitively as superior. Wholesale abandonment of live tissue training is not warranted. We maintain that combined live tissue and simulation-based training add value and should be continued. Congressional mandates may accelerate simulation development and improve performance. (Surgery 2016;160:997-1007.)*

From the Department of Surgery,[a] College of Veterinary Medicine,[b] University of Missouri, Columbia, MO; Department of Surgery,[c] University of Alabama-Birmingham, Birmingham, AL; Department of Surgery,[d] University of South Florida, Tampa; and Information Visualization and Innovative Research (IVIR),[e] Sarasota, FL

The training of medical personnel in high-acuity, low-frequency, life-saving procedural interventions is challenging. An ideal model for training does not exist presently, and the debate between live tissue (LT) and models of inanimate simulation continues. Research designed to compare the educational effectiveness of LT versus simulation training is difficult to perform, and the time needed to assess long-term impacts of training interventions coupled with the fluid landscape of simulation development contributes to a paucity of information on which to base best practice.[1-4] Political pressures about the use of LT in medical training are substantial and have resulted in

Supported by the US Human Research Protections Office of the United States Army Medical Research and Materiel Command Grant W81XWH-10-JPC-MEDSIM-CCTC.

Presented at the Central Surgical Association Annual Meeting, Montreal, Quebec, Canada, March 10–12, 2016.

Accepted for publication April 13, 2016.

Reprint requests: Stephen L. Barnes, MD, Professor and Chief, Division of Acute Care Surgery, Department of Surgery, University of Missouri, One Hospital Drive, MC219, Columbia, MO 65212. E-mail: barnesste@health.missouri.edu.

0039-6060/$ - see front matter

http://dx.doi.org/10.1016/j.surg.2016.04.044

legislation limiting effectively the use of funds for LT training in the Department of Defense, despite current training practices relying heavily on this modality.[5,6]

Medical modeling and simulation technology, while advancing rapidly, still present technical challenges for faithful replication of human anatomy, physiology, and pathology. LT use does not lend itself to repetitive training or extensive throughput and carries with it complex regulatory requirements, extensive life-cycle and logistical support, and the physiologically confounding effects of general anesthesia. The Advanced Trauma Life Support course developed by the American College of Surgeons shifted from LT to simulation-based surgical skills training in the early 2000s; this shift occurred despite an American College of Surgeons position statement classifying animals as "an indispensable element of biomedical research, education, and teaching… and that, wherever feasible, alternatives to the use of live animals should be developed and employed," further highlighting the friction between abandonment of animal models used traditionally and the wholesale adoption of simulation that persists in medical education and training today.[7] The mandated transition from LT to simulation-based skills training in the military has exposed strong views from advocates of both LT and simulation, with the potential for training models to change without solid evidence to guide best practice.

Given the current state of medical simulation technology, the importance of effective training for optimal patient outcomes and the disagreement about the superiority of either LT or simulation-based training environments, the University of Missouri Combat Casualty Training Consortium (MU CCTC) was established to investigate the comparative effectiveness of LT and simulation-based training across a spectrum of 12 emergency trauma procedures.

METHODS

The MU CCTC represents a national coalition of subject matter experts (SMEs) encompassing the areas of battlefield/trauma surgery, surgical education, prehospital/battlefield medical care and training, educational practice and design, statistical analysis, and simulator design. The primary goal of the study was to identify best training practices and modalities to decrease preventable mortality on the battlefield and in civilian practice. This multiacademic and industry effort hypothesized that relevant differences in self-efficacy, cognitive performance (COG), psychomotor performance (PSY), and affective response (AFR) would be observed between subjects trained with LT versus simulation in 12 lifesaving emergency procedures (Table I). In 11 of the 12 procedures (P1−P11), the research design randomized subjects into LT or simulation training arms. For procedure P12 (nerve agent casualty), subjects were randomized into 3 training groups: LT, simulation, or a high-resolution video of the LT training exercise. Additionally, procedure P12 in the PSY assessments were separated into 3 subgroups representing the varying presentations of nerve agent exposure (Fig 1).

Standardization of training was achieved through scripted curricula. LT and inanimate simulation models were selected based on consensus input from the consortium (Tables I and II). To isolate the effect of each training modality, subject performance was assessed in a controlled setting without external stressors. Four animal models were utilized, and related procedures were grouped for sequential performance in both training and testing. Group 1 consisted of procedures P1–P5, group 2 included procedures P6–P10, and groups 3 and 4 each contained a single procedure, P11 and P12, respectively (Table I). This study design was based on logistical considerations and a desire to limit overall LT use. All training and testing was performed in a single day.

The subjects comprised a heterogeneous population of both military and civilian medic volunteers. Self-efficacy was measured through surveys administered pre- and post-training on a 10-point Likert scale. COG was measured by multiple choice assessments given pre- and post-training. PSY was scored by trained observers utilizing standardized checklists composed of readily identifiable and observable decomposed steps for each procedure. These assessments were procedure-specific, scored dichotomously, and designed to capture each subject's ability to perform a related action. PSY was analyzed in 2 ways: the total number of steps completed and the total number of critical steps completed. Step criticality was determined by consensus of the SMEs.

To account for potential inter-rater variation in the scoring of PSY assessment checklists, inter-rater concordance was used as a measure of consistent judgment between raters within each procedural grouping. Observational concordance was aided by strict definition of successful completion of each decomposed, readily identifiable, and observable item in the PSY performance checklist. Concordance between individual raters was achieved and documented via repeated scripted performances with planned omissions for rater training, both

Table I. Live tissue, simulation, video training groupings, subjects, simulators, animal models, and electrodermal activity affective results

Training group	ID	Procedure	N	Simulator		Animal model	Video	Affective response	
				Testing	Training			Finding	P value
		All	742	All	All	All	NHP video	LT>SIM>VID	<.001
Group 1	P1	Adult intubation	147	SimuLab TraumaMan	Trauma FX AirwayPlus	Caprine	NA	LT = SIM	.892
	P2	Surgical airway		Laerdal Airway Management Manikin	Lifecast; intubation amputation				
	P3	Chest seal							
	P4	Needle thoracostomy							
	P5	Tube thoracostomy							
Group 2	P6	Intraosseous device	194	MATT Trauma FX.	Tom000C Custom GSW	Porcine	NA	LT > SIM	.004
	P7	Casualty resuscitation		KGS-TFX-HEMO-1	T.O.M Man				
	P8	CAT tourniquet		KGS-TFX-APL-R					
	P9	Hemostatic dressing							
	P10	Amputation management							
Group 3	P11	Pediatric intubation	207	CAE Healthcare PediaSim2	TruCorp AirSim	Ferrett	NA	LT = SIM	.418
Group 4	P12	Nerve agent casualty	194	CAE Healthcare METIMan prehospital	Laerdal SimMan 3G	NHP	NHP video	LT>SIM>VID	<.001

Simulab, Seattle Washington; ITTS – Innovative Tactical Training Solutions – Crestwood, Kentucky; CAE Healthcare – Sarasota, Florida; KGS – Kforce, Inc, Fairfax, Virginia; Laerdal – Stavanger, Norway; Trucorp – Belfast, Ireland.
CAT, Combat application tourniquet; *ID*, procedural identification; *LT*, live tissue animal training; *NA*, not applicable; *NHP*, non-human primate; *P*, procedure.

prior to the start of the study and mid-study to assure sustainment.

AFR to the training modality was quantified by measuring the electrodermal activity (EDA) of each subject before and during training using the Affectiva Q-Sensor (Affectiva, Waltham, MA). This device was worn on the anterior wrist of the non-dominant hand and allowed for unobstructed free range of motion. EDA was sampled at 8 Hz via 2 12-mm electrodes. Matlab (MathWorks, Natick, MA) was used to analyze time-synched raw data. Baseline and training periods were extracted from each EDA time series. For each subject, the change in EDA response from baseline to training was quantified via the fractional change (FC) defined as FC = (training point estimate − baseline point estimate [BPE])/BPE. The training point estimate was taken as the maximum EDA response recorded in the training period after application of a moving window average to filter pressure and motion artifact. The BPE was taken as the 10% trimmed mean of the baseline period data. Fractional change was found to be skewed positively, and data were normalized using a log transformation prior to statistical analysis.[8,9]

Gap analysis of each training modality, LT and simulation, by procedure was performed by industry SME members of the MU CCTC. These members were not engaged directly in the research conduct and performed the analysis

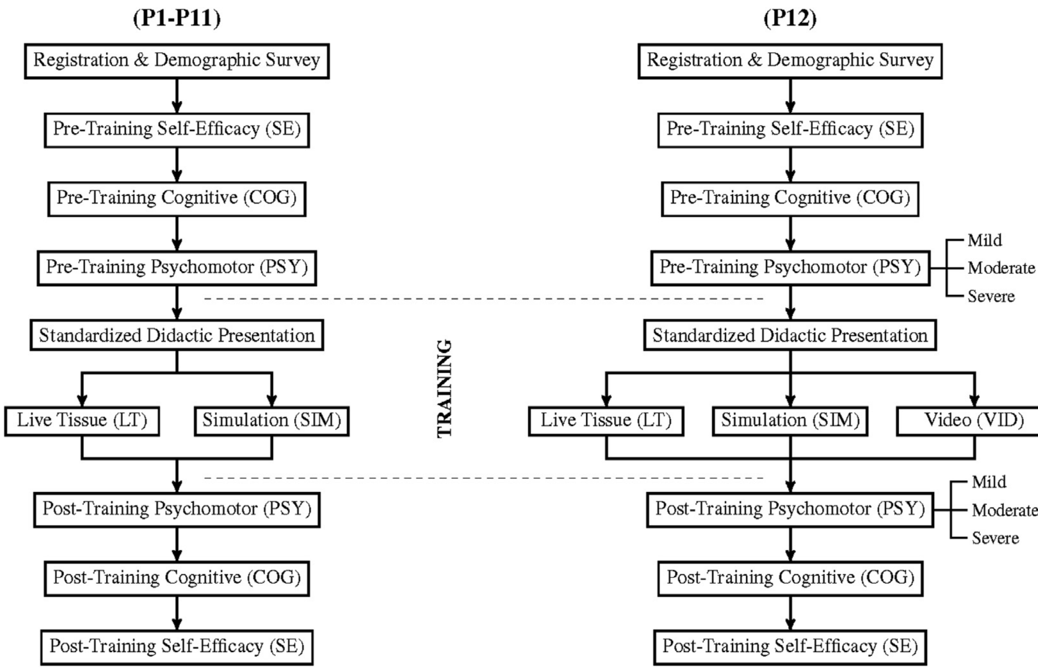

Fig. Flowchart of the research structure.

through direct observation and survey of research teams. For each procedure, PSY checklist steps were analyzed for their ability to be performed consistently without limitations during the pre- and post-testing and training exercises. For each modality, steps were labeled as either equivalent, unable to be accomplished, or accomplished uniquely by either LT or simulation. Based on the proportion of steps accomplished uniquely by modality, gap analysis results were identified as either representing synergy between modalities, favorable to LT, or favorable to simulation.

Data analysis focused on pre- to post-training differences of assessment instrument scores in the self-efficacy, COG, and PSY domains and on the EDA fractional change in the affective domain. Subjects were evaluated as a whole and then divided into high and low performing groups based on pre-training PSY scores. Customary descriptive statistics were calculated. For inferential statistics, general statistical assumptions were appraised for all data, including normality in the scales, linearity in the scores, homogeneity of variance, sphericity, homogeneity of regression, and absence of outliers in the score distributions. Scales were measured at a nominal level of measurement or above, and when summed, interval level was assumed. Outliers were explored with both logistic and ordinary least squares regression by regressing expected normal values onto observed values with a criterion of ±1 standard deviation. Additionally, to identify extreme values not judged to be outliers, z-scores were calculated using criterion ±3 z-score units. These data lead to risk analyses focusing primarily on odds ratios. In addition to these correlational approaches, inferential statistics included *t* tests, analysis of variance, and analysis of covariance (G-Power version 3.1, Kiel, Germany; SPSS version 22, Armonk, New York; SAS version 9.4, Cary, North Carolina).

To complement the research design, highly experienced senior special operations medics underwent structured interviews to evaluate their perceptions of LT and simulation as training tools. Overall responses were categorized as supporting either LT or simulation when responses were >80% in favor, mixed when responses were between 60% and 79%, and equivalent when responses were in the 40–59% range in regards to LT or simulation. Responses to directed questions exploring limitations and benefits of each modality were categorized and reported as a percentage.

All human subject use was approved by both local Institutional Review Boards and the Human Research Protections Office of the United States Army Medical Research and Materiel Command (USAMRMC). Prior to participation and data collection, written informed consent was obtained from each subject. All animal use was approved by both local Animal Care and Use Committees and the Animal Care and Use Review Office of the USAMRMC.

Table II. Live tissue versus simulation medic training: Results and training recommendations

Training group	Procedure ID	Procedure name	Self–efficacy (Analysis of variance)	Cognitive (Analysis of variance)	Psychomotor: All subjects, Analysis of variance: All steps	Psychomotor: All subjects, Analysis of variance: Critical completion Steps	Psychomotor: High performers, Analysis of variance risk estimates: All steps	Psychomotor: High performers, Analysis of variance risk estimates: Critical completion steps	Psychomotor: Low performers, Analysis of variance risk estimates: All Steps	Psychomotor: Low performers, Analysis of variance risk estimates: Critical Completion steps	Affective response	Structured interview	Gap analysis	Training recommendation
1	P1	Adult intubation	ND	ND	ND	ND	ND ND	ND ND	ND SIM	ND SIM	ND	Mixed	Synergistic	Limit or reduce LT
	P2	Surgical airway	ND	ND	ND	ND	ND ND	ND ND	ND SIM	ND ND	ND	LT	LT	Combined SIM and LT training
	P3	Chest seal	ND	ND	ND	ND	ND ND	ND ND	ND ND	ND ND	ND	Equivalent	Synergistic	Limit or reduce LT
	P4	Needle thoracostomy	ND	ND	ND	ND	ND ND	ND ND	ND SIM	ND SIM	ND	LT	Synergistic	Combined SIM and LT training
	P5	Tube thoracostomy	ND	ND	ND	ND	ND ND	ND ND	ND SIM	ND SIM	ND	LT	Synergistic	Combined SIM and LT training
2	P6	Intra-osseous device	ND	ND	ND	ND	ND ND	ND ND	ND ND	ND ND	LT	Mixed	Synergistic	Combined SIM and LT training
	P7	Casualty resuscitation	ND	ND	SIM	ND	ND ND	ND ND	ND SIM	ND ND	LT	LT	LT	Combined SIM and LT training
	P8	CAT tourniquet	ND	ND	ND	ND	ND ND	ND ND	ND SIM	ND ND	LT	Equivalent	Synergistic	Combined SIM and LT training
	P9	Hemostatic dressing	ND	ND	ND	ND	ND ND	ND ND	ND SIM	ND ND	LT	LT	Sim	Combined SIM and LT training
	P10	Amputation management	ND	ND	ND	ND	ND ND	ND ND	ND ND	ND ND	LT	LT	LT	Continue LT training
3	P11	Pediatric intubation	ND	ND	ND	ND	ND ND	ND ND	ND ND	ND ND	ND	Mixed	Synergistic	Limit or reduce LT
4	P12	Nerve agent casualty	ND	ND					ND	ND	LT	LT		Combined SIM and LT training
		Mild			SIM	SIM							SIM	
		Moderate			SIM	SIM							SIM	
		Severe			LT	LT							Synergistic	

ID, Procedural identification; *ND*, no statistically significant differences; *P*, procedure.

Table III. Inter-rater concordance for psychomotor steps evaluation

Training group	Procedure IDs	Inter-rater concordance	
		Pre-study	Mid-study
Group 1	P1−P5	89.3 %	97.3 %
Group 2	P6−P10	89.2 %	91.6 %
Group 3	P11	95.3 %	92.6 %
Group 4	P12	96.7 %	94.6 %

ID, Procedural identification; *P*, procedure.

RESULTS

In the study, 742 subjects had complete data sets for analysis. Age ranged from 18–64 years (32 years, mean age), with 459 (62%) males and 412 (55.5%) military participants; 384 (52%) were randomized to LT, 358 (48%) were randomized to simulation, and 64 (8.6%) were randomized to video for the nerve agent casualty P12 group 4 only. Of the 742 subjects, 238 (32.1%) claimed human experience in at least 1 of the procedures, and 322 (43.4%) claimed no experience in any of the procedures in either simulated or clinical environments. For the 12 procedures tested, all subjects demonstrated assessment gains from pre- to post-training in self-efficacy, COG, and PSY performance, regardless of training modality ($P < .01$ each). Similarly, subjects demonstrated increases in EDA fractional change ($P < .01$) from the baseline to training conditions, demonstrating the study format captured changes in performance effectively in all assessment domains, and the training design was effective in achieving those gains. Inter-rater concordance for evaluation of PSY performance ranged from 89.2−7.3 % throughout the study (Table III).

When all subjects were analyzed, no statistically significant differences in self-efficacy or COG were noted based on LT versus simulation training in any of the 12 procedures. Although there were more gains in those subjects with less prior knowledge or experience as demonstrated by pre-training testing and survey, randomization to LT or simulation-based training did not influence changes in self-efficacy and COG performance in this 1-day testing and training model. No differences were seen with PSY performance in 11 of the 12 procedures based on training modality. Only procedure P7, casualty resuscitation, favored inanimate simulation training for PSY (Table II).

Cognizant that the level of training and experience may play a role in the effectiveness of a given training modality, subjects were classified further into high and low performance groups based solely on pretraining PSY performance. For each procedure, analysis of variance, risk estimates, and odds ratios were calculated for all steps and the subset of critical steps. High performers demonstrated no differences in PSY based on training modality by either analysis of variance or risk estimate analysis. Low performers, while demonstrating no differences by analysis of variance, demonstrated differences by risk estimates and odds ratios based on training modality favoring simulation for 7 of the procedures (P1, 2, 4, 5, 7–9) when evaluating all steps and for 3 of the procedures (P1, 4, 5) when evaluating only the critical steps (Table II).

AFR to each LT and simulation training exercise was evaluated for each of the 4 procedural groupings (Table I). LT training across all modalities demonstrated a greater AFR as measured by EDA FC than that seen with simulation ($P < .001$); however, further evaluation of each LT model demonstrated differences between AFR. The porcine model utilized in procedures P6−P10, and the non-human primate model utilized in procedure P12, demonstrated greater AFR ($P = .004$ and $P < .001$, respectively) when compared with simulation training. The caprine and ferret models, utilized in procedures P1−P5 and P11, respectively, did not demonstrate differences between LT and simulation with respect to AFR. The additional video training modality utilized in procedure P12 only allowed for an additional comparison among VID, simulation, and LT training, with both LT and simulation having greater AFR than VID. Due to the methodology of collection of EDA data and the training curricula utilized, analysis of AFR was limited to training group and specific training modalities and was not specific to an individual procedure (Tables I and II).

Gap analysis of each procedure in the PSY checklist demonstrated significant shortcomings in both the LT and simulation models used in this study for faithful replication of the steps necessary for successful completion. In 11 of the 12 procedures evaluated (P1−P11), significant synergy was demonstrated in 7 (P1, 3–6, 8, and 11), LT was favored in 3 (P2, 7, and 10), and simulation was favored in 1 (P1). In the nerve agent casualty model with its 3 distinct presentations (P12), 2 of the presentations favored simulation, while 1 favored a synergistic approach to training. No single model was able to accomplish 100% of the steps in the PSY checklist for any of the procedures studied, with failure to accomplish rates varying from 8–50% (Tables II and IV).

Twenty-five senior US Army Special Operations medics underwent voluntary structured interviews.

Table IV. Gap analysis by percent decomposed steps accomplished by procedure

Training group	Procedure No.	Procedure Name	Steps, *N*	% (*N*) steps equally accomplished LT or SIM	% (*N*) steps unable to be accomplished by modality: LT	% (*N*) steps unable to be accomplished by modality: SIM	% (*N*) steps uniquely accomplished by modality: LT	% (*N*) steps uniquely accomplished by modality: SIM	Results
Group 1	P1	Adult intubation	28	68% (19)	14% (4)	11% (3)	7% (2)	11% (3)	Synergistic
	P2	Surgical airway	20	70% (14)	10% (2)	30% (6)	20% (4)	0.00% (0)	LT
	P3	Chest seal	7	43% (3)	29% (2)	43% (3)	29% (2)	15% (1)	Synergistic
	P4	Needle thoracostomy	10	40% (4)	30% (3)	30% (3)	30% (3)	30% (3)	Synergistic
	P5	Tube thoracostomy	21	48% (10)	33% (7)	24% (5)	14% (3)	24% (5)	Synergistic
Group 2	P6	Intra-osseous device	13	77% (10)	15% (2)	8% (1)	8% (1)	15% (2)	Synergistic
	P7	Casualty resuscitation	8	50% (4)	25% (2)	50% (4)	25% (2)	0.00% (0)	LT
	P8	CAT tourniquet	7	71% (5)	14% (1)	14% (1)	14% (1)	14% (1)	Synergistic
	P9	Hemostatic dressing	7	86% (6)	14% (1)	0.00% (0)	0.00% (0)	14% (1)	SIM
	P10	Amputation management	9	67% (6)	22% (2)	33% (3)	11% (1)	0.00% (0)	LT
Group 3	P11	Pediatric intubation	28	54% (15)	39% (11)	14% (4)	7% (2)	21% (6)	Synergistic
Group 4	P12	Nerve agent casualty, mild	13	0.00% (0)	100% (13)	23% (3)	0.00% (0)	62% (8)	SIM
		Nerve agent casualty, moderate	13	0.00% (0)	100% (13)	23% (3)	0.00% (0)	69% (9)	SIM
		Nerve agent casualty, severe	13	54% (7)	38% (5)	54% (7)	38% (5)	23% (3)	Synergistic

CAT, Combat application tourniquet; *P,* procedure; synergistic, both LT and SIM demonstrate uniquely accomplished steps.

All participants demonstrated extensive experience in all training environments under investigation (LT and simulation), with 96% reporting experience in these modalities as both students and instructors and 100% having combat experience. All but 1 (96%) reported translation of skills P1–P11 from training to the care of combat-related injuries.

The majority of respondents thought that the LT training was superior to SIM for 7 of the 12 procedures surveyed (P2, 4, 5, 7, 9, 10, and 12). Chest seal placement and tourniquet application (P3 and P8) garnered equal favor for both modalities, suggesting equivalence. Adult and pediatric intubation as well as insertion of the intra-osseous device (P1, 6, and 11) produced a mixed response (Table II). All participants thought LT training should be used in combat medic training, while 96% thought simulation also offered substantial benefit. Limitations and benefits were noted for both LT and simulation modalities by respondents (Table V). Simulators were noted commonly to be of appropriate size and weight, allowed for repeatability, and improved familiarization with equipment and steps producing muscle memory. In contrast, simulators were thought to provide an inaccurate and linear response to treatment, produced no sense of urgency, lacked the appropriate tactile feedback, and provided no visceral response to the trainee. Live tissue was noted most commonly to build confidence, instilling both a sense of urgency and a visceral response in the trainee to a model that can expire with more realistic responses to treatment and better tactile response. Reported limitations of LT were the non-human anatomy, logistic difficulty, and the need for anesthesia and veterinary medicine, thus making decision-making processes inconsistent with human

Table V. Structured interview live tissue versus simulation: Limitations and benefits

$N = 25$	Limitations	% response	Benefits	% response
Simulation	Inaccurate/linear response to treatment	84	Appropriate size and weight: easy logistics	96
	No sense of urgency: not taken seriously	84	Repeatability	92
	Lacks appropriate feel/tactile response	80	Familiarization: muscle memory	48
	No visceral response: no stress inoculation	70	Public perception	20
Live tissue	Inaccurate anatomy	84	Confidence: feel/tactile response	92
	Not appropriate size weight: logistics difficult	52	Accurate/nonlinear response to treatment	84
	Anesthetized/veterinary medicine	50	Instills sense of urgency: focuses training	84
	Public perception	28	Visceral response: stress inoculation	72

medicine. Public perception also was noted by respondents as a benefit of simulation and a limitation of LT.

DISCUSSION

LT always has been used in medical education. Although modern medical simulators came into use in the 1960s, the current climate of medical simulation has gained notable ground with advances in technology, materials, and cost-reduction in the past 25 years.[10] Even with these advances, LT remains an integral part of medical education, and wholesale adoption of the modern simulator for all medical training has not occurred. Few studies comparing directly LT to simulation exist.[11-16] "Political sensitivities" of all aspects (ethical, social, financial, etc) have played a major role in the move to mandate decreased use of LT for trauma training, especially for military medical personnel.[17-19]

In our head-to-head, single-day comparison of LT and simulation in a controlled environment without external stressors, we found no differences in the domains of COG or self-efficacy. Interestingly, those subjects with the greatest opportunities for PSY gain showed no benefit from LT; however, 7 procedures demonstrated some benefit from simulation training, suggesting simulation may be superior to LT for the introduction of a new procedure to the novice learner. Although these results are compelling, additional results from our comprehensive analysis (including the affective domain, gap analysis, and SME structured interviews) do not allow us to support wholesale change to paradigms of simulation-only based training.

AFR measured by EDA represents the emotional response to training and learner engagement. This study represents the first novel application of EDA measurement to medical education for the purpose of quantifying AFR to training. Both the porcine and non-human primate models were dynamic in nature, both in their interaction with the learner and response to treatment, while the ferret and caprine models were static in their presentations to the learner, similar to simulation. These model differences may account for the observation that some but not all LT models induced greater AFR than simulation and may suggest a development strategy for future generation simulation to have a greater impact on learner engagement.

Gap analysis demonstrated substantial and relevant shortcomings in both LT and simulation with no model achieving all educational goals and human-like performance for any of the 12 tested procedures. Many training programs utilize a combined approach of both LT and simulation for the introduction of new medical skills. We suspect that this combined training methodology allows local expertise to incorporate the most effective parts of LT and simulation, so that current technologic gaps are minimized, and training value is maximized.[20]

SME structured interviews demonstrated significant bias for LT while supporting the use of simulation in a combined training model as described above. It is unclear whether this was secondary to the rigors of training for medical care in the austere environment, perceived historic training success, or visceral response to training providing psychologic preparation for providing care under duress. It is clear that current training with combined LT and simulation is effective on the modern battlefield in terms of improved injury survival.[21]

Our analysis has limitations. The subject group was heterogeneous and demonstrated variance in age, prior training, and experience. Gains in those with prior experience or understanding were minimal and thus made significant comparisons of gains difficult to achieve. The pre- and post-training testing was performed with inanimate simulation, which may have allowed for additional learning and may have favored simulation outcomes. Each test occurred on a single day, so comments on retention based on LT or simulation cannot be made. The selection of simulators was based on commonly utilized training paradigms and did not evaluate all simulation technology directly nor technology which is/was in development in this rapidly progressing field. While EDA is an established method for quantifying emotional responses to defined stimuli, the use of EDA estimates of baseline and training points to enable the quantification of affective response to training environment as a fractional change represents a novel technique of EDA analysis.

Traditional medical education of "see one, do one, teach one" has changed with advances in technology and changes in practice and patient expectations.[22] Training environments have grown through the use of LT and simulation prior to direct patient interaction.[23,24] The passion and "politics" around the use of LT is substantial, and none more so than that seen in the training of combat casualty care.[17-19] These types of political issues, while important and relevant to the general population, should not drive best practice because the consequences are too high, most notably in high acuity, low occurrence events. Without the ideal model, educators must choose from those available, and mitigate the negative and accentuate the positive aspects of each. As responsible educators, we must refine, decrease, replace, and whenever possible respect the use of LT protocols for training.[25] As efforts are made to decrease the use of LT, there should be recognition that LT and simulation offer differing benefits to learners for some skills. The combined use of LT and SIM allows curricula to maximize the benefits of both training modalities and lead to the best opportunities for success in the future. Congressionally mandated development programs and ever advancing technology hold promise to continue to decrease and may eliminate ultimately LT in medical and surgical education.

Although simulation has made widespread substantial gains in education, no single training modality can yet be identified definitively as superior for any of the 12 emergency trauma skills we evaluated. Wholesale abandonment of LT is not warranted and should be addressed with caution. Combined LT and simulation-based training adds value to training outcomes. Congressional funding may accelerate simulation development, improve performance, and move medical and surgical education closer to LT-free training environments, but we are not there yet.

Views expressed are those of the authors alone and do not represent the views of the US Government or the Department of Defense. University of Missouri Combat Casualty Training Consortium Investigators: Col. Jeff Bailey, MD, FACS, Walter Reed Army Medical Center, Bethesda, MD; Col. (retired) Warren Dorlac, MD, FACS, USAF CSTARS, Cincinnati, OH; Rob Shotto, Penn State University Simulation Center, Hershey, PA; Jack Norfleet, USA RDECOM/ARL-STTC, TX; Tim Coakley, MD, USN Deputy Force Surgeon, VA; Mark Bowyer, MD, FACS, USUHS National Capital Simulation Center, Bethesda, MD; Bousseau Murray, MD, Penn State University Department of Anesthesia, Hershey, PA; Mark Shapiro, MD, FACS, Duke University Department of Surgery, Durham, NC; Roberto Manson, MD, Duke University Department of Surgery, Durham, NC; Al Moloff, MD, Information Visualization and Innovative Research Inc, Sarasota, FL; Deborah Burgess, MD, The Salus Group, San Antonio, TX; Robert Hester, PhD, University of Mississippi, Jackson, MS; William Lewandowski, 3D Systems Corporation, Rock Hill, SC; Waymon Armstrong, Engineering and Computer Simulations, Inc, Orlando, FL; Jack McNeff, Simulaids Corp, Saugerties, NY; Jan Cannon-Bowers, PhD, Joanne Hardeman, MPA, Jenny Guido, MD, Cole Giering, BS, Robert Rohrlack, BS, Jessica Acosta, BS, Raj Patel, BS, Zachary Green, BS, University of South Florida Center for Advanced Medical Simulation and Learning, Tampa, FL; Ronald Roan, MD, Adam Robinett, MD, Scott Snyder, PhD, Bharat Soni, PhD, Dale Davis, BSN, Lina Rodriquez, BS, BSN, Phillip Shum, MS, University of Alabama-Birmingham, Birmingham Alabama; Steve Osterlind, PhD, Chris Cooper, MD, Rindi Uhlich, MD, Christina Stephan, MD, PhD, John Tucker, MS, University of Missouri, Columbia MO; John Anton, MS, Ray Shuford, BS, Catherine Strayhorn, BS, Emily Anton, BS, Nadine Baez, BS, Erin Honold, BS, Information Visualization and Innovative Research (IVIR), Sarasota, FL.

REFERENCES

1. Lineberry M, Walwanis M, Reni J. Comparative research on training simulators in emergency medicine: a methodological review. Simul Healthc 2013;8:253-61.
2. Van Nortwirck SS, Lendvay TS, Jensen AR, Wright AS, Horvath KD, Kim S. Methodologies for establishing validity in surgical simulation studies. Surgery 2010;147:622-30.
3. Cheng A, Auerbach M, Hunt E, et al. Designing and conducting simulation-based research. Pediatrics 2014;133:1091-101.
4. Cook D. How much evidence does it take? A cumulative meta-analysis of outcomes of simulation-based education. Medical Education 2014;48:750-60.

5. HR 3172. 113th Congress. S.1550. September 25, 2013. Battlefield Excellence through Superior Training Practices Act (BEST).
6. National Defense Authorization Act 2013, Sect 736. Report on strategy to transition to use of human-based methods for certain medical training.
7. Statement on the Use of Animals in Research, Education, and Teaching Bulletin of the American College of Surgeons. Bull Am Coll Surg 2002;87:16.
8. Poh MZ, Swenson NC, Picard RW. A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. IEEE Trans Biomed Eng 2010;57:1243-52.
9. Phitayakorn R, Minehart RD, Pian-Smith MCM, Hemingway MW, Petrusa ER. Practicality of using galvanic skin response to measure intraoperative physiologic autonomic activation in operating room team members. Surgery 2015;158:1415-20.
10. Cooper JB, Taqueti VR. A brief history of the development of mannequin simulators for clinical education and training. Qual Saf Health Care 2004;13(Suppl 1):11-8.
11. da Luz LT, Nascimento B, Tien H, Kim MJ, Nathens AB, Vlachos S, et al. Current use of live tissue training in trauma: a descriptive systematic review. Can J Surg 2015; 58(Suppl 3):S125-33.
12. Hall AB. Randomized objective comparison of live tissue training versus simulators for emergency procedures. American Surgeon 2011;77:561-5.
13. Savage LE, Tenn C, Vartanian O, Blackler K, Sullivan-Kwantes W, Garrett M, et al. A comparison of live tissue training and high-fidelity patient simulator: a pilot study in battlefield trauma training. J Trauma Acute Care Surg 2015;79:S157-63.
14. Hall AB, Riojas R, Sharon D. Comparison of self-efficacy and its improvement after artificial simulator or live animal model emergency procedure training. Mil Med 2014;179:320.
15. Sutherland LM, Middleton PF, Anthony A, Hamdorf J, Cregan P, Scott D, et al. Surgical simulation: a systemic review. Ann Surg 2006;243:291-300.
16. Iverson K, Riojas R, Sharon D, Hall AB. Objective comparison of animal training versus artificial simulation for initial cricothyroidotomy training. Am Surg 2015;81:515-8.
17. Gala SG, Goodman JR, Murphy MP, Balsam MJ. Use of animals by NATO countries in military medical training exercises: an international survey. Military Medicine 2012;177: 907-9.
18. Petition for enforcement. In re: training of Air Force Special Operations Forces at Joint Base Lewis McChord. Physicians Committee for Responsible Medicine. 2013. https://assets.documentcloud.org/documents/700761/pigs.pdf.
19. Martinic G. The use of animals in live-tissue trauma training and military medical research. Lab Anim 2011; 40:319-22.
20. Sohn VY, Runser LA, Puntel RA, Sebesta JA, Beekley AC, Theis JL, et al. Training physicians for combat casualty care on the modern battlefield. J Surg Edu 2007;64:199-203.
21. Butler FK, Blackbourne LH. Battlefield trauma care then and now: a decade of tactical combat casualty care. J Trauma Acute Care Surg 2012;73(6 Suppl 5):S395-402.
22. Ziv A, Wolpe PR, Small SD, Glick S. Simulation-based medical education: an ethical imperative. Acad Med 2003;78: 783-8.
23. Okuda Y, Bond W, Bonfante G, McLaughlin S, Spillane L, Wang E, et al. National growth in simulation training within emergency medicine residency programs, 2003–2008. Acad Emerg Med 2008;15:1113-6.
24. Qayumi K, Pachev G, Zheng B, Ziv A, Valentyna K, Badei S, et al. Status of simulation in health care education: an international survey. Adv Med Educ Pract 2014;5:457-67.
25. Animal Welfare Act, Amendment of Title XVII, Subtitle F, Sections 1751−1759, US Statutes at Large 99:50, 1985.

DISCUSSION

Dr Martin Zielinski (Rochester, MN): Dr Barnes, I would like to thank and congratulate you and your coauthors on an excellent presentation and a timely manuscript. Central Surgical's Program Committee should also be commended on the inclusion of how best to educate our surgical residents.

In the overly and heavily mandated world in which we live, the use of live tissue models has come into question. As the authors point out, the Department of Defense has been limited by political pressure in the use of live tissue models. In fact, their use may become banned altogether.

As a result, the DOD funded Dr Barnes and his colleagues to determine, in a scientific fashion, whether these changes in policy will result in unanticipated changes in practice. The methods employed by the authors are truly awesome and are an example of rigorous scientific methods and analysis. They concluded that any education is good, but some models are more ideally suited to certain parameters and that a live tissue model may be superior in certain circumstances while inanimate simulation may be better in others. I do have a few questions.

Due to the same political sensitivities, you limited your use of the live tissue modeling for this study. How did this affect your results? In other words, did the imposed limitation negatively affect the outcome of the very study designed to determine if the imposed limitation affects education?

Second, we all have residents on our service that are excused from clinical responsibilities to attend these educational activities at our respective simulation centers. Therefore, these experiences must contain the most amount of education in the least amount of time. If you can identify the optimal simulation training program, what would it look like based on your results?

Are there previously live models that may be educational without the negative political pressures? For instance, as Doctors Farnell and Nagorney will remember, I'm sure, I also vividly remember sewing hot dogs with 5-0 Prolene to simulate a pancreas anastomosis.

Are there other educational models which may be feasible using previously living tissue? Perhaps, for the intraosseous placement, a previously living femur, or something similar, could be used.

Instead of the moratorium on live tissues, perhaps a middle ground may be able to be discovered which utilizes the advantages of live tissue when necessary but otherwise uses inanimate simulation when live tissue holds no educational advantages.

I want to thank you for the privilege of reviewing this outstanding work.

Dr Stephen Barnes (Columbia, MO): Thank you Dr Zielinski for those kind words. The ideal structure for this evaluation would have been single animal, single student, single student, crossover analysis, and utilize pre and post testing in both the simulated and live tissue environments. That structure would make the study cost prohibitive and the political sensitivities with animal work would have made IRB and IACUC approvals at both the local and federal levels impossible to achieve in our opinion. Did it affect outcomes? I think you are always learning, even when in the testing environment. So was everybody learning how to interact with the simulator in the first simulated event? Yes. There's a little bit of "gaming" that goes on when you are working on a plastic simulator just like there is a little bit of "gaming" when you're in an animal model. Neither is a perfect representation of human beings. Did the simulator only group get better because they had an extra round of practice in the pre training testing? I cannot tell you, I suspect maybe.

As for resident education, it is all about curriculum. One must identify the objectives of any training and then create a curriculum to satisfy those stated objectives. So to tie into question 3, a hot dog may be all you need to achieve your objectives and can be used as your simulated model in your curriculum as opposed to sacrificing an animal to learn how to perform a pancreatic anastomosis.

I think it's really based on curriculum that can achieve your objectives. Most of us think about having a simulated event for our trainees based on what is available in the simulation lab first, then we think about why we're doing it second. We need to spend more time focused on objects and appropriate curriculum prior to choosing which environment best achieves those goals.

Dr Sally Carty (Pittsburgh, PA): I enjoyed your presentation. It was really interesting to think about.

In your opinion, what is it that makes live training better? Is it the motion? Is it the bleeding? Is it the distress? How would you change simulation to improve it?

Dr Stephen Barnes (Columbia, MO): Thank you Dr. Carty. The more invasive you get with your training, the more educators seem to like live tissue because the interaction and the interface between the learner and what they learning on declines when you're just interacting with multiple layers of plastic that lack the fidelity and interaction of live tissue.

The simulation industry can solve that issue tomorrow, and none of us could afford their simulators. So it becomes a balance between cost and fidelity. I personally believe the type of simulated environment being utilized affects the personality and the atmosphere of the learning environment. When there is a piece of plastic lying in the room, there are a bunch of high fives, there's some joking, there's talk about what we're going to do tonight. Where there is something alive lying in that same room, there is focus, and there is a completely different personality of engagement for that learning event. I believe that has benefit. I don't know that we captured that here, but I think it has benefit.

I personally believe that advances in haptics and virtual reality environments will have to be achieved where learners get the appropriate tactile feedback and can be put into any environment to make live-tissue obsolete. But the simulation industry get closer and closer to the goal every year.