Battlefield Trauma Training: A Pilot Study Comparing the Effects of Live Tissue vs. High Fidelity Patient Simulator on Stress, Cognitive Function and Performance

Oshin Vartanian¹,², Catherine Tenn¹, Wendy Sullivan-Kwantes¹, Kristen Blackler¹, Ingrid Smith¹,
Henry Peng¹, Jerzy Jarmasz¹, Norlene Caddy¹, Ann-Renee Blais¹, Capt Dylan Pannell³,⁴, Fethi Bouak¹, Doug Saunders¹ and Col Homer C. Tien³,⁴,⁵

¹ = Defence Research and Development Canada
² = University of Toronto Scarborough
³ = Canadian Forces Health Services
⁴ = University of Toronto
⁵ = Sunnybrook Health Sciences Centre

Address for correspondence:
Oshin Vartanian
Defence Research and Development Canada; Toronto Research Centre
1133 Sheppard Ave. West
Toronto ON Canada M3K 2C9
Email: oshin.vartanian@drdc-rddc.gc.ca

Author Note

This manuscript is dedicated to the memory of LCol Erin Savage. We would like to acknowledge the following personnel from Defence Research and Development Canada: Stephen Bjarnason, Michelle Garrett, Corry Byrne, Cory Vair, Danielle Poissant, Kendal Weatherby, Ira Hill, Capt Russell McAllister, Aidan Thompson and Blake Martin, as well as Matthew Jubelius from Medicine Hat College, Alberta, Canada. The authors also wish to thank The Technical Cooperative Panel HUM TP12 for their valuable comments and insights on the topic of medical modeling and simulation versus live tissue training.
Public Significance Statement

In the Canadian context, both simulator and live tissue modalities are used to train medics in the treatment of preventable causes of death on the battlefield. Our results demonstrated that there was a positive correlation between stress and cognitive function during training. However, the choice of modality was not associated with different levels of stress or cognitive function in trainees.
Abstract

Within the Canadian Armed Forces (CAF), the Tactical Medicine (TACMED) course is used to train Medical Technicians (i.e., medics) in battlefield trauma care. Although training is administered using both simulators (SIM) and live tissue (LT), little is known about their relative effects on stress and cognitive function in this context. To address this shortcoming, we conducted a pilot study and collected self-report (State-Trait Anxiety Inventory - STAI) and biological measures of stress (salivary cortisol and dehydroepiandrosterone - DHEA), as well as working memory (WM) and short-term memory (STM) data from medics ($N = 20$)—assigned randomly to training and skill assessment using either SIM or LT. Skill assessment resulted in the elevation of STAI scores and salivary cortisol and DHEA levels. WM and STM performance were better at the time of skill assessment, and WM performance exhibited a positive correlation with salivary cortisol level. Salivary cortisol and DHEA levels, STAI scores, and memory performance did not predict pass/fail rates on combat casualty care skills. Although the TACMED course was associated with elevated stress and improved memory performance, those effects were not affected by the training modality. We end by discussing lessons learned from our pilot study, and highlight outstanding questions that remain to be addressed in future studies on this topic.

*Keywords:* Battlefield trauma training, live tissue, simulation, stress.
Within the Canadian Armed Forces (CAF), the Tactical Medicine (TACMED) course is used to train Medical Technicians (medics) in the treatment of preventable causes of death on the battlefield, such as massive hemorrhage, tension pneumothorax and airway obstruction, in preparation for overseas deployment. During training, skill acquisition is assessed based on performance on several combat casualty care skills (Savage, Forestier, Withers, Tien & Pannell, 2011; Savage et al., 2015; Tien, Jung, Rizoli, Acharya & McDonald, 2008). Although training is administered using both simulators (SIM) and live tissue (LT), little is known about the relative effects of each training modality on skill acquisition. A recent review “did not identify a body of evidence robust enough to conclude whether LTT [live tissue training] is better than other simulation methods” (da Luz et al., 2015, p. S130). However, strong inferences were difficult to draw because the available studies suffered from various methodological shortcomings, including small sample sizes, lack of control groups in the majority of studies, and variability in simulators, participants, and outcome measures.

Recently, we took a step towards assessing the relative merits of SIM and LT by studying their effects on performance on five combat casualty care skills in the context of a TACMED course (Savage et al., 2015). Specifically, participants \( n = 20 \) were randomly assigned to train on either a SIM or LT model, following which they were further randomized prior to subsequent skill assessment during a simulated combat scenario on either of the two training models. This design (i.e., prospective, single blind, randomized controlled study) was implemented specifically to assess the degree of skill transfer from one training modality to another in both directions. Although modality had no effect on performance on any specific skill, pass rates were significantly lower if the training and testing modality were dissimilar (summed across all
combat casualty care skills) (Sullivan-Kwantes, Vartanian, Jarmasz, Tenn, & Nazarov, 2015). In addition, semi-structured interviews conducted after final assessments revealed a strong preference for LT over SIM. Specifically, all participants felt that SIM alone would not prepare them adequately to treat causalities in the battlefield due to a variety of reasons, including lack of realism and difficulties with the localization of anatomical landmarks. Critically, consistent with the performance data (i.e., pass/fail rates), most (18/20) participants stated that the transition from one modality to another was problematic, suggesting that forward transfer of learning is hampered when the training and target modalities are not the same (see also da Luz et al. 2015). It is possible that these performance impairments might have been driven by increased stress and/or decreased cognitive capacity as a result of the transition.

In addition to the effects of SIM and LT on pass/fail rates in combat casualty care skills (Savage et al., 2015), we were also interested in exploring possible differences between the two training modalities in terms of the effects they might exert on stress and cognitive function. This is important for four reasons. First, recall of learned material is enhanced to the extent that the retrieval context matches the encoding (i.e., learning) context (Smith & Vela, 2001). Given that treating wounds in the battlefield is likely associated with a high degree of stress, it might be advantageous to train the medics in a context that more closely resembles the higher stress levels encountered in the battlefield. Here it is important to note that LT is not necessarily experienced by trainees as more stressful than SIM. Indeed, our semi-structured interviews revealed that LT can in fact attenuate stress levels by diminishing the subjective sense of time pressure, in turn instilling confidence in trainees. For example, one trainee noted that "...it almost made me more confident in my skills...because...if I am stressed, I can take a step back
and know that I can calm myself down and get back into it...even though it's a sense of urgency that you need...it's better for you to wait...just knowing two seconds isn't going to make a difference...thirty seconds is a big deal...but it's not two seconds." Hence, based on observations from previous research, we predicted that there would be differences in stress levels between LT and SIM—assessed using biological measures of stress including salivary cortisol and dehydroepiandrosterone (DHEA) levels as well as self-reported measures (hypothesis 1). This approach is consistent with recent trends to assess variations in biological and psychological reactivity within combat casualty care scenarios (McGraw et al., 2013).

Second, we were interested in determining the relationship between stress and cognitive performance. Specifically, a review of the literature on the effects of acute stress on the clinical performance of individuals and teams demonstrated that elevated stress levels can impair performance on tasks that require divided attention, working memory (WM), retrieval of information from memory and decision making (LeBlanc, 2009). Given that performance in the TACMED course likely draws on those capacities, we hypothesized that there would be a negative correlation between stress and memory performance (hypothesis 2)—the latter assessed using tests of WM and short-term memory (STM).

Third, although we have shown previously that the transition from one modality to another was associated with lower pass rates (summed across all skills), we have not examined whether these transitions also cause increases in stress and/or decreases in cognitive function—in turn providing possible mechanisms that could be used to explain the lower pass rates associated with modality transition. Thus, we hypothesized that the transition from one modality to another would be associated with increases in stress and decreases in cognitive
function—assessed using self-report and salivary cortisol and DHEA levels (for stress) and tests of WM and STM (for cognitive function) (*hypotheses 3-4* respectively).

Finally, using logistic regression, we tested whether pass/fail rates on five combat casualty care skills (surgical airway insertion, needle decompression, tourniquet application, wound packing and intraosseous line insertion) would be affected by cognitive capacity (STM, WM) and stress (self-report and biomarker levels). Specifically, we predicted that cognitive capacity would contribute positively to pass/fail rates (*hypothesis 5*), but that stress would contribute negatively to pass/fail rates (*hypothesis 6*).

**Method**

This study protocol was approved by Defence Research and Development Canada’s Human Research Ethics Committee (DRDC HREC).

**General overview: TACMED course**

The data were collected in the context of a TACMED course that unfolded over seven consecutive days (for design see Figure 1). The first day consisted of a classroom lecture during which the theoretical and applied aspects of the combat casualty skill sets were reviewed. Following the classroom lecture, participants were assigned randomly to train on LT (*n* = 10) or SIM (*n* = 10). On the second day, participants underwent a familiarization session in the operating room (OR), including individualized training on each skill taught by experienced TACMED instructors. Baseline data were collected on the second day (i.e., prior to any skill assessment). Baseline data represent initial measures. Days 3-4 involved skill assessment on the same modality on which they were trained in the OR. Performance was evaluated by medical officers (MO) using a standardized scoring method routinely used in the TACMED course. Next,
on days 5-6 participants were randomized again and evaluated on either SIM or LT in a simulated combat scenario. Thus, on days 5-6 half of the participants were evaluated on the same modality as their training modality in the OR, whereas the other half crossed over and were evaluated on the other modality. MOs who conducted the evaluations were blind to each participant’s training modality in the OR. The seventh day consisted of debriefing and interviews. Whereas Savage et al. (2015) focused on the effects of training modality on pass/fail rates on days 3-4 (in the OR) and on days 5-6 (simulated combat scenario), here we shifted our focus to assessments of cognitive function and stress during training (in the OR) and skill assessment during the simulated combat scenario (vis-à-vis baseline).

Participants

Twenty CAF medics with no prior TACMED course experience volunteered to participate in this study. There was no significant difference between the two training groups (SIM vs. LT) in gender (14 male, 6 female), age ($M = 27.05, SD = 6.04$), status (19 regular force, 1 reservist), rank (non-commissioned members = 20), years of service ($M = 6.18, SD = 6.77$), education (college diploma = 20), or deployment history (18 deployed, 2 non-deployed) ($p > .05$).

Materials

Self-reported stress. The State-Trait Anxiety Inventory (STAI; Spielberger, 2011) is a measure of anxiety, used here as a proxy measure of situational stress. We used STAI Form Y. A sample item includes I am tense, to which one responds using a scale ranging from 1 (not at all) to 4 (very much so).

Cognitive function. We focused on WM and STM for assessing cognitive function. WM was assessed using the n-back task, which requires that participants decide, on a trial-by-trial
basis, whether a stimulus presented in the current trial matches a target stimulus presented a specific number of trials earlier in the sequence. The letter $n$ denotes the specific number of trials that separate the current trial from the target trial. In this study $n$ ranged from 1 to 2. For 1-back participants would press the spacebar only if the letter presently on the screen matched the letter in the previous trial, whereas for 2-back participants would press the spacebar only if the letter presently on the screen matched the letter two trials earlier. Performance was measured by the number of misses. This task necessitates the maintenance and updating of dynamic rehearsal sets in WM during performance (Kane, Conway, Miura & Colflesh 2007).

Delayed matching-to-sample (dMTS) is a classic measure of STM (Miller, Erickson & Desimone 1996), and includes three phases: encoding, maintenance, and retrieval. During encoding participants memorized the stimulus (an $8 \times 8$ grid of a random distribution of green and red blocks); during maintenance participants maintained the stimulus in short-term memory; during retrieval they pressed the button corresponding to one of the two adjacent stimuli that matched the stimulus presented during encoding. The Cognitive Test Software (Grushcow, 2008) was used to administer the tasks.

**Salivary samples.** Salivary levels of cortisol (and DHEA) are well-established measures of acute stress (Dickerson & Kemeny, 2004). Saliva was collected using cortisol salivettes (SARSTEDT Inc., Montreal, QC, Canada). The samples were analyzed for cortisol and DHEA with enzyme immunoassay kits (Salimetrics, LLC, State College, PA, USA). DHEA and cortisol assays possessed intra- and inter-assay coefficients of variation < 10% based on manufacturer’s kit performance characteristics.
**Live Tissue.** The live tissue used in this study were male castrated York-Landrace cross pigs (20-25 kg). The research was conducted in accordance with the standards of the Canadian Council on Animal Care, and approved by DRDC’s Animal Care Committee.

**Simulator.** We used the CAESAR™ Trauma Patient Simulator (CAE Healthcare, Montreal, QC, Canada). It is a mannequin patient simulator, operated and monitored remotely via a wireless control panel, powered by rechargeable internal batteries, and modeled in the likeness of a 6-foot-4 Caucasian adult male. It is designed with the aim of reproducing the effects of combat trauma, and is able to simulate a number of traumatic injuries to allow trainees to practice a number of TACMED interventions (e.g., tourniquet, cricothyrotomy, and needle chest decompression) (CAE Healthcare, 2014). Simulators were operated by qualified CAE Healthcare personnel and assisted by a trained moulage team.

**Procedures**

Data were collected over the course of TACMED training. On the second day of training we collected baseline saliva in the morning (AM) and in the afternoon (PM) because of naturally occurring circadian variation in salivary cortisol and DHEA levels (Weitzman, 1971). In the AM we also collected STAI, n-back and dMTS data. We administered the STAI, n-back and dMTS and collected saliva samples before and after skill assessment in the OR (days 3-4) and before and after skill assessment in the simulated combat scenario (days 5-6). We focused on measurements obtained immediately after skill assessment (vis-à-vis baseline) because our aim was to capture elevations in stress levels as a result of exposure to the test conditions.

**Results**
Cortisol and DHEA. As a manipulation check, we compared cortisol and DHEA levels in the AM vs. PM at baseline. As predicted, cortisol levels were higher in the AM than the PM at baseline, $t(19) = 4.22, p < .001$, Cohen’s $d = 1.02$. For DHEA there was a similar trend for higher levels in the AM than the PM, but that difference was not statistically significant, $t(19) = 1.74, p < .10$, Cohen’s $d = .34$.

We examined whether salivary cortisol and DHEA levels varied as a function of the testing modality compared to baseline. Each participant’s post-OR and simulated combat scenario levels were compared separately to the closest baseline time point (i.e., AM or PM). A separate mixed-model analysis of variance (ANOVA) with time point (2 levels: baseline, post-OR/simulated combat scenario) as the repeated measure and condition (2 levels: LT, SIM) as the independent measure was conducted for each dependent variable and for each time point (i.e., post-OR or simulated combat scenario). Post-OR levels of cortisol were elevated compared to baseline, $F(1, 18) = 22.85, p < .001$, partial $\eta^2 = .56$. However, there was no main effect for modality, $F(1, 18) = .30, ns$, partial $\eta^2 = .02$, or a modality $\times$ time point interaction, $F(1, 18) = .07, ns$, partial $\eta^2 = .01$. Post-OR levels of DHEA were elevated compared to baseline, $F(1, 18) = 17.94, p < .001$, partial $\eta^2 = .50$. However, there was no main effect for modality, $F(1, 18) = .01, ns$, partial $\eta^2 = .01$, or a modality $\times$ time point interaction, $F(1, 18) = .06, ns$, partial $\eta^2 = .01$.

In regards to the simulated combat scenario, levels of cortisol were elevated compared to baseline, $F(1, 18) = 38.70, p < .001$, partial $\eta^2 = .68$. However, there was no main effect for modality, $F(1, 18) = .04, ns$, partial $\eta^2 = .01$, or a modality $\times$ time point interaction, $F(1, 18) = .02, ns$, partial $\eta^2 = .01$. In addition, levels of DHEA were elevated compared to baseline, $F(1, 18) = 8.30, p < .05$, partial $\eta^2 = .32$. However, there was no main effect for modality, $F(1, 18) = .04, ns$,
partial $\eta^2 = .01$, or a modality $\times$ time point interaction, $F(1, 18) = .33$, ns, partial $\eta^2 = .02$. Thus, for both cortisol and DHEA, there was no interaction between modality and variation in stress hormone levels.

**Self-reported stress.** Post-OR STAI scores were elevated compared to baseline, $F(1, 18) = 4.72, p < .05$, partial $\eta^2 = .21$. However, there was no main effect for modality, $F(1, 18) = .04$, ns, partial $\eta^2 = .01$, or a modality $\times$ time point interaction, $F(1, 18) = 2.91$, ns, partial $\eta^2 = .14$. In regards to the simulated combat scenario, STAI scores were not elevated compared to baseline, $F(1, 18) = .46$, ns, partial $\eta^2 = .03$. In addition, there was no main effect for modality, $F(1, 18) = .21$, ns, partial $\eta^2 = .01$, or a modality $\times$ time point interaction, $F(1, 18) = 3.12$, ns, partial $\eta^2 = .15$.

**Cognitive function.** Post-OR performance on dMTS was marginally better ($F[1, 18] = 4.26, p = .05$, partial $\eta^2 = .19$. However, there was no main effect for modality, $F(1, 18) = 1.72$, ns, partial $\eta^2 = .09$, or a modality $\times$ time point interaction, $F(1, 18) = 3.36$, ns, partial $\eta^2 = .16$. Post-OR there was no difference in 1-back performance compared to baseline, $F(1, 18) = 1.47$, ns, partial $\eta^2 = .08$, no main effect for modality, $F(1, 17) = 1.52$, ns, partial $\eta^2 = .08$, or a modality $\times$ time point interaction, $F(1, 18) = 1.06$, ns, partial $\eta^2 = .06$. Post-OR 2-back performance was better compared to baseline, $F(1, 18) = 10.41, p < .01$, partial $\eta^2 = .86$. However, there was no main effect for modality, $F(1, 17) = .12$, ns, partial $\eta^2 = .01$, or a modality $\times$ time point interaction, $F(1, 18) = .05$, ns, partial $\eta^2 = .06$.

In regards to the simulated combat scenario, dMTS performance was better compared to baseline, $F(1, 18) = 9.07, p < .01$, partial $\eta^2 = .34$. In addition, there was a main effect for modality, such that accuracy was higher for LT compared to SIM, $F(1, 18) = 10.64, p < .01,$
partial $\eta^2 = .37$. However, there was no modality $\times$ time point interaction, $F(1, 18) = .01$, ns, partial $\eta^2 = .01$. Regarding performance on 1-back, there was no difference compared to baseline, $F(1, 18) = 1.56$, ns, partial $\eta^2 = .08$. In addition, there was no main effect for modality, $F(1, 18) = 1.54$, ns, partial $\eta^2 = .08$, or a modality $\times$ time point interaction, $F(1, 18) = .77$, ns, partial $\eta^2 = .04$. Finally, 2-back performance was better compared to baseline, $F(1, 18) = 10.14$, $p < .01$, partial $\eta^2 = .85$. However, there was no main effect for modality, $F(1, 18) = .09$, ns, partial $\eta^2 = .01$, or a modality $\times$ time point interaction, $F(1, 18) = .26$, ns, partial $\eta^2 = .01$. Thus, for both dMTS and n-back, modality did not interact with variation in cognitive function.

Relationship between stress and cognitive function. To examine the relationship between stress and cognitive function, we calculated the zero-order correlations between salivary cortisol and DHEA levels and performance on n-back and dMTS. Because measurements of salivary cortisol and DHEA levels and n-back and dMTS performance were obtained at multiple time points throughout the study (Figure 1), an analytic procedure was implemented in order not to violate the independence assumption when computing correlation coefficients (Bland, 2000). Specifically, for each participant, we computed average cortisol and DHEA scores across all time points (i.e., baseline, post-OR and simulated combat scenario). Similarly, for each participant we computed average dMTS (accuracy) and n-back scores (number of misses separately for 1-back and 2-back) across the same three time points. In turn, correlations were computed involving those averaged data points. As expected, there was a positive correlation between salivary cortisol and DHEA levels, $r(18) = .73$, $p < .001$. Also as expected, performance on dMTS (accuracy) was correlated negatively with n-back (number of misses) for both 1-back ($r[18] = -.52$, $p < .05$) and 2-back ($r[18] = -.61$, $p < .01$). In addition,
there was a negative correlation between cortisol level and 1-back misses ($r_{18} = -0.45, p < .05$), but no correlation between cortisol level and 2-back misses ($r_{18} = -0.33, p = .16$) or dMTS accuracy ($r_{18} = 0.28, p = .22$). Thus, greater stress levels as measured by salivary cortisol were associated with better WM performance on the 1-back task. There was no correlation between DHEA levels and performance on dMTS ($r_{18} = 0.06, p = .82$), 1-back ($r_{18} = -0.26, p = .27$) or 2-back ($r_{18} = -0.19, p = .43$).

**Transition from one modality to another.** There was no difference in cortisol ($t_{18} = -1.19, ns, d = -0.45$), DHEA ($t_{18} = 0.25, ns, d = 0.17$) or STAI ($t_{18} = 0.23, ns, d = 0.10$) levels between those participants who were trained and tested on the same modality in the OR and in the simulated combat scenario ($n = 10$) vs. those participants who transitioned to a different modality in the simulated combat scenario ($n = 10$). Similarly, there was no difference in performance on the dMTS ($t_{18} = -0.50, ns, d = -0.22$), 1-back ($t_{18} = -1.36, p = ns, d = -0.61$), or 2-back ($t_{18} = 0.38, p = ns, d = 0.17$) between the group that transitioned and the group that did not transition between modalities from the OR to the simulated combat scenario.

**Effects of stress and cognitive function on course performance.** Using logistic regression and the stepwise method of entry, we examined the effects of STAI, n-back, dMTS, and salivary cortisol and DHEA levels on pass/fail rates on each of the five combat casualty care skill in post-OR and in the simulated combat scenario. None of these factors predicted pass/fail rates ($p > .05$).

**Discussion**

In this pilot study we tested the relative effects of two different training modalities (LT vs. SIM) on performing five specific combat casualty care skills in the TACMED course. Whereas
earlier we reported our findings on the effects of training modalities on pass/fail rates for each skill (Savage et al., 2015), here we focused on examining the effects of LT and SIM on stress and cognitive function, and their relationship with performance during skill assessment. Our first hypothesis stated that LT and SIM would be associated with different stress levels. This hypothesis was not supported. Specifically, although levels of cortisol and DHEA were elevated both post-OR and during the simulated combat scenario, there was no interaction with the training modality. In addition, although STAI scores were elevated post-OR, there was likewise no interaction with the training modality. These results suggest that SIM and LT do not exert varying effects on stress in the present context.¹ From a training perspective, the present findings do not support the argument that SIM or LT exerts a greater level of stress during training, and thereby more closely resembles the higher stress level encountered in the battlefield.

Interestingly, performance on 2-back—the more cognitively demanding level of the n-back task—was better than baseline at both post-OR and during the simulated combat scenario. The same was true for dMTS, although the difference at post-OR was only marginally significant (i.e., $p = .05$). These improvements in cognitive function corresponded to elevations in cortisol and DHEA levels post-OR and during the simulated combat scenario. According to our second hypothesis, we expected to see a negative correlation between stress and memory performance. In fact, contrary to our expectation we found that higher cortisol levels were associated with better performance on the 1-back task. One possible explanation for the

¹ However, this conclusion must be qualified because cortisol and DHEA do not represent the full spectrum of the body’s response to stress.
observed correlation between cortisol and WM function is offered by Porcelli et al. (2008), who observed elevated signals in the prefrontal cortex (PFC) in response to exposure to an acute stressor while performing a WM task. They suggested that “stress-related PFC activation increases reflect operations in the service of maintaining organized behavior during stress. Behavioral research highlights the role of PFC in resistance to interference or distraction via executive processes fundamental to WM capacity” (Porcelli et al., 2008, p. 287). In other words, in the presence of acute stress PFC might exert top-down control over behavior, reflected by preserved or improved performance. From a training perspective, our results suggest that the current TACMED course is inducing stress levels in trainees that could be beneficial to cognitive function, although the specific source of the measured stress remains to be determined. However, this inference must be qualified by two factors. First, given that participants were tested at multiple time points, learning and practice effects cannot be ruled out. Second, given that this is a pilot study, the results will have to be replicated in a larger study to determine their robustness (see section below on Limitations).

Next, we tested the hypotheses that transitioning from one modality to another would result in increases in stress and decreases in cognitive function—assessed using salivary cortisol and DHEA levels (for stress) and tests of WM and STM (for cognitive function) (hypotheses 3-4 respectively). Our results did not support these hypotheses. Interestingly, we had previously shown that the transition from one modality to another was associated with lower pass rates on TACMED skills (Sullivan-Kwantes et al., 2015). The present results suggest that mechanisms other than stress and short-term and working memory might underlie this transition-related impairment in skill performance.
Our fifth and sixth hypotheses involved whether pass/fail rates on combat casualty care skills are related to cognitive function and stress, respectively. These hypotheses were not supported: Cognitive function and levels of self-reported and biomarkers of stress did not predict pass/fail rates on any skill. These results are inconsistent with findings linking hypothalamus-pituitary-adrenal (HPA) axis and autonomic nervous system reactivity to poor task-specific performance during combat casualty simulation (McGraw et al., 2013). Given that a critical question in designing the TACMED course involves the induction of the optimal level of stress for maximizing learning and forward transfer of skills, additional research is needed to determine more conclusively the relation between stress and skill acquisition in this context.

Limitations. The three major limitations of this study involve low statistical power due to the small sample size, possibly suboptimal criteria for skill assessment (i.e., performance on combat casualty care skills), and limited options regarding the SIM patient model. We offer the following considerations in relation to each limitation. First, our sample size was small because our effort involved a pilot study. In turn, the small sample size reduced the statistical power associated with our analyses. We hope that future studies will attempt to replicate our findings using larger sample sizes (see Future research).

Second, because skill assessment was based on pass/fail rates, our measures likely did not have the necessary level of granularity to assess more subtle variations in the quality of performance. This leaves open the possibility that modality might have an effect on how well aspects of each skill were performed, rather than simply the overall pass/fail outcome. Related to this issue, it is likely that each combat casualty care skill incorporates multiple surgical competencies. As such, any possible effect of modality might be more detectable if the focus
were shifted to more specific sub-skills in future studies. In addition, our study did not address whether the training modality—SIM or LT—provides sufficient mechanical and tactile cues for trainees to develop the “tissue feel” required to perform any TACMED task. That is, does the modality allow trainees to learn how to identify and locate critical tissue structures and damage in a major wound by touch, in cases where visual cues are insufficient and/or unavailable? Broadening the assessment criteria to include these more comprehensive and fine-grained measures in future studies will likely contribute to a better understanding of the relative merits of SIM and LT for training medics.

Third, the CAESAR™ was selected primarily because (a) our resources did not allow us to test or employ more than one patient simulator, and (b) the CAESAR™ was the most easily accessible high-fidelity patient simulator for combat casualty care in Canada at the time. In addition, it was beyond the scope of this study to evaluate those features of the available patient simulators that best suited the TACMED course requirements. For example, we had no a priori way of knowing whether the fixed set of bleeding locations comprised an ideal match for training the required wound care skill set. Along similar lines, there is no published research concerning whether this simulator, or any other commercial patient simulator, has sufficient fidelity to support the TACMED training. Thus, this study was a first step toward obtaining insights on the fidelity requirements for simulators for TACMED training.

Lessons learned. We employed the present design (i.e., prospective, single blind, randomized controlled study) because we were particularly interested in the effect of transition between modalities on forward transfer of skills (Figure 1). However, the present design could not address what is perhaps the key question of interest, namely the extent to which skills
learned on either modality contribute to better treatment of preventable causes of death on the battlefield. In this sense, an important lesson learned from the present study is to improve the linkage of data from similar studies to battlefield metrics of ultimate importance. An important starting point would be engagement with subject matter experts with experience in the treatment of battlefield injuries to acquire a better sense of how variations in TACMED training features contribute to on-the-job performance.

Second, we collected data on the relevant measures at multiple time points throughout the study (see Savage et al., 2015 and Figure 1). Follow-up research building on this work should consider whether the number of data collection points can be reduced in light of the specific hypotheses under consideration. For example, if the key question of interest is further examination of transition-related changes in performance, stress, and cognitive function, it might not be necessary to collect any data at OR, limiting the focal analyses to a comparison of data collected following the simulated combat scenario with baseline data.

Future research. da Luz et al. (2015) ended their recent review of this literature by offering a list of suggestions for future research. They noted that although SIMs have begun to replace the use of live animals in trauma training, problems remain due to limited realism and authenticity. For example, patient simulators that do bleed do not respond in the same biological way that bleeding patients do in clinical settings. Likewise, research in bleeding and hemorrhage control is never conducted using mannequins because they cannot simulate bleeding accurately. These shortcomings highlight a need to develop better clinically relevant simulation models. LT has its own shortcomings, suggesting that LT and SIM could contribute synergistically to training (Barnes et al., 2016). da Luz et al. (2015) also noted a need for
properly powered randomized control trials, and the use of outcome measures implemented in actual operative circumstances. Echoing that last point, the combined results of Savage et al. (2015) and the present study highlight the urgent need to study the extent to which training on any model—animal or otherwise—transfers to battlefield trauma care involving real casualties.
References


meeting proceedings *Health surveillance and informatics in missions: Multidisciplinary approaches and perspectives*. Paris, France: STO/NATO.


Figure 1. A graphical representation of the study design.

Notes. LT = live tissue, SIM = simulation, OR = Operating Room. We employed a prospective, single-blinded, randomized controlled study with a 2–arm parallel group design to assign participants to train on one of two modalities in the OR. The participants were again randomized before being assessed during a simulated combat scenario on either the same or different modality than what they had trained on—giving rise to a $2 \times 2$ design with equal participants in each cell (see text).