MEETING WARFIGHTER MEDICAL CHALLENGES
Proceedings of the 2014 Military Health System Research Symposium

Foreword
Military Medicine: The Future and Beyond
  Dallas C. Hack

Issue Overview
Pushing Boundaries
  Bruce A. Doll

Editorial
Publishing Military Medical Research: Appreciating the Process
  Jennifer J. Hatzfeld, Ramin A. Khalili, Teresa L. Hendrickson, Patricia A. Reilly

Editorial
Military Graduate Medical Education Research: Challenges and Opportunities
  Matthew A. Borgman, Eric A. Elster, Clinton K. Murray, Jonathan Forsberg, Arthur L. Kellermann, W. Scott Jones

2014 Plenary Presentation
A Large-Scale Informatics Database to Advance Research and Discovery of the Effects of Mild Traumatic Brain Injury

2014 Plenary Presentation
The Natural History of Acute Recovery of Blast-Induced Mild Traumatic Brain Injury: A Case Series During War
  David T. Larras, Walter Carr, Elizandro G. Gonzales, Jason S. Hawley

Repealed Low-Level Blast Exposure: A Descriptive Human Subjects Study
  Walter Carr, James R. Stone, Tim Wallikko, Lee Ann Young, Tianlu Li Snook, Michelle E. Paggi, Jack W. Tsao, Christopher J. Jankosky, Robert V. Parish, Stephen T. Ahlers

Simple and Procedural Reaction Time for Mild Traumatic Brain Injury in a Hyperbaric Oxygen Clinical Trial
  Susan Churchill, R. Scott Miller, Kayla Deru, Steffanie H. Wilson, Lindell K. Weaver

Neurosensorily Assessments of Concussion
  Mark Tommerdahl, Robert G. Dennis, Eric M. Francisco, Jameson K. Holden, Richard Nguyen, Oleg V. Favorov

2014 Plenary Presentation
An Interim $L_{AEq8}$ Criterion for Impulse Noise Injury
  Brissi Zagadou, Philemon Chan, Kevin Ho

2014 Plenary Presentation
Impulse Noise Injury Model
  Philemon Chan, Kevin Ho, Allen F. Ryan

A Comparative Case Study of Risk, Resiliency, and Coping Among Injured National Guard
  Lisa A. Gorman, Angela J. Huebner, Mara K. Hirschfeld, Sudha Sankar, Adrian J. Blow, Danielle Guty, Michelle Keen, Joel S. Ketner

Physical Training Outcome Predictions With Biomechanics, Part I: Army Physical Fitness Test Modeling
  Bryant L. Sih, Charles H. Negus

Physical Training Outcome Predictions With Biomechanics, Part II: Overuse Injury Modeling
  Charles H. Negus, Bryant L. Sih

Effectiveness of an Injury Prevention Warm-up for Unit Physical Training: A Case Series of Two Flying Squadrons
  Joshua D. Shumway, Danielle N. Anderson, Brett Bishop

2014 Plenary Presentation
Pulmonary Emboli and Deep Vein Thromboses: Are They Always Part of the Same Disease Spectrum?
  Nicole T. Gordon, Martin A. Schreiber

Dose Responses of Ibuprofen In Vitro on Platelet Aggregation and Coagulation in Human and Pig Blood Samples
  Wenjun Z. Martini, Cassandra M. Rodriguez, Rodolfo Deguzman, Jessica B. Guerra, Angela K. Martin, Anthony E. Pusateri, Andrew P. Mag, Michael A. Dubick

In Trauma, Conventional ROTEM and TEG Results Are Not Interchangeable But Are Similar in Clinical Applicability
  Sandro Rizoli, Arimie Min, Adic Perez Sanchez, Pang Shik, Richard Grodecki, Precilla Veigas, Henry L. Peng

Bayesian Scoring Systems for Military Pelvic and Perineal Blast Injuries: Is it Time to Take a New Approach?
  Somayeh Mossadegh, Shan He, Paul Parker

Preflight Variables Are Associated With Increased Ventilator Days and 30-Day Mortality in Trauma Casualties Evacuated by Critical Care Air Transport Teams: An Exploratory Retrospective Study
  Ed Barnard, Alejandra G. Mora, Vikhyat S. Beharta
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dysphagia Management and Research in an Acute-Care Military Treatment Facility: The Role of Applied Informatics</td>
<td>138</td>
</tr>
<tr>
<td>En Route Use of Analgesics in Nonintubated, Critically Ill Patients Transported by U.S. Air Force Critical Care Air Transport Teams</td>
<td>145</td>
</tr>
<tr>
<td>Is Hydroxyethyl Starch Safe in Penetrating Trauma Patients?</td>
<td>152</td>
</tr>
<tr>
<td>Performance of Portable Ventilators Following Storage at Temperature Extremes</td>
<td>156</td>
</tr>
<tr>
<td>Evaluation of Oxygen Concentrators and Chemical Oxygen Generators at Altitude and Temperature Extremes</td>
<td>160</td>
</tr>
<tr>
<td>2014 Plenary Presentation Recognition and Treatment of Nerve Agent Casualties: Evidence of Reduced Learner Engagement During Video-based Training</td>
<td>169</td>
</tr>
<tr>
<td>System Design Verification for Closed Loop Control of Oxygenation With Concentrator Integration</td>
<td>177</td>
</tr>
<tr>
<td>An Overview of the Efficacy of a Next Generation Electroceutical Wound Care Device</td>
<td>184</td>
</tr>
<tr>
<td>Telehealth Coaching: Impact on Dietary and Physical Activity Contributions to Bone Health During a Military Deployment</td>
<td>191</td>
</tr>
<tr>
<td>Evaluation of Miniature Wireless Vital Signs Monitor in a Trauma Intensive Care Unit</td>
<td>199</td>
</tr>
<tr>
<td>Developing a Cognitive and Communications Tool for Burn Intensive Care Unit Clinicians</td>
<td>205</td>
</tr>
<tr>
<td>Adaptive Virtual Reality Training to Optimize Military Medical Skills Acquisition and Retention</td>
<td>214</td>
</tr>
<tr>
<td>Face Transplantation in a Highly Sensitized Recipient</td>
<td>221</td>
</tr>
<tr>
<td>Development of Lyophilized Loop-Mediated Isothermal Amplification Reagents for the Detection of Leptospira</td>
<td>227</td>
</tr>
<tr>
<td>Magnetic Nanoparticles as a Potential Vehicle for Corneal Endothelium Repair</td>
<td>232</td>
</tr>
<tr>
<td>Evaluation of a Bovine Vascular Graft in Sheep</td>
<td>240</td>
</tr>
<tr>
<td>2014 Plenary Presentation Exploratory Laparotomy for Proximal Vascular Control in Combat-Related Injuries</td>
<td>247</td>
</tr>
<tr>
<td>Intra-Arterial Perimortem Resuscitation Using a Micellar Colloid</td>
<td>253</td>
</tr>
<tr>
<td>Screening of Nanoemulsion Formulations and Identification of NB-201 as an Effective Topical Antimicrobial for Staphylococcus aureus in a Mouse Model of Inflamed Wounds</td>
<td>259</td>
</tr>
<tr>
<td>The 2014 Military Health System Research Symposium Awards: Recognizing Those Pushing the Science Envelope</td>
<td>265</td>
</tr>
</tbody>
</table>
As we close another successful Military Health System Research Symposium (MHSRS), I am struck by the advancements in military medicine over the past decade. There are a number of parallels between the Department of Defense’s (DoD) flagship medical research symposium and the entire field of military medicine. In both cases, the whole is greater than the sum of its parts. In both cases, the many successes have come about only through the continuous and dedicated work of thousands of military and civilian researchers and health care providers.

These are the changes that bring me the greatest satisfaction as I reflect on my tenure in leading both the MHSRS (and its predecessor, the Advanced Technology Applications in Combat Casualty Care [ATACCC] conference) and the DoD’s Combat Casualty Care Research Program.

Medical breakthroughs are the result of hard work, not luck. However, if we limit ourselves to a single and confined location without reaching out across organizations, we deny ourselves of a key requirement of success: synergy. By that I mean the new ideas that develop when people of varying backgrounds meet to discuss their common mission and collaborate. At the MHSRS, research is shared, relationships are developed, and partnerships are formed—all with the intent that we can improve the lives of the military warfighter. I think of our work in the field of traumatic brain injury (TBI) and how tireless we’ve been in the pursuit of answers, solutions, and of guidelines. A decade ago, we were trying to elbow our way into the conversation. Today we are the acknowledged leader in nurturing change in the field of TBI.

Progress in TBI is but one example. In this Supplement to Military Medicine you’ll see burgeoning research in biomechanics and overuse injuries, as well as a focus on virtual reality and telemedicine; both emerging frontiers that we must begin to fully understand and incorporate into the work we do. In fact, MHSRS is at once both a product of its time and a step ahead of its time. We do not dwell on past successes here, but prefer to build upon them. We must take today’s marvel and develop tomorrow’s “next big thing.” We must take time to nurture fledgling ideas, yet always keep an eye on where those ideas may lead us.

Seven years ago, it would have been difficult to imagine that this work and this symposium would accelerate at such a pace. In 2014, more than 1,630 military, civilian, and academic scientists registered for MHSRS—a 28% increase over the previous year’s tally. In addition, more than 1,130 abstracts were submitted along with more than 400 posters—by far the highest numbers, respectively, in this event’s history. We look forward to more attendees, more knowledge, more science, and more energy in the years to come.

As I wind down my own involvement with MHSRS I am particularly struck by the progress we’ve made and by the coalescence of our collective efforts. The work we’ve done has certainly made a difference to the warfighter, in the broader context of warfighter health, and how we prevent and treat deployment-related medical issues. Such achievement is rightfully to be expected, given the resources applied and the extent of our efforts, the dynamic nature of our work, and the rapid cross-fertilization of ideas that MHSRS provides. As I look back, I am deeply gratified at the broad progress that has been made. I hope you all share that same thankfulness, that same appreciation for the past—and that same hope for the future.

I have hundreds of people to thank as I reflect on the work that has led to this success. Their dedication to the mission has changed medicine for the better, both for the military warfighter and for the civilians who suffer trauma, whether physical or psychological, or are exposed to environmental challenges and infectious agents. I would also like to thank everyone who committed their time to the success of MHSRS and ATACCC. There are far too many to enumerate here, but the following individuals have been especially supportive over the past 7 years: Robert Sarvaideo, Dr. Pat Reilly, Dr. Dave Baer, Maj. Kevin Kupferer, and Peggy Thomas. Your dedication has been vitally important in achieving such progress.

This teamwork has made, and continues to make, the MHSRS so productive. The entire continuum, from the assessment of the needs to the generation of ideas to the entire proposal generation and review process—all of which includes fastidious attention to detail in their execution; including the overall round-the-clock effort required to make even the smallest change or improvement—these advancements cannot be conceptualized, validated, or implemented by one person alone. This is the main reason MHSRS continues to be such a success. This is just one more reason we continue to be—as a team, as a unit, as a military—greater than the sum of our parts.
Pushing Boundaries

RADM Bruce A. Doll, USN

Years ago, I had the opportunity to meet former medical leaders from countries where military personnel did not benefit from close access to medical care while on a mission. This lack of access affected the performance and willingness of their service members to perform. The prospect of being injured with little or no access to life-saving medical care—let alone cutting edge diagnostics and treatment—presented a dismal picture to those asked to go in harm’s way. This is not the prospect our military medical community accepts. This is a mandate for the military medical research community. We must always press for continual improvement. The Military Health System Research Symposium (MHSRS) accelerates recognition of where we stand and where we will head to meet new threats.

Excelling at the tip of the spear for medical research means pushing boundaries on a day-in and day-out basis. It’s what we do—well! Staying on the leading edge, developing the expertise, obtaining the funding, and maintaining the facilities and technologies to meet this opportunity is demanding and exciting. As you read this supplement, you will realize this effort is simply part of our collective job description, performed by thousands of dedicated scientists and clinicians like you, and a vital way that we support the efforts of the military warfighter. Regardless of your role within military medicine, I suggest this issue is particularly relevant to your ongoing commitment to stay on the leading edge.

This supplement primarily contains information presented at the 2014 MHSRS, which established new attendance and submission numbers for poster and podium presentations. We are growing in numbers and impact! Over the past few years, MHSRS has developed into a widely attended and informative experience for the future of our research community and has also provided aspiring graduate students and young researchers the opportunity to interact with the military medical science community. There are numerous individuals and organizations that have contributed to the success of this annual event, which has been steadfastly coordinated by COL Dallas Hack and Dr. Terry Rauch and a dedicated team of professionals. Together with this supplement, MHSRS responds to critical need that not only disseminates research findings, but also contributes to the development of a future military medical research agenda and, in addition, lays the foundation for a continual refreshment of talented, insightful researchers.

As addressed in many of the presentations and captured in this collection of articles from 2014, there was a clear sense at the meeting that military medicine must stay current and relevant despite the drawdown of combat operations. Military policies and training efforts are an important way to maintain currency, but research has the potential to revolutionize military medicine for the future. Now, in retrospect, it is clear that MHSRS provided not only an important opportunity to assess the currency and relevance of ongoing medical research, but it also affirmed our current tack on areas that were anticipated to be important in the near and long term.

The Defense Health Program has designated six different broad categories to fund research, including Medical Simulation and Information Sciences (Joint Program Committee [JPC]-1), Military Infectious Diseases (JPC-2), Military Operational Medicine (JPC-5), Combat Casualty Care (JPC-6), Radiation Health Effects (JPC-7), and Clinical and Rehabilitative Medicine (JPC-8). Each one of these areas addresses an important element of military medicine, and aggressively serves as a catalyst for further innovation. In many cases, these JPCs have funded the work presented at MHSRS and included in this supplement. Furthermore, these programs ensure that the research findings inform operationally relevant clinical practice and advance the broad scope of military medicine.

This particular supplement covers this same broad spectrum of topics, but the individual articles have been further categorized to capture the common themes of Human Performance, Brain Injuries, Trauma System, Leveraging Technology, and Emerging Care, as highlighted in Table I. Together with the call to publish military medical research findings and the need to actively support research within a Graduate Medical Education program, each article contributes to a comprehensive effort to further improve the medical care we provide to the warfighter. Our shared synergy enhances a ready medical force and a medically ready force.

The articles included under the “Human Performance” category focus on the human body and its abilities as a wholly functioning unit. Not only do the topics of these articles address physical fitness, they also cover resilience and stress as well as responses to noise injuries. The focus here is on overcoming physical and psychological stressors and the human body’s ability to react and return to action.

“Brain Injuries” is another important category to military medicine, as we acknowledge that these injuries occur in many different settings and, as such, make it difficult to fully realize the impact of traditional training environments on brain function. From focused studies on low-level blast exposure to specific case studies on mild traumatic brain

U.S. Army Medical Research and Materiel Command, 810 Schenep Street, Fort Detrick, Frederick, MD 21702-5000.

The views expressed in this article are those of the author and do not necessarily represent the official position or policy of the U.S. Government, the Department of Defense, or the U.S. Army. doi: 10.7205/MILMED-D-15-00520
injuries occurring in combat, this year’s supplement poses questions that push beyond what we currently know in order to focus on what we need to know for the future.

The articles within the “Trauma System” category focus on developing knowledge that can be applied throughout the continuum of combat casualty care. Not only does this involve selecting the appropriate intravascular fluids for trauma patients with penetrating injuries, but also identifying factors that influence the risks and assessment of hemorrhage and coagulopathy, systems to predict outcomes following pelvic and perineal blast injuries, the impact of preflight variables and analgesic use during critical care transport, and the value of data during rehabilitation. As learned from 13 years of caring for combat injuries, all of these tools and techniques contribute to the ultimate recovery of warfighters with traumatic injuries.

“Leveraging Technology” is a category focused on both developing and adapting new technology to improve the military health system. In some cases, this technology has been developed to better train or support health care teams via automated communication systems and video training or virtual reality. Other technologies such as telehealth coaching can provide important medical advice to military members in the deployed environment, even when their access to direct medical care is limited. Additional efforts have been made to improve the performance of existing medical devices for vital signs monitoring, wound care, and ventilation. Understanding the performance of these devices in more austere environments
is particularly helpful in planning for future contingencies, and to ensure that safe patient care is being provided.

The articles in the “Emerging Care” category push the boundaries of how we currently provide medical care, including the use of new assays to detect important organisms, nanoparticles to treat infections and repair injuries, and micellar colloids for resuscitation after blood loss. New surgical techniques—including some of the challenges and difficulties—are also described in this section, including those focusing on hemorrhage control, blood vessel repair, the use of regenerative medicine modalities, and even techniques to support successful face transplantation.

These topics areas do not encompass the whole of our efforts. However, they clearly support the sustained direction and commitment of the military medical research agenda. Many other efforts, including the control of bleeding, en route care, precision medicine, pandemic surveillance, global health engagement, and combating antibiotic resistance are also important elements of the ongoing and future directions of research. Just as in the operational setting, efforts routinely blend across dividing lines, blurring definition, and distinction. And yet we also know that only through such a harmonized effort—an effort so reliant on such a unique combination of specificity and collaboration—can our goals truly be accomplished. It is integral then that we focus on the discovery, development, and delivery of medical products, knowledge and material, as we anticipate the medical needs of our service members.
Publishing Military Medical Research: Appreciating the Process

Lt Col Jennifer J. Hatzfeld, USAF NC*; Ramin A. Khalili, BA*; Teresa L. Hendrickson, MAT*; Col Patricia A. Reilly, USAF BSC (Ret.)†

There are many compelling reasons to publish the findings of medical research. Perhaps the most important reason can be found in the need to provide evidence-based recommendations which contribute to improved population health. Aside from being a key driver to better medical care, such research also serves to identify tangible ways to prevent illnesses and injuries, and to enhance individual physical and mental performance. Once completed, the process of disseminating these data must be one that enables easy accessibility to clinicians and thought leaders. Without this effort, medical research findings have a limited ability to influence current and future practice, training, and even policy. As a medical publishing community, effective dissemination of this practice-changing information may be our most important responsibility.

In a military setting, the publishing of evidence-based data in peer reviewed journals is particularly critical. The health of military members, as well as their families, directly influences the ultimate success of the military mission. The phrases “return to duty,” “resilience,” “force health protection,” and even “esprit de corps” all serve to address seemingly trivial activities such as eating, sleeping, fitness, and the balancing of work and family responsibilities. Yet as military clinicians, we recognize that these directly influence the ability of a service member to excel in their assigned tasks or to effectively deploy at a moment’s notice. Just as importantly, this evidence about health also includes the ability to recover from life-threatening injuries in a way that inspires the other members of a military unit to continue a potentially dangerous mission, and also maintain important and healthy personal relationships. It is for these very reasons that publishing military medical research is so important. Why then, does it take so long to get results published?

Research itself is methodical work. It begins with the careful development of a question to be addressed, then assembling a study team that brings the right balance of skills and expertise to the project, followed by spending time carefully crafting a research plan that can adequately answer the question. Executing this plan—collecting the information at the right times and in the right way, meticulously entering and maintaining the data, and thoughtfully analyzing the results—can be just as difficult with many competing priorities and unanticipated events that occur during the course of the study. This is particularly true for military clinicians, whose research responsibilities are often assumed in addition to their primary clinic or hospital duties.

Disseminating research findings is just as challenging. Both military and academic medical researchers cite challenges, delays, and pressures related to publishing completed research. Producing an article that frames the results of a complex research study in such a way that truthfully captures the implications, while at the same time putting the results into an operationally relevant context, is a skill that must be carefully developed over time. Just as importantly, the study results must also be integrated with other evidence to ultimately influence clinical practice, training requirements, and policy decisions. This final step is critical to truly advance the mission of military medicine.

Looking back on the construction of this particular supplement, the editorial process tackled similar challenges. The publication of this supplement extended beyond simply compiling dozens of articles and making sure each conformed to a specified format—although those were important tasks. It is especially important to recognize the efforts of the scientific and medical personnel providing a critical peer review of the evidence reflected in each study. These reviewers willingly took time out from their daily routines to process, analyze, and critique each article with the hopes of achieving—in tandem with the study author, of course—an article that could contribute to the science of military medicine in a meaningful way. In addition to the reviewers, there were many others who endeavored to make sure every “i” was crossed and every “t” was dotted; even the seemingly simple steps such as tracking author responses and creating the layout of this year’s cover illustration were labored over to ensure this supplement reflected the breadth and depth of the research it represents.

There are many correlations between research and publishing: the moving parts, the attention to detail and schedule, and the way incremental steps are required for even the smallest major change. There is an important takeaway in establishing that connection—one that benefits us all as we strive toward our mutual goal of aiding the warfighter: an appreciation in contributing to the “process.” There is an understated effectiveness of methodical work. Much like the turn of an ignition switch that starts a car engine, which then drives effortlessly down the highway, or a single vaccine that initiates a cascade of events, which ultimately provides...
immunity from a deadly disease, each of the parts of a given system are complex, but must move in perfect synchronicity to achieve the desired outcome. It is easy to appreciate the final result; it is much less common to appreciate all of the steps required in between. Perhaps, the effectiveness and complexity of the process needs a brighter spotlight from time-to-time, a greater appreciation for the hard work from a multitude of contributors that result in a final product that can change the science of military medicine.

The Military Health System Research Symposium is the Department of Defense research community’s opportunity to shine the spotlight on the effectiveness of militarily relevant research and recognize those who have contributed to it. In that same spirit, we take this opportunity to acknowledge the research publishing process; whether staff or volunteer reviewers, we are all humbled by the opportunity to contribute to the military medical mission. It is with a renewed respect for this greater mission that we are proud to present the Military Health System Research Symposium 2014 Supplement to Military Medicine.

REFERENCES
Military Graduate Medical Education Research: Challenges and Opportunities

LTC Matthew A. Borgman, MC USA†; CAPT Eric A. Elster, MC USN†; COL Clinton K. Murray, MC USA†; CDR Jonathan Forsberg, MC USN††; Arthur L. Kellermann, MD, MPH†; Col W. Scott Jones, USAF MC (Ret.)††

INTRODUCTION
Over the past 150 years, military medical research has made significant contributions to American health care. In the past decade alone, military medicine has achieved significant advances in combat casualty care\(^5\)\(^6\) by demonstrating the value of tourniquets\(^5\)\(^6\), damage control resuscitation and surgery\(^7\)\(^9\), new infection management guidelines\(^10\)\(^11\), transcontinental critical care transport\(^12\), advanced prosthetics\(^13\), and new techniques in rehabilitation\(^14\)\(^16\). This pioneering research cannot occur in a vacuum. As in civilian settings, medical research must be nested in an academic architecture. In the Military Health System (MHS) the foundation for cutting-edge research is built at the Uniformed Services University of the Health Sciences and through military Graduate Medical Education (GME).

BACKGROUND
The MHS is undergoing unprecedented changes. Force shaping, restructuring, the evolving role of the Defense Health Agency, and the constant pressures of fiscal restraint provide both challenges and opportunities for military GME research. Military GME research represents the scholarship produced by the Accreditation Council for Graduate Medical Education (ACGME) recognized faculty and residents in military GME programs. In this article, we make the case for the value of military GME research, its role in maintaining program accreditation, the ingredients of successful military research, the importance of adequate mentorship, and offer recommendations for the future.

The most obvious rationale for military GME research is that it is required to maintain accreditation by the ACGME. Said another way, military health research cannot continue without GME, and military GME cannot maintain accreditation without research. ACGME guidance dictates that “the faculty must establish and maintain an environment of inquiry and scholarship with an active research component ... [And] should encourage and support residents in scholarly activities.”\(^17\) In addition, at least some faculty in a residency program must demonstrate scholarship through: “peer reviewed funding, publication of original research ... publication or presentation ... at local, regional, or national professional and scientific society meetings ... participation in national committees or educational organizations.”\(^17\) This expectation cannot be met with broad statements of assurance; it must withstand scrutiny through annual, detailed tracking and ongoing accreditation reviews.

Beyond the necessity to maintain program accreditation, military research should be sustained to complement and build on civilian medical research. The nature of the military’s mission requires its physicians to be more adaptable, resourceful, and resilient than their civilian counterparts.\(^18\)\(^20\) In addition, the issues that engage the civilian research establishment are often quite different than those faced by military physicians. For this reason, civilian research universities are neither inclined nor equipped to meet the time-critical requirements of military health research, which constantly evolve to meet new and emerging threats.

In light of these needs, the challenges facing military GME accreditation is concerning. In the fall of 2013, ACGME Residency Review Committees began tracking each resident’s and faculty’s scholarly activities through a Web database (WebADS). A careful review of the 2013 San Antonio Uniformed Services Health Education Consortium (SAUSHEC) WebADS reported scholarly activity determined that only 19% of the 432 core faculty at SAUSHEC had two or more U.S. National Library of Medicine (PubMed) referenced publications. Only 16% had published a textbook chapter or reported leadership through extramural grants. Given that many of SAUSHEC faculty’s scholarly activities are the products of recent graduates of fellowship programs, the percentage of established core faculty producing meaningful scholarship is estimated to be less than 10% based on WebADS data and those with an academic appointment of Associate Professor or higher. One large department had the most appointments to full Professor (6/126 faculty or 4.8%), compared to 65/295 (22%) in the same department at a local civilian academic medical center in South Texas.

The declining level of scholarship in many military GME programs may get worse as a result of increasingly stringent
restrictions on travel to scientific meetings. In 2012, the Office of Management and Budget directed federal agencies to reduce their travel cost by 30%. Almost simultaneously, fiscal challenges secondary to sequestration impacted many federal agencies.

In response to these pressures, the three services implemented travel restrictions making it extremely difficult for military physicians to present their research at non–Federal Entity conferences. This reduced the number of opportunities military faculty need to garner peer review, find mentors and build collaborative partnerships. Connections like these are essential for academic advancement. Although the precise nature of barriers varies by service, the end result is the same: a complicated, costly and time-consuming process that lacks any assurance of success. A team that mapped the Air Force conference approval process for SAUSHEC estimated the local time-cost involved exceeded $1,600; an amount comparable to the average cost of attending a typical non–Federal Entity conference! A SAUSHEC-conducted, confidential survey completed by 21 (81%) Tri-Service Directors of Medical Education found 62% reported most of the faculty and residents were unable to successfully navigate the processes at their institution. Most Directors of Medical Education in the same survey reported that they have “no confidence” in retaining highly qualified faculty or growing future scholars under the current travel restrictions, and most rated the policy’s impact on retention as “medium” or “large.” Travel restrictions were the number two issue listed as impacting satisfaction of 187 SAUSHEC faculty responding to an anonymous institutional survey in December 2013. The following is a representative additional comment from this survey: “Given my desire for a career in academic medicine, I do not intend to extend my service beyond my current obligation . . . .” The policy is having a negative effect on the morale of mid-career military doctors—the backbone of the MHS.

Scholarly productivity is also affected from 2013 to 2014, SAUSHEC Faculty WebADS data indicated a 19% drop in faculty making national presentations and 22% decrease in publications. These observations, combined with data from other sources indicate that restrictions of military physician participation in academic meetings may be having a significant negative impact on their productivity, morale and retention in the MHS. Recognizing these concerns, the Defense Health Agency is in the process of developing new processes.

DISCUSSION

Despite these challenges, many military physicians continue to build their careers by capitalizing on the unique opportunities and mentorship that are still available in the MHS. The successful military academic physician has four key elements: clinical excellence, teaching, research, and operational experience. While the first three are shared similar essentials with the civilian community, operational experience is unique to the military and can inform and motivate a research career with exposure to the medical challenges.

The core competencies of a successful researcher include strength in a content area, the ability to critically review literature, excellent communication skills (presentations and manuscript/grant writing), a solid grounding in research ethics and responsible conduct, and an understanding of clinical epidemiology, biostatistics and study design. Finally, successful researchers need to develop 1 or more sources of funding, whether through the traditional National Institutes of Health pathway (T, K, and R grants), society-specific grants, or Department of Defense sources such as Defense Health Programs and Medical Research and Material Command. It is also important to find one or more research mentors early in the process. Finding selfless mentors, comfortable with their own accomplishments and willing to help others attain academic independence, is crucial to professional growth of young investigators.

Mentorship is challenging within the military system. Although the Department of Defense has a remarkably committed clinical faculty, they tend to be younger and often lack the mentors they need for academic growth. As a result, it can take much longer to accrue a competitive academic portfolio compared to their civilian counterparts. In a confidential survey completed by 187 (42%) SAUSHEC faculty, a lack of research mentors was ranked the fourth most significant issue negatively impacting the training environment trailing only travel restrictions, excessive computer training and inadequate or unqualified support staff. The 20 to 30 year military career is condensed compared to that of a typical academic physician mentor, where a full Professor at an affiliated medical school may enjoy a career that spans 40 to 50 years. Although the military has clear expectations and pathways for developing its operational and health facility leaders, there is a less well-developed track for cultivating its future academic leaders. The U.S. Air Force Medical Service established the “Academic Grand Master” Special Experience Identifiers analogous to the Army “A” professional designator. The U.S. Air Force Medical Service also recently designated Master Academician billets for Colonels at key GME platforms, a very positive step to keep senior scholars doing scholarship. While the other two services have a rich history of a number and broad range of research billets, a recent focus on career broadening assignments even for proven researchers threatens the future of venerable research institutions. Importantly, academic accomplishments that are vital for civilian academic promotion (peer-reviewed manuscripts, society involvement, speaking engagements, mentorship, grant support, etc.) are not consistently valued in the military medical corps promotion process.

In military and civilian academic medical centers alike, aspiring clinician–researchers need the internal motivation to overcome system constraints and persevere over the “long haul” required to produce high-value research. However, institutions can and should play an important role by creating
and fostering a research culture necessary for developing researchers. This entails assigning value to the research process in much the same way hospitals quantify clinical practice in “relative value units” and measuring and reporting research productivity. This would codify the critical element of “time” allocated to productive researchers. In the civilian world, top academic medical centers are regarded as the best places to receive care. The same can and should be true in the MHS.

If the MHS wants to maintain its capacity for innovation demonstrated in the last 10 plus years of conflict, a cultural shift is required. This must begin by identifying individuals with the motivation, interest and requisite clinical skills so they can be groomed early in their career as future academic leaders. These individuals should be connected early on to a highly capable mentor who will maintain contact with them, even when they move from one MTF to another. Faculty development should also foster an environment where select individuals are placed in advance training masters programs in public health or clinical investigation. Researchers must also be provided with the appropriate resources, including dedicated time, support staff, and funds. Barriers should be minimized accessing existing extra/intramural funding. To facilitate this process, the MHS should consider setting up a MHS “venture capital” fund to support targeted seed grants, promising translational research, and support the best and brightest military residents and faculty to present important research findings at national meetings.

CONCLUSION

The cultural shift we envision will be strengthened if the MHS’ top treatment facilities see themselves as “academic medical centers” comparable if not better than the top academic medical centers in the civilian world (Table I). To facilitate this process, all three services should create a better defined and valued research career pathway. Although some military health care providers end their military careers as soon as their service commitment is met, the MHS can create the conditions to retain and develop top doctors, nurses, and other health care professionals to be research leaders. This can be accomplished by offering incentives (award), recognition (titles and promotions), career stability, and intentional mentorship. Finally, efforts to ease the restrictive travel policies that hinder military researchers from collaborating with their civilian counterparts and each other must be successful. If these actions are taken with the ongoing support of the senior MHS leadership, the positive aspects of a military medical career and the commitment to mission that characterizes military health care professionals will carry the day. This will ensure that the MHS always has the cadre of academics it needs to train skilled and versatile providers, manage complex and challenging medical centers, deploy when needed and generate the high-value research required to meet evolving and emerging threats to the health and safety of our armed forces.

REFERENCES


---

**TABLE I.** Key Recommendations for Improving Military Graduate Medical Education Research

| 1. Well-defined and Valued Military Service Research Career Pathways |
| 2. Early Identification of Talented Potential Researchers |
| 3. Align Potential Researchers with Senior Mentors |
| 4. Provide Resources for Researchers: Training, Dedicated Time, Support Staff, Funding Source |
| 5. Recognize Researchers Through Promotion, Titles, and Career Stability to Retain Senior Research Leaders |
| 6. Minimize Restrictive Travel Barriers to Promote Collaboration With the Civilian Community |

**MILITARY MEDICINE**, Vol. 181, May Supplement 2016
A Large-Scale Informatics Database to Advance Research and Discovery of the Effects of Mild Traumatic Brain Injury

Jesus J. Caban, PhD*; Col Albert Bonnema, USAF MC†; Eddy R. Bueno, MBA*; Thomas DeGraba, MD*; COL Geoff Grammer, MC USA*; CAPT Walter Greenhalgh, MC USN*; CAPT Sara Kass, MC USN*

ABSTRACT Clinical research advances in traumatic brain injury (TBI) and behavioral health have always been restricted by the quantity and quality of the data as well as the difficulty of collecting standardized clinical elements. Those barriers, together with the complexity of evaluating TBI, have resulted in serious challenges for clinicians, researchers, and organizations interested in analyzing the short- and long-term effects of TBI. In an effort to raise awareness about existing and cost-effective ways to collect clinical data within the Department of Defense, this article describes some of the steps taken to quickly build a large-scale informatics database to facilitate collection of standardized clinical data and obtain trends of the longitudinal outcomes of service members diagnosed with mild TBI. The database was built following the Defense of Health Agency guidelines and currently has millions of longitudinal clinical data points, Department of Defense-wide clinical data for service members diagnosed with mild TBI to support population studies, and multiple built-in analytical applications to enable interactive data exploration and analysis.

INTRODUCTION
The pathophysiological changes in the brain following traumatic brain injury (TBI) remain poorly understood. Post-traumatic alterations in the brain have been difficult to quantify in part because of the lack of standardization across clinical disciplines, the frequency of incomplete data collection efforts, and the lack of unique metrics to measure changes after TBI events. Clinical practice guidelines recommend examining TBI patients with physical, cognitive, behavioral, imaging, and neuropsychological evaluations, which results in a large collection of data elements that should be analyzed collectively to better establish an understanding of the patient’s condition.

Despite the existence of detailed clinical practice guidelines for managing patients with mild TBI (mTBI),2–4 most clinical evaluations still rely heavily on behavioral observations such as the Glasgow Coma Scale and on the patient’s subjective recall of post-traumatic amnesia and loss of consciousness. Since these assessment tools depend on environmental and subjective recollections, they often yield highly variable ratings that can complicate the evaluation process by potentially adding uncertain or conflicting information to the patient record.

The complexity, heterogeneity, and often subjective properties of the data elements related to mTBI, combined with the shortage of HealthIT systems to support the data collection, standardization, and analysis of mTBI data, have created significant challenges. First, in order to develop a comprehensive understanding of the patient’s condition, providers must integrate many multimodal measurements and current clinical systems do not facilitate the collection and analysis of the widely disparate data obtained during the evaluation of TBI. Second, researchers typically must use their own data collection strategies to build a database to facilitate their research. The resulting datasets are not integrated with other researchers, potentially limiting the scope and applicability of some findings.

In general, a key factor that has limited research and innovation in mTBI across the Department of Defense (DoD) has been the lack of standardized large-scale databases integrated with the clinical workflow where hundreds of clinical variables can be collected, analyzed simultaneously, and used to discover new complex patterns.

A 2014 memorandum by the assistant Secretary of Defense for Health Affairs (Woodson J: Guidance for the Management of Registries in the Military Health System. Affairs ASoDfH, ed 2014) provides specific guidance about how DoD registries should employ reusable interfaces and data service. In an effort to show how those guidelines can be followed to enable the development of cost-effective data repositories, this article describes a large-scale database that was designed to collect multidisciplinary clinical elements from dozens of external sources and medical devices. The data registry, which combines DoD-wide TBI clinical data with multiple built-in analytical applications, was designed in partnership with the Defense of Health Agency (DHA) with the goal that other organizations or individuals under the
DoD could leverage the concept, technology, and systems to quickly create their own large-scale data repositories or registries.

BACKGROUND
Clinicians and researchers face several challenges when trying to diagnose patients that have sustained concussions. During the last decade, a significant amount of attention has been given to the acquisition of clinical data from patients who have suffered from mTBI and/or psychological health (PH) conditions. In addition, the DoD as well as many other government and private organizations have been leading efforts to raise awareness about the long-term effects of concussions. This has both increased and influenced the collection and the quality of the data being collected.

To provide a comprehensive interdisciplinary treatment for service members diagnosed with mTBI, the DoD supports several organizations that perform extensive clinical evaluations of patients diagnosed with TBI/PH. At organizations such as the National Intrepid Center of Excellence (NCoE), active duty service members diagnosed with mTBI and PH conditions undergo a 4-week intensive treatment program by a team of experts, including weekly appointments with internists and approximately 105 encounters with diverse clinical staff. This dedicated process produces an extensive amount of data points and measurements that must be collected and used to create rich research datasets.

Despite the efforts to support more acquisition of clinical data, there are still many challenges and questions about how to collect standardized data while not disrupting the clinical workflow and avoiding the clinical staff to enter the same information multiple times in different systems. Currently, most of the attention has been given in creating centralized databases where researchers, once they collect the different clinical elements, can share the data with other collaborators and organizations. However, the question about cost-effective approaches to collect standardized data from providers without disrupting the clinical workflow remains open.

There are four well-known databases that are currently assembling clinical information from TBI patients: “Traumatic Brain Injury Model Systems (TBIMS), Federal Interagency Traumatic Brain Injury Research (FITBIR),” the “Center for Neuroscience and Regenerative Medicine Database (CNRM), and the Gemini program.” Given the limited information publicly available about the technical and engineering design of each of those databases, an in-depth comparison of the databases is not within the scope of this article.

The TBIMS program is funded by the U.S. Department of Education’s National Institute on Disability and Rehabilitation Research, to examine the course of recovery and outcomes following a moderate or severe TBI.5 The program operates a database that receives information from 16 different centers around the country. The primary objective of the database is to generate new and useful knowledge about the short- and long-term effects of TBI based on 327 clinical elements that are collected during inpatient rehabilitation and after discharge.

To address some of the existing gaps surrounding TBI, the National Institutes of Health, in partnership with the DoD, is building a centralized database for TBI research named the FITBIR database.6 The primary objective of FITBIR is to serve as a central repository for TBI data to enable the comparison of results across studies. Since the data going to FITBIR will come from individual federally funded research projects that might not be collecting the same data elements or might not share the data until the end of the project, it may take years to produce a database with sufficient homogeneity and quantity to perform some of the desired population-based studies. To mitigate a prolonged period before a return on investment, strategies are being implemented, such as the development of common data elements (CDEs) to reduce variability in data element reporting and incentives to collect legacy data.7

A similar effort is being driven by the CNRM, which is a federal medical research program operated by the Uniformed Services University of the Health Sciences that brings together clinicians and scientists across different disciplines for innovative TBI research, including approaches for diagnosis and recovery. Internally, CNRM maintains a comprehensive database with de-identified information regarding study participants. The clinical and research elements that are incorporated into the database are predominantly driven by clinical research protocols conducted within the organization.8

Finally, the “Gemini” program is a project under One Mind for Research focusing on exchanging and gathering data from different organizations performing research with patients who have TBI.9 The “Gemini” program employs the principle of open science to collect data from different funded projects with the goal of making them available to the broader population. This database faces the same challenges as other systems as well as the challenge of obtaining data from third-party organizations collecting the information.

Recently, there have been other efforts to build additional large-scale databases such as the Department of Defense and Department of Veteran Affairs Psychological Health & Traumatic Brain Injury registry that is scheduled to be operational by fiscal year 2017.10 Despite the great work that has been done on repositories to store clinical data on TBI patients, there are many challenges with the design and development of many of the existing databases. First, most of the current databases depend on individually funded projects that, as part of their project deliverables, have agreed to contribute their data with one of the existing databases before the completion of the project. Since funding can vary between different clinical disciplines, the focus of research programs can change on an annual basis, research protocols have a specific target population, and different research studies can collect distinctive clinical data; combining and comparing the data produced by individual research studies will be extremely challenging. Second, most of the data
collection and analysis is focused on particular aspects of TBI (e.g., neuropsychological, neuroimaging, sleep, vestibular, etc.) and not necessarily on comprehensive datasets that can enable research on multivariate approaches to model the complex biological state of mTBI patients. Third, none of the existing databases have been incorporated with the clinical workflow. Fourth, most of the existing databases assume that organizations or researchers have data collection mechanisms compatible with existing case report forms such as those used by CNRM or that have e-forms that follow the CDEs suggested for TBI. Fifth, none of the existing databases provide built-in analytical techniques that can be used to analyze, model, forecast, and mine the heterogeneous data. Finally, the existing databases do not incorporate population data such as DoD-wide clinical data for TBI patients.

The primary focus of our work was to fill the gap between the clinical workflow and existing repositories by building a system that follows the DHA guidelines and can be seamlessly integrated with the clinical workflow, used to collect standardized clinical elements, and provide a platform where clinical staff, researchers, and administrators can analyze, explore, and review large collections of interdisciplinary data of mTBI patients.

All of the different components that were used to create our database are owned and operated by the DoD and are available to any other organization that might be interested in leveraging them to create their own database. To disseminate information and help researchers learn about how existing DHA guidelines can be used to create cost-effective systems to collect clinical data, this article describes some of the steps that were taken to quickly build a large-scale informatics database to enable in-depth analysis of patients diagnosed with mTBI. Note that the article describes a platform and database system that was built using applications, technologies, and resources that the DoD currently has available at the enterprise level. Given the rapid changes and evolution of HealthIT systems, some of the design and/or technologies might evolve; however, the overall framework has been designed flexible enough to support any potential changes.

**APPRAOCH**

Instead of relying on investigators to collect and contribute data or on researchers to extract information from clinical notes, we created a clinical repository that automatically receives and integrates data from disparate clinical sources. Figure 1 shows a diagram of our large-scale database. The diagram is divided into two sections to the left of Figure 1 (gray-shaded areas) are the “producers”—those systems that can generate clinical data—and to the right (green-shaded areas) are the “consumers”—those systems that can use the clinical data.

**Producers**

To develop the comprehensive database, we first identified the different sources of information that can produce clinical and

![FIGURE 1. Diagram illustrating the overall design of our large-scale informatics system to support research and innovation in mTBI. The system is capable of aggregating clinical information from DoD-approved systems as well as from local data repositories. AHTLA, Armed Forces Health Longitudinal Technology Application; CHCS, Composite Health Care System; CDR, Clinical Data Repository; DEERS, Defense Enrollment Eligibility Reporting Systems.](image-url)
research data for patients diagnosed with mTBI. The different sources of clinical data were divided into two groups: “trusted clinical sources” and “external/unharmonized sources.”

**Trusted Clinical Systems**

Clinical data collected from electronic health records (EHRs) such as the Armed Forces Health Longitudinal Technology Application (AHLTA) or the Composite Health Care System (CHCS) that have a provider, organization, and/or individual signing the information and is accountable for the accuracy of the data. Those systems cannot be altered, thus the data should be considered as accessible information from a known “trusted source.” Other data coming from additional DoD sources such as the MHS Mart (M2), Pharmacy Data Transaction Service (PDTS), Clinical Data Repository (CDR), Defense Enrollment Eligibility Reporting Systems (DEERS), and TRICARE Encounter Data Institutional (TED_I) should also be considered secure and trusted information.

The DoD has multiple systems used to capture and integrate data from different “trusted” clinical sources. Three of the most popular databases are the Health Services Data Warehouse (HSDW), the Medical Data Repository, and the COHORT Database. After considering the flexibility, scalability, and capabilities of the different systems, we partnered with HSDW to provide us access to the data from different clinical sources. Currently, following the DHA registry guidelines, HSDW is becoming the adopted warehouse to provide clinical information to DHA (Woodson J: Guidance for the Management of Registries in the Military Health System. Affairs ASoDfH, ed 2014).

HSDW is a DHA effort to create a repository of the DoD health information. The system uses multiple internal and commercial tools to quickly integrate and query data from different sources. Typically, EHRs contain billions of data records as well as significant amounts of sensitive information. Due to the size of the data and the necessity of combining data from multiple sources, the clinical data in HSDW flows from different databases to a temporary database where the data is transformed into a star schema that consist of multiple fact tables referencing many different dimension tables. The benefits of the star schema include enabling simple queries, simplified logic, fast performance, and quick aggregations.

Currently, HSDW has connectivity to over 12 different data sources including DEERS, M2, CHCS, COHORT, Picture Archiving and Communication System, and PDTS. Figure 2 lists some of the data sources currently available through HSDW. From those sources we can obtain patient and encounter level information such as demographics, deployment information, type of encounters, diagnosis, procedures, and medication changes. Such information has proven to be valuable for epidemiology studies, trending information, and to obtain indirect measurements of outcomes. HSDW has over 9 billion records containing clinical information and metadata of all encounters that have taken place since 2006.

Table I illustrates the magnitude of the data that is currently available in HSDW. It is important to note that these numbers are DoD-wide numbers, not only for those patients diagnosed with mTBI.

Although encounter metadata and coding information have proven to be beneficial in many population studies, there are also many limitations regarding the uncertainty of the data and the validity of translating coding information into research data. In addition, standardized information or CDEs must be collected within every discipline to successfully accomplish research in mTBI that cannot be done by simply relying on International Classification of Diseases (ICD) or Current Procedural Terminology (CPT) codes. In general, detailed information need to be collected through standardized forms.

To create electronic forms that can be used by providers to collect clinical and research data, the system must be integrated with the clinical workflow and should not add any burdens to the providers. In general, the primary concerns of incorporating standardized forms within clinical settings

**TABLE I. Sample List to Illustrate the Comprehensive Data Available Through Health System Data Warehouse**

<table>
<thead>
<tr>
<th>Type of Records</th>
<th>Number of Records (2006–2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Encounters</td>
<td>2,295,080,000</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>4,092,881,000</td>
</tr>
<tr>
<td>Lab</td>
<td>225,709,110</td>
</tr>
<tr>
<td>Radiology</td>
<td>35,695,174</td>
</tr>
<tr>
<td>Detailed S/O Notes</td>
<td>1,409,000,000 (Between7/13–7/14)</td>
</tr>
<tr>
<td>(Split by MEDCIN)</td>
<td></td>
</tr>
</tbody>
</table>
include the potential of workflow disruptions, concerns about security and privacy, lack of incentives for additional effort that might be required to collect or input the data, the learning curve required to learn a new system, potential of dual entry, and the lack of interoperability between different systems.\textsuperscript{19}

In the DoD, the primary outpatient EHR has been AHLTA and most providers are required to write a clinical note for every encounter. To document the subjective (S) and objective (O) portion of an encounter, AHLTA S/O notes use MEDCIN terminology.\textsuperscript{20} MEDCIN is a commercial product employed in AHLTA that uses a hierarchical structure to enhance the codification of information in EHRs. Aside from MEDCIN terms, AHLTA supports Alternate Input Method (AIM) forms that are templates that can be used to document S/O notes.\textsuperscript{21} AIM forms can be built around a theme, a diagnosis, or a symptom.

To enhance the data collected by HSDW, we built AIM forms for multiple disciplines including Neurology, Nursing, Social Workers, Audiology, Speech, Vestibular, Complementary Alternative Medicine, and other disciplines. By using AIM forms the S/O notes are not just a text field where providers can write free-form text, but instead standardized forms where each button, field, and checkbox has a corresponding MEDCIN term. Figure 3 shows some screenshots of an existing AHLTA AIM forms. Figure 3A shows a screenshot of the Audiology AIM form. The form has different tabs across the top as well as different discrete fields within each tab that have unique MEDCIN terms associated with them. The benefit of using AIM forms is that the providers complete their S/O notes using a form that has specific and standardized data elements. When providers submit a clinical note, each button, checkbox, and text field has a

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{aim-form-screenshot.png}
\caption{(A and B) Sample screenshots of existing Alternative Input Method (AIM) forms used to standardize the collection process within the Armed Forces Health Longitudinal Technology Application (AHLTA).}
\end{figure}
particular MEDCIN term associated with them, thus making the parsing and querying of the data possible.

In collaboration with DHA, we added a new data source and comprehensive database to HSDW—"CDRMart." The CDRMart contains all the S/O notes written in AHLTA, each note can be linked to a specific encounter, and the text of each note can also be split into the corresponding MEDCIN terms. This new functionality currently allows us to pull encounter information (e.g. date, ICD, CPT, etc.) and augment that information with clinical notes either collected through AIM forms or as free-form text. A combination of parsing and natural language processing (NLP) techniques are used to extract relevant clinical information from unstructured data sources. For example, an AIM form containing an injury tab, like the one illustrated in Figure 3B, enables the collection of standardized injury information following the CDE recommendations by seamlessly integrating the forms to the system that all providers use to document their encounters.

In the design and development of the AIM forms, we partnered with the Tri-Service Workflow—a DHA team of clinicians and information technology professionals that standardize clinical documentation and workflows while enhancing provider efficiency and patient safety. All of the AIM forms that were created as part of this project are currently in the enterprise and available to anyone with an AHLTA account and can be easily recreated in other commercial electronic health records (EHRs).

**External Sources**

In addition to trusted sources such as EHRs, hospital organizations, and clinics always have other sources of clinical data that are used by the clinical staff, but often the data does not make it into the patient records. For example, sleep studies, neuropsychological evaluations, and audiology examinations are some of the assessments that employ specialized devices to test patients. Each of the individual computer system collects a number of measurements about patients and often generates summary reports with overall findings. Generally, providers review the aggregate reports and write a subjective description of the findings and diagnosis. Frequently, the data collected by the individual devices or the individual reports are not included in the patient records. Those systems that do not have a connection to the EHRs are considered as "external/unharmonized" sources.

In order to build a comprehensive dataset that can be used to explore the short- and long-term effects of mTBI, data generated by different clinical devices or external sources must be collected. In our approach, we connect medical devices and external data collection sites to a “data augmentation” database that works as an external database to enhance the information collected from HSDW. The augmentation database also works as a temporary platform where parsing, filtering, and standardization of the data produced by different systems that are not connected to any EHR takes place. Figure 1 shows the overall concept of the augmentation system used to collect external data.

Some of the sample external sources that have been incorporated into our augmentation database include data from independent DoD databases such as CarePoint and the Wounded, Ill, and Injured Registry (WIIR).

CarePoint is a centralized platform and hosting environment owned and operated by DHA that was built on top of Microsoft Office SharePoint Services to consolidate multiple DoD applications such as the Military Health Service Population Health Portal or other web-based applications. To provide clinical staff and researchers dynamic forms with built-in decision rules and interactive interfaces that AHLTA AIM forms cannot provide, we designed multiple clinical forms in CarePoint. Figure 4A shows a referral form that was developed to collected in-depth historical information from TBI patients before they attend the NICoE program. Figure 4B shows an interdisciplinary form that was created to effectively collect information from multiple providers belonging to different disciplines.

In addition, data from medical devices such as polysomnogram, actigraphy, and audiogram devices are being pulled and incorporated into the augmentation database. Similarly, to create a comprehensive database that can benefit different researchers and providers interested in understanding the effects of mTBI, many other databases and clinical devices that do not have connectivity to AHLTA have been incorporated into the augmentation database.

Given the importance of collecting short- and long-term outcome measures, within the DoD there are several systems that can be used to collect standardized questionnaires including the Behavioral Health Data Portal, the WIIR, mCare, After Deployment website, and many others. We partnered with the Army Analytics Group in developing a module on the WIIR that can be used to collect longitudinal information from patients that receive treatment at the NICoE. The decision of using the WIIR was made given that it was a very cost-effective and flexible system that could be quickly incorporated into our clinical workflow to collect outcome data directly from the patients. The WIIR, a system owned and operated by the DoD, has many unique features including all the information assurance (IA) certifications, a Common Access Card-enabled interface for providers, reachable from any MTF, and accessible to patients regardless what network, computer, phone, or tablet device they use. The NICoE environment within the WIIR currently has over 30 standardized surveys including PCL-M, NSI, GAD-7, and AUDIT-C. Currently, our clinical staff requests specific surveys to be completed at different time intervals, the patients are automatically notified and/or reminded about the pending surveys, on submission the surveys are scored by the system, standardized plots and tables are generated to show changes, and the clinical staff is notified that the information is available for their review. The use of the WIIR
FIGURE 4. (A) A referral form that was developed within CarePoint to collected in-depth historical information from TBI patients before they attend the NICoE program. (B) An interdisciplinary form that was created in CarePoint to effectively collect information from multiple providers belonging to different disciplines.
has significantly improved our ability to collect outcomes data without adding any significant burden to the patients or providers. The data collected from the WIIR are automatically transferred to our augmentation database and can then be merged with any other clinical element. Figure 5 (left) shows a screenshot of the WIIR module that is currently being used to automatically collect self-reports at the NICOe. Figure 5 (right) lists some of the self-reports that are currently available through the WIIR.

Now that we have an enhanced version of HSDW with structured and unstructured clinical information from different EHRs and a comprehensive augmentation database with data collected from external sites or devices, we need to integrate the data so it can be used to better understand the status of our patients and the effects of mTBI. Because of IA regulations, data from external sources should not be merged with clinical data. To comply with these regulations, although avoiding the expense of having to extract, transform, copy, and load data from multiple systems into a single database, we developed, in collaboration with DHA, a “virtual layer.” The uniqueness of the virtualization layer is the ability for multiple data sources to be brought together through a centralized tool. Instead of copying the data from disparate sources, the metadata from different databases is

---

FIGURE 5. (Left) Screenshot of the NICOe environment within the WIIR that is used to collect patients’ self-reports. (Right) Lists of some of the questionnaires that are currently available in the WIIR.

FIGURE 6. To show the value of our database, the clinical diagnosis of a large DoD population of 89,840 service members that have sustained an mTBI between 2007 and 2014 was analyzed to better understand the differences in symptoms prior and post mTBI. Eight symptoms commonly associated with mTBI are presented with the corresponding percentage of patients with such a clinical diagnosis prior and post their brain injury. It is important to note the significant changes that are apparent in headaches, sleep, neurological disorders, and depression.
copied, the virtual layer hides system-specific data schemes, and during the query process, the data are brought together. The virtual layer has proven to be a very cost-effective way to scale our large-scale informatics database given that additional sources of data can be added at very minimum expense or impact to the system.

**Consumers**

Figure 1 shows the virtual layer as the gate to consumption of the multimodal clinical data coming from different database systems. The two primary ways the data are consumed by users are (a) as “queries” or (b) by “business intelligence (BI) dashboards.” Queries can be written using a variety of applications and tools that can connect, communicate, and query the virtual layer, which connects to many different databases. In addition, the BI Dashboards are created using a combination of popular tools including Statistical Analysis System (SAS), Microsoft Power BI, or Tableau.27,28 The results section will discuss some cases where the data from our large-scale repository were quickly used to extract new knowledge about mTBI.

**RESULTS**

The database described above has been used to support multiple clinical and research studies about the effects of mTBI.29–31 While the main objective of this article is to describe the process that was followed to create a large-scale informatics databases for mTBI patients and encourage other organizations to follow a similar approach, this section will briefly describe some of the results that we have been able to obtain from the database.

Some of our preliminary results include findings in speech and pathology data, neurology, audiology, vestibular, observations about neuroimaging lesions, validation of some self-reports, and the description of the short- and long-term effects of TBI. In addition, comprehensive dashboards have been developed to assist with program evaluation, longitudinal outcomes analysis, analyze DoD-wide trends, and determine organizational metrics.

As an example of the effectiveness of the database, a research query was done to quickly look at the speech-language pathology findings. A total of 485 patients (age $M = 32.47$, $SD = 8.60$) were found with all the information needed for the study. Then, by leveraging the technology of our database system, we were able to immediately estimate that 67% of our patients complained of word-finding difficulties and that the average scores for cognitive evaluations that are related to verbal memory were low or below average while those related to prospective, visual, and spatial memory show strengths among the same service members. The results of combining different databases provide objective support that verbal memory and new learning are consistently the most frequent area of weakness in formal speech-language pathology assessment and that there is a significant difference between verbal and visual memory that becomes more evident among patients who have sustained a concussion.

During the last decade, the sequel of symptoms that occurs most commonly after mTBI has been of great interest.1,2 It is known that patients that suffer a brain injury are more likely to suffer from depression,32 headaches,33 sleep disorder, and other conditions. To describe the difference in symptom configuration within patients before and after their TBI, we used our database to quickly pull the symptoms for a large DoD population of 89,840 service members that have sustained an mTBI between 2007 and 2014. Figure 6 shows eight symptoms commonly associated with mTBI and the...
percentage of patients with such a clinical diagnosis prior and post their brain injury. Note the significant changes that are apparent in headaches, sleep, and neurological disorders. From Figure 6, it is also important to see the significant increase that speech and language disorders have after an mTBI.

Health care utilization has been another important topic that has received a significant amount of attention during the last few years.34,35 Another way to utilize our database is to look at health care utilization trends after mTBI and analyze the impact that multiple concussions have on the health care system. Figure 7 shows the average number of TBI-related clinical encounters service members had 30 days after their first, second, third, fourth, and fifth mTBI event. From the results shown in Figure 7, we can see that on average service members have 3.9 encounters 30 days following their first TBI; however, that number goes to 8 encounters 30 days following their third TBI event.

Another way to consume data is through dashboards. The dashboards provide a single analytical framework where administrators, clinicians, and researchers can analyze and explore clinical data to better understand the effects of TBI, mTBI, post-traumatic stress disorder, and comorbid PH. Multiple dashboards have been created to better understand the effects of concussion as well as the effects that a specific treatment has on patients diagnosed with mTBI. Figure 8 shows four dashboards generated using the data obtained from our comprehensive mTBI database. Figure 8A shows a dashboard with nine different metrics used to monitor outcome measure and monitor patients that have received a specific treatment. Longitudinal data are automatically pulled from the different databases and shown to providers so they can better understand the effectiveness of different treatments among subjects with mTBI. Figure 8B shows a dashboard illustrating the changes of a specific patient. For this particular patient, the users can interactively explore the data and see that the patient is improving in most of the different dimensions, but the severity of the headache seems to be getting worse. Figure 8C shows a screenshot of an interactive tool to explore the health care utilization of mTBI patients.
pre- and post-treatment. Figure 8D shows a dashboard used to monitor the incidence of mTBI across the DoD and forecast demand for TBI-related clinical services.

CONCLUSION

In an effort to raise awareness about how existing DHA guidelines can help organizations design cost-effective ways to collect clinical data within the DoD, this paper described some of the steps that were taken to quickly build a large-scale informatics database to facilitate collection of standardized clinical data and obtain trends of the longitudinal outcomes of service members diagnosed with mTBI. All the systems, components, and applications that were used to build our large-scale data repository are available to organizations and researchers interested in creating their own data repository and trending applications.

ACKNOWLEDGMENTS

We thank the Information Delivery Division (IDD) at the Health Information Technology (HIT) Directorate at DHA, the Space and Naval Warfare Systems Command (SPAWAR), the different contractors, and the NCoE staff for their continued support and contribution to this project.

REFERENCES

in service members receiving intensive treatment for combat-related mTBI. The J Head Trauma Rehabil 2015; 1: 1–8.


The Natural History of Acute Recovery of Blast-Induced Mild Traumatic Brain Injury: A Case Series During War

MAJ David T. Larres, MS USA*; MAJ Walter Carr, MS USA†; SSG Elizandro G. Gonzales, USA‡; LTC Jason S. Hawley, MC USA§

ABSTRACT  Traumatic brain injury (TBI) secondary to blast exposure is a common injury in the Global War on Terrorism, but little is known about the acute effects, recovery, pathophysiology, and neuropathology of blast-induced mild TBI (mTBI) in humans in a battlefield environment. Moreover, there is ongoing debate whether blast-induced mTBI is a different injury with a unique pathophysiology compared with mTBI from blunt trauma. In the case series reported here from Craig Joint Theater Hospital at Bagram Airfield, Afghanistan, 15 military service members with acute concussion/mTBI associated with blast exposure were evaluated within the first 24 hours after concussion and on days 2, 3, 5, and 7 with a Graded Symptom Checklist and a balance assessment, the Balance Error Scoring System. These data suggest that the recovery in blast-induced mTBI follows the pattern of recovery in sports-related concussion reported in The National Collegiate Athletic Association Concussion Study. In this retrospective case series, we provide the first description of the natural history of acute recovery in blast-induced mTBI, and we suspect, given our experience treating military service members, that further observations of the natural history of recovery in blast-induced mTBI will continue to mirror the natural history of recovery in sports concussion.

INTRODUCTION
Traumatic brain injury (TBI) secondary to blast exposure has been a common injury in the Global War on Terrorism, specifically Operation Iraqi Freedom (OIF) and Operation Enduring Freedom (OEF). Of 8,000 cases of brain injury reported by the Defense and Veterans Injury Center, 90% were consistent with mild TBI (mTBI)/concussion. The majority of those injuries were caused by blast mechanisms.1–3 There is a growing body of literature on the physics of blast-induced brain injury, primarily using animal models and computer simulations.4–7 However, data pertaining to human exposure to blast and corresponding effects on the central nervous system are more limited. Within the Department of Defense (DoD) and Department of Veterans Affairs, studies of blast-induced mTBI mainly involve patients who have redeployed with chronic complaints after remote exposure to blast.8–10 Further, studies of blast-related mTBI are frequently limited in specificity due to polytrauma and effects from extracranial injury.11,12 Studies of combat theater mTBI without polytrauma have begun to emerge,13 but those studies have typically not concentrated on mTBI from blast relative to mTBI from other causes.14–17 It remains the case that little is known about the acute effects, recovery, pathophysiology, and neuropathology of blast-induced mTBI in humans in a battlefield environment.18,19 Moreover, there is ongoing debate whether blast-induced mTBI is a different injury with a unique pathophysiology compared with mTBI from blunt trauma (motor vehicle accidents, sports concussion).20,21

To our knowledge, the natural history of acute recovery in blast-induced mTBI has not been reported, so that is the purpose of the present study. This the first report of natural history of recovery in blast-induced mTBI in a combat theater of operations. There are extensive studies, predominantly in the sports concussion literature, detailing the natural history of recovery of sports-related concussion. In 2003, McCrea et al22 published a widely cited study of college football players sustaining concussion (The National Collegiate Athletic Association/NCAA Concussion Study). The authors used a Graded Symptom Checklist (GSC), the Standardized Assessment of Concussion, and a balance assessment, the Balance Error Scoring System (BESS) to outline the recovery time course in symptoms, cognition, and balance proceeding in the days and weeks following concussion. Their results showed that most athletes with concussion recover from symptoms, cognitive dysfunction, and postural instability within 7 days. That study confirmed previous reports including a review from the World Health Organization indicating strong evidence for full recovery of symptoms within weeks after concussion/mTBI.23

*Occupational Therapy Clinic, Tripler Army Medical Center, 1 Jarrett White Road, Honolulu, HI 96859.
†Department of Behavioral Biology, Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910.
‡Occupational Therapy, Guthrie Army Health Clinic, 11050 Mount Belvedere Boulevard, Fort Drum, NY 13602.
§Department of Neurology, Walter Reed National Military Medical Center, 8901 Rockville Pike, Bethesda, MD 20889.

This work was presented at the Military Health System Research Symposium, Fort Lauderdale, FL, August 18–21, 2014 and at the National Capital Area Traumatic Brain Injury Research Symposium, Bethesda, MD, March 3–4, 2014.

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Army, Department of Defense, or the U.S. Government. The authors are military service members. This work was prepared as part of official duties.

Title 17 U.S.C. §105 provides that “Copyright protection under this title is not available for any work of the U.S. Government.” Title 17 U.S.C. §101 defines a U.S. Government work as a work prepared by a military service member or employee of the U.S. Government as part of that person’s official duties.

doi: 10.7205/MILMED-D-15-00152
In 2010–2011, the authors were deployed to OEF to establish a subspecialty team at the Craig Joint Theater Hospital at Bagram Airfield, Afghanistan with the mission of providing subspecialty evaluation and treatment for service members sustaining mTBI. We had the opportunity to evaluate service members with blast-induced mTBI acutely and track symptom resolution. In our treatment program, we used the standardized metrics of the GSC for symptom recovery, and the BESS for evaluating for postural instability. Soldiers were treated per DoD Directive Type Memorandum 09-033, which specifies the acute management algorithms for soldiers sustaining concussion.24 Our experience allowed us to evaluate soldiers sustaining acute concussion from blast and monitor the natural history of recovery in blast-induced mTBI. We report a case series of 15 soldiers, who sustained a blast-induced mTBI. We were able to track their recovery with metrics (GSC and BESS) that have been standard assessment tools in sports concussion and report our observations here.

**METHOD**

We performed a retrospective review of cases in which soldiers sustained a blast-induced mTBI from March to June 2011. The diagnosis was based on the DoD and American Congress of Rehabilitation Medicine definition of concussion/mTBI, which defines the injury as mechanical energy (blast, blunt, fall, etc.) applied to the head with the presence of alteration of consciousness (AOC), loss of consciousness (LOC), or the presence of post-traumatic amnesia after the incident. For all the individual cases in our report, patients were seen within 24 hours of injury and the diagnosis was made by the treating neurologist (JSH). We excluded cases in which the mechanism was not related to a blast explosion.

We were interested in the natural history of recovery in acute concussion. Previous reports indicate that individuals are most symptomatic in the first 48 hours after concussion25; therefore, we only included cases in which our team was able to conduct assessments on soldiers within 24 hours after the injury. As we were using a balance test as a key assessment of recovery, we excluded any cases in which there were secondary injuries to the lower extremities that would impair the ability to complete the balance assessment.

In our center’s evaluation and treatment algorithm, service members with acute concussion/mTBI were evaluated daily by staff. Assessments were made via clinical interview along with the individuals’ self-report on the GSC and a balance assessment with the BESS scored by occupational therapy staff (DL, EG).

The GSC lists 26 symptoms associated with concussion and respondents rate each symptom on a scale from 0 to 6 according to severity, with the sum across all symptom ratings yielding a total score. Reviews of concussion symptom checklist literature showed the GSC to have good sensitivity and specificity (0.94 and 1.00, respectively), whereas psychometric properties for most such concussion symptom checklists have not been published.26,27 Reliability measures for such concussion symptom scales are expected to be a function of administration time point following injury, as concussion symptoms will typically resolve over time, as observed in the NCAA Concussion Study.22

The BESS uses three different stances on two different surfaces; patients maintain a steady posture in each condition with their eyes closed their hands on their hips. The score is the sum of the patient deviations from the proper stance within the allotted time. A review of BESS literature showed a wide range for reliability across published studies with best reliability observed when BESS administration included training for raters and fewer raters.28 In the present study, BESS was administered by only two raters, each of whom had BESS administration experience by virtue of their assignment as occupational therapy staff. One member (DL) had supervisory responsibility for procedures and standardization. The BESS has been used as an outcome measure in studies of fatigue, dehydration, and neuromuscular exertion, so its validity for concussion may be compromised when administered in conditions that include such stressors.28 In the present study, the BESS was administered in a controlled clinical environment in which environmental stressors were minimized and, therefore, not expected to be an influence on BESS performance.

Patients were assessed within the first 24 hours after concussion (day 1), then on days 2, 3, 5, and 7 with the GSC and BESS. We performed a record review for the circumstances of the blast injury, recording LOC or AOC when present. We also recorded the type blast—examples include improvised explosive device (IED), suicide bomber explosion, and mortar explosion—and whether the soldier was in a vehicle (mounted) or outside a vehicle (dismounted). The data were analyzed and are presented descriptively.

**RESULTS**

Fifteen cases of blast-induced mTBI were identified. None of these 15 cases included other major musculoskeletal injury or positive neuroimaging results. The mean age was 28.5 years. There were 14 men and 1 woman described in the series. The most common blast scenario, seen in 11 of 15 cases, was a mounted IED blast that destroyed the vehicle in which the soldier was the driver or passenger. Of the remaining cases, 3 of 15 individuals were dismounted when a suicide bomber detonated a bomb in close proximity, and 1 case resulted from a mortar round that struck within 10 m of the soldier’s standing position. Table 1 shows demographic information and character of concussion as well as each soldier’s scores on the GSC and BESS measures.

Assessment of symptoms using the GSC showed a rapid recovery over a 7-day period of assessment (Fig. 1). Baseline preinjury scores were not available, so, to approximate a return to baseline,22 we examined day-to-day resolution of symptoms as a marker of recovery from concussion, with discontinuation of change in consecutive assessments indicating no further recovery and return to practical baseline.
There were differences over the 5 assessment days (repeated measures analysis of variance [ANOVA], $F_{13,4} = 12.06$, $p < 0.001$), and $t$ test pairwise comparisons showed day-to-day symptom differences only between assessments on days 1 and 2 [$t(13) = 3.5$, $p = 0.009$] and days 2 and 3 [$t(13) = 3.2$, $p = 0.006$]. There were no further symptom differences between days 3, 5, and 7, meeting our criterion for recovery. An alternate approach, Dunnett's test calculation, comparing each day to the final day (day 7) yielded the same result.

It should be noted that 4 cases were recovered and discharged from the program before the day 7 assessment, leading us to carry forward the day 5 score on the GSC and BESS to day 7. On average, symptoms in soldiers with concussion had resolved by day 7 (GSC mean change day 1 to day 7 was from 41 to 8, with only 4 of 15 showing any further resolution in symptomology from day 5 to day 7). Also, in the 15 cases, by day 7, 14 of 15 service members had resumed exercise, and only 1 of 15 (7%) had not returned to full duty by day 14.

On examination of postural stability using the BESS, patients showed a recovery pattern that reflected recovery in symptom reporting (Figs. 2 and 3). Across the five time points these two measures, GSC and BESS, showed positive pairwise correlation ($r = 0.73$, $p < 0.001$). For the differences in BESS measurements over the 5 assessment days (repeated measures ANOVA, $F_{14,4} = 43.21$, $p < 0.001$), $t$ test pairwise comparisons showed day-to-day postural stability differences between assessments on days 1 and 2 [$t(14) = 3.9$, $p = 0.002$], days 2 and 3 [$t(14) = 4.9$, $p < 0.001$], and days 3 and 5 [$t(14) = 3.8$, $p = 0.002$], with no further differences between days 5 and 7. An alternate approach, Dunnett’s test calculation, comparing each day to the final day (day 7) yielded the same result.

When examined as covariates, age, type exposure (mounted versus dismounted), and LOC did not reveal differences in GSC and BESS day 1 measures [all $t$ values $< 1.5$ and $p$ values $> 0.05$, regression (age) and ANOVA (LOC and type exposure)]. The sample size in this case series does not
afford sufficient statistical power for a thorough examination of differences from covariates, but the differences in means among these covariates resolved along with the recovery seen in the overall sample. There appeared to be no qualitative differences in the pattern of recovery.

DISCUSSION
In this retrospective case series, we provide the first description of the natural history of recovery in blast-induced mTBI. Our series suggests that the recovery in blast-induced mTBI follows the pattern of recovery in sports-induced concussion reported by McCrea et al.22 By using the same assessment tools used in the 2003 NCAA Concussion Study, we revealed similar and comparable patterns of recovery. As in sports concussion, there was near complete symptom and balance recovery by day 7 in blast-induced mTBI, with only 1 of 15(7%) not returned to full duty by day 14 (comparable to the 10% reported by McCrea et al, who needed more than 7 days to resolve symptoms from sports concussion). Our data and evaluation help us make the argument that while the mechanism of injury varies between blunt head injury and blast-induced mTBI, the profile of recovery is qualitatively similar. The similarity in recovery reported here in acute patients is consistent with Belanger et al,20 who showed that, in chronic patients, blast is not categorically different from other TBI mechanisms, at least with regard to cognitive sequelae on select measures during an intermediate to long-term assessment.

Our data also suggest that there is a clear neuropsychologic injury occurring acutely after blast concussion. The physics of blast neurotrauma have been well described in comparative research29 and current theories of injury that generalize to blast-induced mTBI. In the animal model,25,30 in our series, subjects reported marked symptoms immediately after blast-induced mTBI and there were objectively measureable balance deficits. Both symptoms and balance deficit quickly resolved within days following time of injury, with patient-reported symptoms resolving more rapidly than balance deficit. This pattern of recovery is consistently shown in the sports concussion literature31 and has also been seen in deployment concussion not specific to blast,16 and it is now reflected in blast-induced mTBI.

The long-term symptoms of blast-induced mTBI, however, are the subject of significant debate. Hoge et al and others have argued that long-term symptoms in soldiers exposed to blast or reporting a history of mTBI are more a function of psychological factors like anxiety, depression, and post-traumatic stress disorder.32 Similarly, a recent study of long-term postconcussion symptoms showed that long-term symptom reporting in service members deployed to OEF/OIF correlated mainly with stress levels.33 Moreover, as reviewed by McCrea et al,25,36 noninjury factors appear to play the dominant role in long-term outcomes after concussion/mTBI. We suspect that studies on individuals with symptoms after remote blast-induced mTBI (concussion/mTBI occurring greater than 6 months before evaluation) will continue to have null results when psychological factors are accounted for. Although there are limitations to inference from the clinical observations reported here such as the potential for some blunt force trauma to these soldiers’ helmets. Blast-induced neurotrauma without comorbid physical impact–induced injury has been described, but is not common.34 Regardless of the rarity of pure blast mTBI, the association of blast with neurotrauma includes considerations that warrant the study of blast injury as potentially separate from blunt injury without blast association.35,36 Further limitations to inference from the present study are omission of vision, hearing, or anosmia observations beyond the GSC in our case review, the nonavailability of psychological assessment results and baseline or normative values in a combat zone, and greater variability in time of day of combat injury and subsequent assessment than in sports concussion. Despite limitations, it is reasonable that our data suggest the time to study the true neurologic effects and dysfunction caused by blast-induced mTBI is in the first several days after concussion.

ACKNOWLEDGMENT
There was no separate funding source other than institutional resources available to the authors as part of their military positions.

REFERENCES
Natural History of Acute Recovery of Blast-Induced mTBI

Repeatead Low-Level Blast Exposure: A Descriptive Human Subjects Study

MAJ Walter Carr, MS USA*; James R. Stone, MD, PhD†; Tim Walliko, PhD‡; Lee Ann Young, MS§; Tianlu Li Snook, MD†; Michelle E. Paggi, MA¶; CAPT Jack W. Tsao, MC USN‖; CAPT Christopher J. Jankosky, MC USN (Ret.)**; LTC Robert V. Parish, MS USA††; CAPT Stephen T. Ahlers, MSC USN (Ret.)‡‡

ABSTRACT The relationship between repeated exposure to blast overpressure and neurological function was examined in the context of breacher training at the U.S. Marine Corps Weapons Training Battalion Dynamic Entry School. During this training, students are taught to apply explosive charges to achieve rapid ingress into secured buildings. For this study, both students and instructors participated in neurobehavioral testing, blood toxin screening, vestibular/auditory testing, and neuroimaging. Volunteers wore instrumentation during training to allow correlation of human response measurements and blast overpressure exposure. The key findings of this study were from high-memory demand tasks and were limited to the instructors. Specific tests showing blast-related mean differences were California Verbal Learning Test II, Automated Neuropsychological Assessment Metrics subtests (Match-to-Sample, Code Substitution Delayed), and Delayed Matching-to-Sample 10-second delay condition. Importantly, apparent deficits were paralleled with functional magnetic resonance imaging using the n-back task. The findings of this study are suggestive, but not conclusive, owing to small sample size and effect. The observed changes yield descriptive evidence for potential neurological alterations in the subset of individuals with occupational history of repetitive blast exposure. This is the first study to integrate subject instrumentation for measurement of individual blast pressure exposure, neurocognitive testing, and neuroimaging.

INTRODUCTION

The relation between primary blast exposure and traumatic brain injury (TBI) is not well understood and remains controversial. Approaches used to explore this connection include experimental animal models and human surrogates. Both are limited by the requirement for appropriate scaling to relate observations to human real-world blast scenarios and to the clinical setting. The generalizability of experimental observations is of particular importance when determining whether blast exposure can cause cognitive dysfunction in humans without evidence of injury on either clinical examination or conventional neuroimaging.

To investigate whether repeated exposures to low-level blast events are associated with alterations in neurocognitive function or neuroimaging in an exposed human population, a population of breachers was identified at the U.S. Marine Corps Weapons Training Battalion (WTB) Dynamic Entry School (DES). Breachers are a unique population who, as part of their regular training, are exposed to series of controlled blasts under supervised conditions that minimize the risk of injury from debris, fragments, or whole-body translation. Breachers participating in DES training are shown in Figure 1.

Breaching is the practice of using a variety of methods to gain entry to secured structures. Methods range from lock picking to controlled use of explosives. Explosive breaching was the focus of this investigation. Because blast exposure is controlled during breaching, the effects from a single exposure to blast overpressure do not result in an injury event that would prompt immediate concussion screening. Safety measures in place include use of safe standoff distances and hearing protection. However, anecdotal reports of neurological symptoms in experienced military and law enforcement breachers have emerged. Those anecdotal reports are echoed in a recent survey of symptoms among this population.¹

¹Symptoms reported in this population include headaches, sleep disturbances, and working memory impairment and are similar to those reported by service members with...
mild TBI who also report greater numbers of blast exposures. Among breachers, symptoms are more frequently reported by experienced personnel and are temporally correlated with periods of active field explosive training. Of note, such experienced personnel are those who become breaching Instructors and thus participate in multiple courses throughout the year and, by virtue of their job description, are routinely exposed to low-level blast. The tendency for symptoms to be reported by experienced breachers has raised the question as to whether cumulative injury may occur following repetitive exposure to low-level blast. To date, studies addressing this question of effects in humans from breaching blast have been equivocal, yielding no supporting evidence or evidence for isolated effects.

The scientific objective of this study was to determine whether measurable changes in neurological function, as assessed through behavioral and neuroimaging examinations, are associated with the repeated low-level blast exposures that occur as part of standard breacher training. At the outset, the working hypothesis that primary blast exposure can cause cumulative neurological injury included an expectation that such changes would be small. A measurable effect yielding conventional evidence of serious injury would have been recorded by training command personnel during routine operations and would have been prevented from recurrence through revision in procedures. However, a small effect or subclinical injury, developing slowly over repeated exposures and manifested to differing degrees amongst a group of individuals, might escape notice. As such, capture of a signal associated with a slowly developing, smaller effect might only be possible through targeted objective measurements.

METHODS

This study was approved by the Institutional Review Boards of the Naval Medical Research Center and the University of Virginia (UVA) in compliance with all applicable Federal regulations governing the protection of human subjects. The study was conducted with two separate 2-week breacher basic training courses at the WTB/DES in Quantico, Virginia.

During the breacher basic course, the first 2 weekdays are classroom-based, the next 6 weekdays are breaching practical, and the last 2 weekdays are for administrative tasks (Fig. 2). The intervening weekend days are liberty days, with no structured activities scheduled so as to leave time for the Students to review coursework. For the study reported here, the weekend days before and after the course, and a 30-minute period at the end of every training day, were dedicated to subject evaluations specific to the research protocol.

Participants underwent neurobehavioral testing, blood toxin screening, vestibular and auditory testing, and neuroimaging during the weekends before and after the 2-week training course. The first weekend provided a baseline condition for the test subjects, against which the post-exposure results could be compared. Computerized neurobehavioral testing and limited vestibular/auditory assessments were performed during the training period, thus facilitating the detection of acute neurological and auditory changes. To correlate human response measurements and the blast loading environment, participants were instrumented with pressure gauges.

Subjects

The subjects studied consisted of 40 research volunteers who were the U.S. Marines in one of three study groups: Students, Instructors, or Controls. Students \((n = 28)\) were those enrolled in the breacher basic course. Instructors \((n = 5)\) were those who taught DES breaching courses. Controls \((n = 7)\) were recruited from Marine Corps Base Quantico, but were not connected to the breacher basic course and not exposed to blast during the period of the study. Of the 28 Students, 26 completed the training course. The two Students who did not complete the course were not included in the final analysis. Another Student was excluded from the analyses because of medical issues unrelated to this protocol, yielding a total of 25 subjects for the Student group in the final analysis. Baseline evaluations of the Control subjects revealed that five of the seven were below the range presented by the 30 blast-exposed subjects and, thus, were not equivalent to the experimental groups; for this reason, none of the Control subject results are presented.

Among those included in the analyses, the mean age and education for Students and Instructors were similar (26 and 28 years old, respectively, and between 12 and 13 years of education for both groups). Ethnicity/race was predominantly Caucasian (27 of 30) with English as the self-reported native language for all. All subjects were right handed. Subjects provided histories indicating prior blast exposures ranging...
from 0 to 700 blasts. Only two of the subjects in the study were blast naive before entering the study.

Criteria for magnetic field safety were applied separately to the 30 subjects and are described in the Neuroimaging section, as those criteria affect only the neuroimaging data.

**Blast Characterization**

The predominant blast components responsible for the neurological sequelae of mild blast-induced TBI are unknown, but blast overpressure is largely accepted as an important traumatic mechanism given that up to 90% of the energies released on detonation of an uncased charge are converted into the formation of the shock wave. As such, breacher Students and Instructors were each instrumented with 4 pressure gauges (on their helmets above each ear and on each shoulder) and a 3 degree-of-freedom inertial cube (on the apex of their helmets). Blast exposures were further characterized using a free-field pressure gauge placed in the vicinity of each group of breachers to provide a reference data point. Sensor data were recorded at a rate of one million samples per second after passing through a 200 kHz antialiasing filter. The data were stored electronically for postprocessing.

**Neurocognitive/Neuropsychological Evaluation**

As a starting point, we conducted a review of the subjects’ medical records for information related to history of head injury or neurological impairment. In conjunction with this medical record review, all subjects were examined by a board-certified neurologist for signs of neurological abnormality, inclusive of mental status, cranial nerve function, motor strength, sensation, coordination, deep tendon reflexes, and gait. The neurocognitive assessments that followed consisted of technician-administered measures, self-report inventories, and computer-based cognitive performance measures. The assessments selected for this study (Table I) were based on a review of head injury literature and were intended to reveal and quantify changes in specific neurocognitive domains.

**Technician-Administered and Self-Report Neuropsychological Measures**

Technician-administered measures were conducted with each subject on an individual basis and completed in approximately 2 hours. The order of these measures was counterbalanced across subjects and held constant within subject from pre-exposure to post-exposure. Alternate forms of measures were used for pre-exposure and post-exposure when appropriate. On completion of the session, the test administrator provided the subject with self-report inventories and left the room. Subjects turned in the completed self-report inventories to a central location.

**Computer-Based Neuropsychological Measures**

Computer-based testing was administered 13 times to the subjects throughout the 2-week course, at the end of each training day. During each group testing session, subjects completed two different computerized batteries: the Automated Neuropsychological Assessment Metrics (ANAM) TBI Battery and a modified version of the Delayed Matching-to-Sample
TABLE I. Neuropsychological Test Library

<table>
<thead>
<tr>
<th>Technician Administered</th>
<th>Computer Based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trails A and B</td>
<td>ANAM TBI Battery</td>
</tr>
<tr>
<td>Controlled Oral Word</td>
<td>Mood Affect Score</td>
</tr>
<tr>
<td>Association Test</td>
<td>Stanford Sleepiness Scale</td>
</tr>
<tr>
<td>California Verbal Learning</td>
<td>Simple Reaction Time</td>
</tr>
<tr>
<td>Test II</td>
<td>Code Substitution</td>
</tr>
<tr>
<td>Weschler Adult Intelligence</td>
<td>Code Substitution Delayed</td>
</tr>
<tr>
<td>Scale-III Subtests</td>
<td>Procedural Reaction Time</td>
</tr>
<tr>
<td>Digit-Symbol Coding</td>
<td>Match-to-Sample</td>
</tr>
<tr>
<td>Symbol Search</td>
<td>Mathematical Processing</td>
</tr>
<tr>
<td>Digit Span</td>
<td>Delayed Matching-to-Sample</td>
</tr>
<tr>
<td>Letter-Number Sequencing</td>
<td></td>
</tr>
<tr>
<td>Wechsler Test of Adult Reading</td>
<td></td>
</tr>
<tr>
<td>Test of Memory Malingering</td>
<td></td>
</tr>
</tbody>
</table>

Self-Reported Symptoms
- Beck Depression Inventory-II
- State-Trait Anxiety Inventory
- Self-Report Blast Exposure
- Post-traumatic Stress Disorder
- Checklist-Military Version

(DMTS) test. These batteries were always completed in the order listed and the combined testing was completed in approximately 23 minutes. Each of these tests, ANAM and DMTS, is designed for repeated administration, varying test stimuli in a controlled manner between sessions.

The ANAM TBI Battery subtests assess different cognitive abilities (Table I) and are selected to be sensitive to effects of brain injury. ANAM subtests in general involve visually presented stimuli and computer mouse responses. Responses are recorded by the computer and scored for accuracy and response time. Instructions to subjects are to be both “fast and accurate.”

The DMTS assessment is a separate software application than the ANAM TBI Battery. The test paradigm is similar to the ANAM Match-to-Sample (M2S) subtest with two key differences: (1) DMTS stimuli are more complicated than M2S stimuli (8 × 8 matrix vs. 4 × 4 matrix, respectively) and (2) DMTS trials had a variable interstimulus delay (1 second or 10 seconds) instead of the static 5-second delay in M2S. The additional level of complexity in the DMTS served as an additional assessment of working memory, to mitigate potential ceiling effects in ANAM M2S. The variable interstimulus delay in DMTS means that there are 4 scores: accuracy and response time scores for the 1-second delay (easier condition) and accuracy and response time scores for the 10-second delay (harder condition). DMTS has demonstrated sensitivity to detect impairment of working memory in a field situation with operational personnel.

Subjects who failed the safety questionnaire were excluded from the neuroimaging portion of the study. Of the 25 Students and the 5 Instructors, 5 Students and 1 Instructor were excluded from MRI procedures because of the possibility of metal in the region of the head. One Student’s late arrival to the training course precluded his travel to UVA for MRI at pretest and he was excluded from the MRI data set. In sum, 19 Students and 4 Instructors were included in the MRI procedures and analyses.

MRI examinations included blood oxygen level-dependent (BOLD) sequences, acquired during the conduct of within-scanner tasks for functional magnetic resonance imaging (fMRI), diffusion weighted imaging across multiple vectors for diffusion tensor imaging, T1-weighted, T2-weighted, fluid attenuation inversion recovery, and susceptibility weighted imaging sequences. The objective of the current communication is to examine measureable changes in neurological function, and as such, this article provides descriptions of the fMRI data only. Structural neuroimaging will be reported separately.

For fMRI, n-back working memory and sentence completion language comprehension/word generation tasks were employed. For both tasks, stimulus presentation and response delivery were performed using the Eloquence fMRI system (Invivo, Orlando, Florida), consisting of a liquid-crystal display screen integrated into the head coil and a keypad for subjects’ responses. The n-back task was chosen based on anecdotal reports from breacher Instructors of working memory difficulties and on multiple reports in the literature describing the utility of this task in the characterization of mild TBI. The sentence completion task was chosen based on anecdotal reports from breacher Instructors of word finding difficulties and from members of the research team who reported observations of characteristic word finding deficits in patients with mild TBI associated with blast exposure. Each task consisted of five alternating 30-second baseline and activation epochs for a total of 5 minutes per run and runs were repeated three times. The n-back task involved serial presentation of letters and subject recognition of a target letter presented two letters previously (2-back condition). The sentence completion task involved observation of a coherent sentence with a blank and subjects were asked to covertly think of the appropriate word to substitute for the blank. All subjects received task training on a separate computer workstation just before entering the scanner.

Images were acquired with repeated single-shot echoplanar imaging: echo time = 30 ms, flip angle = 90°, matrix = 64 × 64, field of view 192 × 192 mm², slice thickness = 3.0 mm, repetition time = 3,000 ms, and 36 slices. A three-dimensional magnetization-prepared rapid acquisition with gradient echo T1-weighted isotropic whole brain data set was acquired for detailed anatomical correlation: repetition time = 1,900 ms, echo time = 1.89 ms, flip angle = 9°, matrix = 256 × 256, slice thickness = 1 mm, and 192 sagittal slices with no gap. All magnetic resonance images were collected using a 3T Siemens Trio scanner (Siemens AG,
Repeated Low-Level Blast Exposure

Erlangen, Germany) located within the MRI department of the UVA Health System (Charlottesville, Virginia).

Data analysis was performed using BrainVoyager 1.10 (Brain Innovation, Masstricht, The Netherlands) in a fashion similar to that described by Dricot et al.\textsuperscript{14} Thresholds of at least \( t > 2.87 \) (one-tailed, \( p \) [Bonf] \(<0.01) \) with \( q \) (false discovery rate) \(<0.05 \) were utilized. Differing thresholds were employed to adjust for anatomically meaningless clusters of activation. Regions of significance were defined based on clusters of activation observed within the statistical map generated from a two-factor repeated analysis of variance (ANOVA).

Other Evaluations

Other evaluations in this study included serum-based lead toxicity assessments and environment lead measurements, to address potential confounding effects associated with the use of lead-cased charges at the WTB/DES training site. Audiology and vestibular measures were assessed as well. Measured lead levels were well below the Occupational Safety and Health Administration permissible limit of 50 \( \mu \)g/m\(^3\). Audiology and vestibular findings from this study were communicated elsewhere.\textsuperscript{15}

Analysis

The planned between subjects comparison groups were Student and Instructor. The repeated measures design afforded two separate within subjects evaluations of change associated with exposure to blast. One evaluation compared assessments conducted on the weekend days before the 2-week training to the weekend days after the 2-week training (Fig. 2, Saturday/Sunday Pre-Exposure vs. Saturday/Sunday Post-Exposure). The other evaluation used the daily administration of computer-based neuropsychological testing to compare baseline performance to performance within hours following blast exposure, rather than performance on the final day of the protocol. As such, these computer-based neuropsychological testing data may reflect transient neurological alterations not present at the time of the technician-administered, self-report, and fMRI assessments. The pressure gauge recordings among individuals were used to reveal the time point with the greatest magnitude blast exposure and, thus, the time point most likely for acute effects among the volunteers. Because these measurements were made in context of a standard military training protocol (rather than exposing humans to experimentally scripted blasts, which would not be ethical), the time point to examine acute effects was not known a priori.

The sample size available for the Instructor group was quite small (\( n = 5 \)), limited by the inherent low density of the qualified Instructor population. Accordingly, analyses were considered descriptive in nature, useful for characterizing data from this type of study, but limited as support for reliable inference. Means were used to summarize the available data and repeated measures ANOVA was used for evaluations of blast-related change, comparing Student and Instructor groups on assessments conducted before and after exposure to blast. The analyses were planned comparisons and limited to descriptive purposes, so we did not use correction for multiple comparisons nor did we combine separate types of assessments into composite variables.

An accuracy criterion of \( \geq 57\% \) was applied to the ANAM TBI Battery as a subject exclusion criterion. This criterion is recommended by the software distributor and is the criterion applied in the Department of Defense use of the ANAM TBI Battery.

RESULTS

Breaching Environmental Characterization

Over the 2-week breacher training course, each subject was exposed to 40 low-level blasts. Over 4,400 pressure traces were collected, with an 80\% successful data capture in Phase 1 and a 98\% successful data capture in Phase 2. Sixty-one percent of the measured peak incident pressures were below 1 pound per square inch (psi); 4\% were above 4 psi. The highest peak incident pressure measured on all breachers was 13.0 psi. The minimum peak incident pressure was 0.1 psi.

The variance in cumulative impulse and peak pressure day-to-day was related to the size and number of charges detonated on a particular day. The average day-to-day variance is shown in Figure 3. The greatest exposure was on day 4 of the practicum exercises (Monday of the second week) when the two largest charges of the training program were detonated. The cumulative impulse on that day represents 25 to 30\% of the total exposure over the 2-week training course and was over two times larger than the exposure on any other day, verified by the total explosive material detonated on that day. If cognitive changes are related to the level of blast exposure, it would be expected that the greatest change in function should be detected on this day or shortly thereafter; thus, Day 4 of the practicum became

![FIGURE 3. Cumulative impulse energy received each day. The greatest exposure observed occurred on day 4. On any given day, the Instructors standing at the back of the stack typically had a lower exposure level than the Students.](image-url)
the basis for our comparison of the daily computer-based neuropsychological testing.

Daily and cumulative impulse was selected as the preferable measure of an individual’s overall exposure because it accommodates variables of peak pressure, duration, and pressure wave irregularities generated from being inside structures. Comparing exposures for the 2-week period using cumulative impulse, the average exposure for the Students was 0.0510 (standard deviation 0.014) psi-s compared to 0.0433 (standard deviation 0.002) psi-s for the Instructors. The standard deviation for the Student group is larger (27%) because the Students rotated among different positions in the breacher “stack,” whereas exposure between Instructors was more consistent (2%) because they consistently stood at the rear of the stack. On a day-to-day basis, the Instructors consistently had the smallest exposure overall (Fig. 3).

**Neurobehavioral Assessments**

**Baseline Neuropsychological Function**

At pretest, all subjects in the Student and Instructor groups were within normal ranges on neurological examination and Post-traumatic Stress Disorder Checklist-Military Version, Beck Depression Inventory-III, and State-Trait Anxiety Inventory (STAI) measures (Table II). Wechsler Test of Adult Reading-based estimates for intelligence quotient were above 90 for all subjects and group means were 105 and 108 for Students and Instructors, respectively. All subjects met Test of Memory Malingering criteria for sufficient effort. In the computer-based testing, for each ANAM subtest, all subjects showed response times for correct responses and percentage of correct responses that were equal to or above mean performance levels for a comparable cohort. There are no normative values available for the DMTS test used in this study.

**Technician-Administered Neuropsychological Measures**

The battery of technician-administered neuropsychological tests before and after the 2-week training period did not show change from pre-exposure to post-exposure or showed change consistent with practice effects expected in repeated administration of tests (Tables II and III). This pattern was observed for the sample as a whole and by group. An exception to this characterization of results is California Verbal Learning Test II (CVLT II) (Table III). The CVLT II recognition accuracy (hit rate) and d’ scale, a measure of sensitivity in cued recognition after long delay, showed a deficit for the Instructor group at post-exposure, following blast exposure.

**Self-Report Measures**

The changes in symptom reporting (on the 5-point scale History and Symptoms Questionnaire) between pre-exposure and post-exposure showed reporting differences among symptoms and among individuals. The symptom “Headaches” showed the greatest incidence of severity increase, with 8 of 25 (32%) Students reporting an increase and 3 of 5 (60%)...
### TABLE III.  
**Administrator-Led Neuropsychological Test Mean Scores Pre-exposure and Post-exposure**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Pre</th>
<th>Post</th>
<th>Δ</th>
<th>Δ±</th>
<th>Practice Effect</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive Function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COWA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>43.54</td>
<td>49.93</td>
<td>6.39</td>
<td>9.00</td>
<td>0.006</td>
<td>0.68</td>
</tr>
<tr>
<td>Student</td>
<td>45.50</td>
<td>49.13</td>
<td>3.63</td>
<td>0.00</td>
<td>0.415</td>
<td></td>
</tr>
<tr>
<td>Animals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>41.00</td>
<td>44.20</td>
<td>3.20</td>
<td>0.53</td>
<td>0.474</td>
<td>0.14</td>
</tr>
<tr>
<td>Student</td>
<td>44.48</td>
<td>45.52</td>
<td>1.04</td>
<td>0.14</td>
<td>0.714</td>
<td></td>
</tr>
<tr>
<td>Trails B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>54.00</td>
<td>58.40</td>
<td>4.40</td>
<td>3.01</td>
<td>0.094</td>
<td>0.02</td>
</tr>
<tr>
<td>Student</td>
<td>56.92</td>
<td>60.72</td>
<td>4.20</td>
<td>0.02</td>
<td>0.900</td>
<td></td>
</tr>
<tr>
<td>Attention/W. Memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digits Span</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>55.60</td>
<td>55.60</td>
<td>0.00</td>
<td>0.21</td>
<td>0.653</td>
<td>0.21</td>
</tr>
<tr>
<td>Student</td>
<td>51.88</td>
<td>53.48</td>
<td>1.60</td>
<td>0.21</td>
<td>0.653</td>
<td></td>
</tr>
<tr>
<td>Letter-Number Sequencing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>51.40</td>
<td>57.20</td>
<td>5.80</td>
<td>5.45</td>
<td>0.027</td>
<td>1.46</td>
</tr>
<tr>
<td>Student</td>
<td>55.80</td>
<td>57.64</td>
<td>1.84</td>
<td>1.46</td>
<td>0.237</td>
<td></td>
</tr>
<tr>
<td>Speed of Processing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Digit Symbol Coding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>53.80</td>
<td>62.40</td>
<td>8.60</td>
<td>13.30</td>
<td>0.001</td>
<td>0.86</td>
</tr>
<tr>
<td>Student</td>
<td>55.96</td>
<td>61.08</td>
<td>5.12</td>
<td>0.86</td>
<td>0.363</td>
<td></td>
</tr>
<tr>
<td>Symbol Search</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>54.40</td>
<td>61.20</td>
<td>6.80</td>
<td>14.72</td>
<td>&lt;0.001</td>
<td>0.05</td>
</tr>
<tr>
<td>Student</td>
<td>58.68</td>
<td>64.76</td>
<td>6.08</td>
<td>0.05</td>
<td>0.832</td>
<td></td>
</tr>
<tr>
<td>Trails A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>56.40</td>
<td>60.80</td>
<td>4.40</td>
<td>5.08</td>
<td>0.032</td>
<td>0.19</td>
</tr>
<tr>
<td>Student</td>
<td>55.32</td>
<td>61.84</td>
<td>6.52</td>
<td>0.19</td>
<td>0.665</td>
<td></td>
</tr>
<tr>
<td>Verbal Episodic Memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CVLT Trial 1 Free Recall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>47.00</td>
<td>49.00</td>
<td>2.00</td>
<td>0.22</td>
<td>0.647</td>
<td>1.38</td>
</tr>
<tr>
<td>Student</td>
<td>52.20</td>
<td>47.60</td>
<td>4.60</td>
<td>1.38</td>
<td>0.250</td>
<td></td>
</tr>
<tr>
<td>CVLT Trial B Free Recall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>44.00</td>
<td>44.00</td>
<td>0.00</td>
<td>0.55</td>
<td>0.464</td>
<td>0.55</td>
</tr>
<tr>
<td>Student</td>
<td>52.60</td>
<td>48.40</td>
<td>4.20</td>
<td>0.46</td>
<td>0.646</td>
<td></td>
</tr>
<tr>
<td>CVLT Trials 1–5 Free Recall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>54.60</td>
<td>58.00</td>
<td>3.40</td>
<td>0.31</td>
<td>0.579</td>
<td>1.06</td>
</tr>
<tr>
<td>Student</td>
<td>56.84</td>
<td>55.84</td>
<td>1.00</td>
<td>1.06</td>
<td>0.313</td>
<td></td>
</tr>
<tr>
<td>CVLT Short Delay Free Recall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>54.00</td>
<td>56.00</td>
<td>2.00</td>
<td>0.31</td>
<td>0.582</td>
<td>0.04</td>
</tr>
<tr>
<td>Student</td>
<td>55.20</td>
<td>56.20</td>
<td>1.00</td>
<td>0.04</td>
<td>0.854</td>
<td></td>
</tr>
<tr>
<td>CVLT Short Delay Cued Recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>56.00</td>
<td>55.00</td>
<td>1.00</td>
<td>0.34</td>
<td>0.563</td>
<td>1.16</td>
</tr>
<tr>
<td>Student</td>
<td>55.40</td>
<td>58.80</td>
<td>3.40</td>
<td>1.16</td>
<td>0.292</td>
<td></td>
</tr>
<tr>
<td>CVLT long Delay Free Recall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>47.00</td>
<td>55.00</td>
<td>8.00</td>
<td>4.94</td>
<td>0.035</td>
<td>3.66</td>
</tr>
<tr>
<td>Student</td>
<td>55.20</td>
<td>55.80</td>
<td>0.60</td>
<td>3.66</td>
<td>0.066</td>
<td></td>
</tr>
<tr>
<td>CVLT Long Delay Cued Recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>52.00</td>
<td>52.00</td>
<td>0.00</td>
<td>&lt;0.01</td>
<td>0.965</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Student</td>
<td>55.00</td>
<td>55.20</td>
<td>0.20</td>
<td>&lt;0.01</td>
<td>0.965</td>
<td></td>
</tr>
<tr>
<td>CVLT Hit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>50.00</td>
<td>42.00</td>
<td>8.00</td>
<td>1.84</td>
<td>0.186</td>
<td>5.69</td>
</tr>
<tr>
<td>Student</td>
<td>49.40</td>
<td>51.60</td>
<td>2.20</td>
<td>5.69</td>
<td>0.024</td>
<td></td>
</tr>
<tr>
<td>CVLT False Positive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>50.00</td>
<td>53.00</td>
<td>3.00</td>
<td>2.36</td>
<td>0.136</td>
<td>0.31</td>
</tr>
<tr>
<td>Student</td>
<td>46.80</td>
<td>48.20</td>
<td>1.40</td>
<td>0.31</td>
<td>0.581</td>
<td></td>
</tr>
<tr>
<td>CVLT d'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instructor</td>
<td>54.00</td>
<td>47.00</td>
<td>7.00</td>
<td>3.47</td>
<td>0.073</td>
<td>7.80</td>
</tr>
<tr>
<td>Student</td>
<td>54.60</td>
<td>56.00</td>
<td>1.40</td>
<td>7.80</td>
<td>0.009</td>
<td></td>
</tr>
</tbody>
</table>

Administrator-led neuropsychological test mean T scores at pre-exposure and at post-exposure presented by group (data columns 1 and 2, respectively; Instructor n = 5, Student n = 25). Difference between each pre-exposure–post-exposure pair of these mean scores is presented in data column 3. Column 4, “Δ±,” indicates if the change value (column 3) represents an improvement (+) or a decline (−) in performance. Results of repeated measures ANOVA are presented in data columns 5 and 6 for repeated testing effect (practice effect) and in data columns 7 and 8 for practice effect interaction between subject groups.
Instructors reporting increase over the 2-week period. Of the remaining symptoms, no symptom was reported by more than 20% of the sample. In general, across all 22 queried symptoms, 11 of 25 (44%) Students reported increase following blast exposure as compared to 4 of 5 (80%) Instructors.

Computer-Based Neuropsychological Measures
As described in the Analysis section, a key difference between the computer-based assessments and the technician-administered neuropsychological assessments was in frequency of administration and consequent availability of data in close temporal proximity to blast exposure (<3 hours) for computer-based assessments. Computer-based performance results are reported here as three time points, characterizing (1) baseline, (2) following exposure to largest blast (day 4 of explosives practicum; see Fig. 3), and (3) 2 days after final exposure to blast. Baseline was defined as mean performance across Sessions 3 and 4 (the Monday and Tuesday of the first week of the data collection), which is before blast exposure for the course, but after Sessions 1 and 2, allowing for known practice effects.18 Similarly, the comparison time points also represent scores from two consecutive sessions, averaged to reduce effects of day-to-day variance in performance and in consideration of the relatively small sample size. Figures 4 to 6 are group means for each of these three time points.

The 57% accuracy criterion affected 3 of the 7 ANAM subtests (Table I). Specifically, in M2S, 1 of the 30 subjects was excluded (1 Student) and in Mathematical Processing (MTH) and Code Substitution Delayed (CDD), 3 of 30 subjects were excluded (2 Students). The 1 Student removed from ANAM M2S was also 1 of the 2 Students removed from the ANAM MTH and CDD subtests. The DMTS is a much more difficult test than the ANAM subtests. In previous use of versions of DMTS with samples of military populations, the prediction was that 20 to 25% of the population find the test too difficult.10 On the basis of DMTS developer a priori experienced-based predictions, we included all 30 subjects rather than use a 24% subject exclusion rate.

Results consistent with practice effect are reported elsewhere.18 Primary results reported here are those that show group difference, Instructor vs. Student. Statistical power in these analyses was low, as a function of small sample size and also of the hypothesized small effect size. Mean differences reported here illustrate a pattern of responses for the Student and Instructor groups across the three time points, specifically, a pattern of Instructors’ decreased accuracy observed among tasks with a memory demand and not observed among tasks without a memory demand.

As with the battery of technician-administered neuropsychological tests at pre-exposure and post-exposure, ANAM subtests Code Substitution (CDS), procedural reaction time (PRO), and MTH showed general ceiling effects for accuracy (<5% error) and a repeated administration practice effect in response time for correct responses. Simple Reaction Time (only a single choice response and thus no errors or accuracy for the course, but after Sessions 1 and 2, allowing for known practice effects.18 Similarly, the comparison time points also represent scores from two consecutive sessions, averaged to reduce effects of day-to-day variance in performance and in consideration of the relatively small sample size. Figures 4 to 6 are group means for each of these three time points.

The 57% accuracy criterion affected 3 of the 7 ANAM subtests (Table I). Specifically, in M2S, 1 of the 30 subjects was excluded (1 Student) and in Mathematical Processing (MTH) and Code Substitution Delayed (CDD), 3 of 30 subjects were excluded (2 Students). The 1 Student removed from ANAM M2S was also 1 of the 2 Students removed from the ANAM MTH and CDD subtests. The DMTS is a much more difficult test than the ANAM subtests. In previous use of versions of DMTS with samples of military populations, the prediction was that 20 to 25% of the population find the test too difficult.10 On the basis of DMTS developer a priori experienced-based predictions, we included all 30 subjects rather than use a 24% subject exclusion rate.

Results consistent with practice effect are reported elsewhere.18 Primary results reported here are those that show group difference, Instructor vs. Student. Statistical power in these analyses was low, as a function of small sample size and also of the hypothesized small effect size. Mean differences reported here illustrate a pattern of responses for the Student and Instructor groups across the three time points, specifically, a pattern of Instructors’ decreased accuracy observed among tasks with a memory demand and not observed among tasks without a memory demand.

As with the battery of technician-administered neuropsychological tests at pre-exposure and post-exposure, ANAM subtests Code Substitution (CDS), procedural reaction time (PRO), and MTH showed general ceiling effects for accuracy (<5% error) and a repeated administration practice effect in response time for correct responses. Simple Reaction Time (only a single choice response and thus no errors or accuracy...
score) response times did not show practice effect or reliable deficit across the three time points.

ANAM subtests M2S and CDD, on the other hand, showed a relatively greater error rate (>5%). For these 2 subtests, there was a pattern of mean differences in the Percent Correct variable suggesting an effect of exposure (Figs. 4 and 5). This pattern, however, did not meet the criterion for statistical significance in either M2S or CDD in this descriptive analysis.

DMTS is similar to the ANAM M2S subtest but is more difficult and consequently yields much greater error rates (33%) and much longer response times. Response times were mostly beyond 4 seconds, especially in the more difficult 10-second delay condition. Interpreting response times of this magnitude...
becomes unclear, as the task becomes more contemplative. In this consideration, Percent Correct is the DMTS variable presented in Figure 6. DMTS 10-second delay, as with ANAM M2S and CDD subtests, showed mean differences for the Instructor group in Percent Correct that were inconsistent with the more accurate responses of a practice effect. Like M2S and CDD, this pattern did not meet criterion for statistical significance when multiple comparisons are considered.

**Neuroimaging**

For n-back and sentence completion tasks, the general linear model (GLM) was calculated from BOLD sequences and compared to the expected hemodynamic response function for the stimulus paradigm. Analysis of covariance was performed to assess group differences. Activated clusters across all subjects during baseline and activation phases of the sentence completion task are shown in Figure 7 (random effects GLM: $t_{(54)} \geq 3.68$ or $\leq -3.68$, $p \leq 5.25 \times 10^{-4}$). Areas of activation are seen primarily within the left frontoparietal and bilateral occipital regions. Areas of decreased activation are seen along the right lateral convexity, the paramedian frontal and parietal lobes, and within bilateral temporal lobes. Activated clusters across all subjects for the n-back task are shown in Figure 8 (random effects GLM: $t_{(54)} \geq 3.65$ or $\leq -3.65$, $p \leq 6.0 \times 10^{-4}$). Areas of activation are seen primarily within frontoparietal regions along the convexities with some activation in occipital lobes. In addition, areas of activation are seen with respect to the basal ganglia. Areas of decreased activation are seen within frontal, parietal, and occipital lobes near the interhemispheric fissure. Decreased activity is also seen within bilateral temporal lobes.

Repeated measures ANOVA comparisons were performed to determine whether cluster differences existed between Student and Instructor post-exposure examinations. No significant group differences were seen with the fMRI sentence completion task on either baseline to post-exposure comparisons by group or when comparing between groups. With the n-back task, comparison of baseline and post-exposure evaluations in Students demonstrated no significantly different activated clusters (random effects GLM: $t_{(37)} \geq 3.00$ or $\leq -3.00$, $p \leq 0.005$). However, a single significantly different cluster was seen in the baseline to post-exposure Instructor comparison (random effects GLM: $t_{(7)} \geq 4.00$ or $\leq -4.00$, $p \leq 0.005$). This cluster demonstrated increased activity in the post-exposure studies as compared to the baseline evaluations. As shown in frame A of Figure 9, this cluster localized to Brodmann area 47, which has been shown to be responsible for processing of syntax in spoken and signed languages. Also, comparison of Instructors to Students (random effects GLM: $t_{(22)} \geq 4.26$ or $\leq -4.26$, $p \leq 3.5 \times 10^{-4}$) demonstrated increased activity within the Instructor group. The Instructor to Student group demonstrated four volumes of interest.
with increased activity. Two volumes of interest were located within Brodmann area 3 of the anterolateral postcentral gyrus, one within Brodmann area 7 of the anterior parietal cortex and the other in Brodmann area 40 of the anterior supramarginal gyrus (Fig. 9).

**DISCUSSION**

The measures employed before and after the 2-week breacher basic course and during the 6 days of exposure to repeated blast did not yield clear evidence for neurological impairment in breacher Students or Instructors. The results observed, however, did yield a description suggestive of blast-induced impairment in selected domains of cognition among individuals subject to sustained repetitive blast exposure. This description includes a reduction in accuracy of Instructor responses on tasks that place demand on memory ability, specifically, the 2 ANAM subtests M2S and CDD, the DMTS 10-second delay condition, and recognition for CVLT II auditory stimuli. Taken together, these results can be said to describe a pattern. It is the similarity across these tests that prompts our consideration. The fMRI BOLD signal increase during the n-back task for Instructors in their post-exposure to environmental insult, it is hypothesized that this increased signal represents less-efficient metabolic function within the traumatically affected brain. While the findings of the current study are intriguing, they must be validated with a larger sample size in a longitudinal study.

The data collected in this study were sufficient for descriptive purposes. It may be that in future studies of the hypothesized relation between blast exposure and subclinical neurological insult in humans, an effective approach may depend on deep evaluation at the single subject level. The addition of breacher instrumentation in the current study provided critical data on subject load exposure, which has not been available in previous clinical and epidemiological work. When comparing the group neurocognitive and neuroimaging results in context of the recorded exposure data, the findings suggest that an individual’s sensitivity to blast may be enhanced after multiple exposures. The period of time over which these exposures must occur for the sensitivity to manifest appears to be longer than the 2-week period of this study. This is supported by the fact that the findings in the neurocognitive tests and fMRI were among the Instructors, who, on average, had the lowest blast exposure levels over the 2-week period but who would have had a history of...
chronic blast exposure (at minimum, by virtue of prerequisites to become an Instructor and experience as an Instructor). The Instructors had more self-reported symptoms and greater mean difference changes in the targeted objective measures than the Student group. Based on the trends observed during this study, long-term chronic exposure to blast environments may increase the likelihood of developing neurological signs or symptoms consistent with concussion injury during subsequent blast re-exposure.

ACKNOWLEDGMENTS

This research could not have been completed without the collaboration of the Breacher Injury Study Consortium including Natalie Carver, Gary Ogg, Bob Lynch, Christopher P. Holstege, MD, Carol Chancey, PhD, LTC Lynnette Bardolf, MS USA, Melinda Hill, PhD, Fred Brozoski, PhD, Jeannine B. Mielke, PhD, Alicia L. Janos, Victoria B. Anderson-Barnes, LTC Margaret Swanberg, MC USA, and CW5 Brad Garfield, USMC (Ret.). This work was funded by work unit number 603729N.2914.W300.A0512.

REFERENCES

Simple and Procedural Reaction Time for Mild Traumatic Brain Injury in a Hyperbaric Oxygen Clinical Trial

Susan Churchill, APRN-NP†; COL R. Scott Miller, MC USA (Ret.)‡; Kayla Deru, BA†; Steffanie H. Wilson, PhD§; Lindell K. Weaver, MD†∥

ABSTRACT Simple reaction time (SRT) and procedural reaction time (PRT) are speed-of-processing tasks in the Automated Neuropsychological Assessment Metrics (ANAM) that may be sensitive to mild traumatic brain injury (mTBI). The investigators measured SRT and PRT throughput (correct responses per minute) at baseline, 6 weeks, and 13 weeks in military personnel with mTBI randomized to local care or 40 chamber sessions (sham-1.2 atmospheres absolute [ATA] air, hyperbaric oxygen-1.5 ATA O₂). Scores were assessed at baseline using univariate analysis of variance and across time with repeated measures methods. Data reported as throughput standard scores (mean = 100, SD = 15). Seventy-two participants with ongoing symptoms after mTBI enrolled in the study (three female, median age 31 years, mean three lifetime concussion events, most recent mTBI 23 months prior). Sixty-four had Automated Neuropsychological Assessment Metrics data at 13 weeks. SRT and PRT throughput standard scores were comparable across groups at baseline. Over time, SRT scores did not change in the hyperbaric oxygen or sham groups and decreased in the local care group. PRT throughput standard scores increased from baseline to mid-intervention and decreased from mid-intervention to postintervention in all groups. Repeated measures change over time in SRT (p = 0.23), and PRT (p = 0.17) scores were not different among groups. This study may be underpowered to detect statistically significant change.

INTRODUCTION The Automated Neuropsychological Assessment Metrics (ANAM) software (Vista LifeSciences, Parker, Colorado) is a computerized battery of cognitive tests that has been administered to U.S. service members before and after deployment. Two speed-of-processing tasks in the ANAM (version 4.0) test battery, the simple reaction time (SRT) and procedural reaction time (PRT), may be sensitive to cognitive deficits after mild traumatic brain injury (mTBI).1–5 though the effect size in traumatic brain injury is modest when the injury is remote.6

In this article, the investigators will report using the SRT and PRT subtests of the ANAM4 as outcome measures. The ANAM4 was administered as a secondary outcome measure in a randomized, double-blind, sham-controlled clinical trial: “A Pilot Phase II Study of Hyperbaric Oxygen for Persistent Post-Concussive Symptoms After Mild Traumatic Brain Injury (HOPPS).”7,8 The utility of the ANAM4 in screening for mTBI may be limited beyond the acute phase,9 though this conclusion may be limited by small sample size.10 This tool may also lack the reliability and precision needed to direct clinical care after mTBI.11 Nevertheless, as a measure of cognitive performance across time, it may have potential as an outcome assessment in clinical trials, but only if sensitive to subtle cognitive deficits that may be obscured by traditional neuropsychological testing. Other potential attributes of the ANAM4 include its wide availability within the Department of Defense and the years of experience with this software, as well as the potential availability of the preinjury ANAM4 scores for clinical trials in service members with brain injury.

METHODS Following institutional review board approval, active duty military personnel with a history of mTBI (occurring at least 4 months before enrollment) and ongoing symptoms were recruited at four military sites: Evans Army Community Hospital (Fort Carson, Colorado), Eisenhower Army Medical Center (Fort Gordon, Georgia), Naval Hospital Camp Pendleton (North Carolina), and Naval Hospital Camp Lejeune (North Carolina). Individuals with a history of moderate or severe traumatic brain injury or current illicit drug use were excluded, as were those for whom hyperbaric oxygen (HBO₂) might carry significant risk.

Study participants were randomized to receive local care (no chamber sessions) or 40 hyperbaric chamber sessions over a 10-week period breathing either air at 1.2 atmospheres absolute (ATA) (sham control) or HBO₂ (100% oxygen at 1.5 ATA) (active intervention) for 60 minutes. Participants randomized to chamber sessions and the research teams were blind to allocation.7,8
At baseline, 6 weeks (mid-intervention), and 13 weeks, participants were assessed by standardized symptom questionnaires, traditional neuropsychological tests, and ANAM4. The primary results of the study (clinicaltrials.gov identifier: NCT01306968) have been presented elsewhere. For this brief report, the investigators examined population characteristics and changes over time in processing speed using the ANAM4 SRT and PRT subtests (Fig. 1).

The ANAM4 test battery was conducted on identical, dedicated laptops provided by the Neurocognitive Assessment Branch within the Army, who installed and configured the ANAM4 software. The site research coordinators acted as proctors for ANAM4 testing after standardized training and delegation of authority. Per protocol analysis was performed for participants with complete ANAM4 SRT and PRT scores across time, using throughput (number of correct responses per minute) standard scores (mean =100, SD = 15) and percentile ranking. The ANAM4 scoring defines “clearly impaired” as a standard score of less than 70 (at least two SDs below the mean).

Descriptive statistics are presented for each time point, and SRT and PRT scores were assessed for comparability among intervention groups at baseline using univariate analysis of variance (ANOVA) and across time with repeated measures ANOVA using PROC MIXED in SAS Version 9.3.

RESULTS
Seventy-two participants enrolled in the study (three female, median age 31 years, mean three lifetime concussion events [range 1–11], with the most recent mTBI occurring on average 23 months before randomization). Fifty-six (78%) participants had injures with loss of consciousness. The most severe duration of loss of consciousness was less than 5 minutes in 37 participants (51%), and between 5 and 30 minutes for 17 participants (24%). Seventeen participants (24%) had initiated medical disability evaluation boards (a panel of at least two physicians who determine medical fitness for retention in the military). Sixty-six participants (92%) scored at least 45 (the cutoff for normal, healthy individuals) on the Test of Memory Malingering at baseline. Full baseline characteristics of the enrolled population, including education, deployments, service branch, medication usage, and post-traumatic stress disorder can be found in the primary publication.

The per protocol population considered in this analysis includes 64 participants with complete ANAM4 data at 13-week follow-up (local care group \(N = 20\), HBO2 group \(N = 23\), sham group \(N = 21\)). In this population, again, three were female, the mean age was 33 years, participants had a mean of four lifetime concussion events, and the most recent mTBI occurred an average of 2 years before enrollment.

Baseline SRT throughput standard scores (Table I) were similar in the sham (mean = 85, 95% confidence interval [CI] = 71–100) and local care groups (mean = 85, 95% CI = 72–98) and lower in the HBO2 group (mean = 79, 95% CI = 64–94), but differences among the groups at baseline were not statistically significant (\(p = 0.78\)). Additionally, baseline mean SRT percentile throughput ranks were higher in the sham group (mean = 39, 95% CI = 24–54) than in the HBO2 (mean = 23, 95% CI = 9–36) and local care (mean = 33, 95% CI = 18–49) groups. As seen in the profile plot (Fig. 2), mean SRT throughput standard scores remained fairly constant across time in the HBO2 and sham groups and decreased at both 6 weeks and 13 weeks in the local care group.

Repeated measures analysis indicated slight increases in estimated differences in SRT standard scores between baseline and 13 weeks for the HBO2 (mean difference = 1.7, SE = 5.5) and sham groups (mean difference = 1.7, SE = 5.7) and a decrease in estimated difference for the local care group (mean difference = –15.0, SE = 5.8). However, the time by intervention group interaction in the repeated measures

![FIGURE 1. Representative test stimuli for Automated Neuropsychological Assessment Metrics (ANAM4) Simple Reaction Time (SRT) and Procedural Reaction Time (PRT) subtests. (A) In the SRT, a series of * (star) symbols is presented on the display. The user pushes a computer mouse button as quickly as possible each time the * appears. (B) In the PRT, a dot matrix number is presented on the display. The user presses the right or left computer mouse buttons to indicate whether the number is “low” (2 or 3) or “high” (4 or 5).](image-url)
model did not indicate that changes over time in SRT standard scores were significantly different among intervention groups ($p = 0.23$).

Baseline PRT throughput standard scores (Table I) were greater in the sham group (mean = 87, 95% CI = 76–99) than the HBO2 (mean = 81, 95% CI = 68–94) or local care groups (mean = 86, 95% CI = 74–97), but similar to results of SRT analysis, the differences in PRT scores among the groups at baseline were not significant ($p = 0.71$). Mean PRT percentile throughput ranks at baseline were higher in the sham group (mean = 35, 95% CI = 19–51) than in the HBO2 (mean = 31, 95% CI = 15–47) or local care (mean = 29, 95% CI = 15–44) groups. PRT throughput standard mean scores increased in all intervention groups between baseline and 6 weeks and then decreased between 6 weeks and after the intervention at 13 weeks (Fig. 3). Overall, PRT scores decreased from baseline to 13 weeks in the local and sham groups and increased in the HBO2 group.

Repeated measures ANOVA showed decreased estimated differences between baseline and 13 weeks for sham (mean difference = −8.7, SE = 5.2) and local care (mean difference = −4.5, SE = 5.3) and an increased difference in the HBO2 group (mean difference = 7.2, SE = 5.0). However, similar to the results of SRT analysis, the time by intervention group interaction term in the repeated measures model did not show evidence of significant differences in changes over time in PRT scores ($p = 0.17$).

In both SRT and PRT, more participants in the local care group met “clearly impaired” criteria (a standard score of less than 70, at least two SDs below the mean) at 13 weeks than at baseline. This change was not observed in the HBO2 and sham groups, where performance was variable (Figs. 4 and 5).

**DISCUSSION**

The diagnostic utility of the ANAM for mTBI and cognitive impairment and its role in clinical decision making have not yet been well-established nor universally embraced.\(^9\)\(^–\)\(^11\)

The ANAM user manual itself cautions that test results for any given patient should be “considered within the broader framework of information regarding the test taker” (ANAM).

### TABLE I.

<table>
<thead>
<tr>
<th>Test</th>
<th>Interval</th>
<th>Local Care (No Chamber)</th>
<th>HBO2 (1.5 ATA)</th>
<th>Sham (1.2 ATA air)</th>
<th>Total</th>
</tr>
</thead>
</table>

\(^{9}\)\(^–\)\(^11\) MILITARY MEDICINE, Vol. 181, May Supplement 2016
TBI-MIL Battery: User Manual. 2 ed. 2008). However, computer-based tests have some advantages over traditional assessments of cognitive function, such as paper-and-pencil neuropsychological tests, in that they accurately capture response time in milliseconds and can present stimuli in a randomized order, allowing the tests to be repeated frequently. These features make the ANAM4 test battery attractive as an outcome measure in an interventional trial, and it has been used occasionally for this purpose.13,14

The group mean scores for these ANAM4 subtests are consistent with the primary findings of the study that receiving HBO2 did not result in greater symptom relief or improved function compared to receiving sham chamber sessions for service members with mTBI.7 It is possible, however, that this study was underpowered to detect statistically significant change with these tests, and it is important to note that because the nature of the analysis was exploratory, no adjustments were made for multiple testing.

Other weaknesses of this study include the heterogeneity of the population (with respect to age, service branch, injury etiology, and time from injury) and concerns with participant performance. Individuals with mTBI may be likely to have more variability in ANAM4 performance across time than normal controls,5 and this is one possible explanation for the variability in the ANAM4 scores observed in this study. As with all neuropsychological testing, suboptimal effort can impact test results of computerized assessment. The deterioration in SRT and PRT throughput of many of the participants randomized to receive local care may reflect suboptimal effort, since these individuals were unblinded and were not receiving chamber sessions, or it may reflect a true weakness in the available mTBI care to stabilize or improve reaction time. The ANAM4 itself has no internal test of validity or optimal effort.

Test performance may also be influenced by internal and external confounders, such as poor sleep quality,15 headache,16 and dehydration,17 though in a randomized trial design, one would expect these confounds to be nearly equally represented in all study groups, and in this study, baseline SRT and PRT scores were comparable across all study groups. Other investigators have addressed the natural variability of computerized neuropsychological testing with reliable change calculations.12,18 The investigators elected not to perform these calculations because trends seen in SRT and PRT scores were consistent with those of the other outcome measures of the clinical trial, particularly symptom reporting, i.e., no benefit of HBO2 over sham chamber sessions and because these calculations were unlikely to yield information of added value.

Other studies have reported a practice effect on these subtests when administered multiple times over a short time period (daily or multiple times per day).19 Because the participants in this study had previously taken the ANAM4 and were tested at least 6 weeks apart, and because across-the-board improvement in scores was not observed, it is unlikely that practice effect influenced the study results.

In this study, baseline throughput percentile ranks for SRT and PRT were below 50 in all groups, indicating underlying dysfunction in the group, a trend seen also on standard scores. For both SRT and PRT, 30% of participants met the “clearly impaired” criteria at baseline. These findings support that individuals with postconcussive symptoms are more likely to score below average on these ANAM4 subtests than normal controls. This assertion could be strengthened by comparison to predeployment ANAM4 results, required for all service members since 2008.20 Unfortunately, the predeployment ANAM4 test results for this study’s participants were not available to the investigators. However, the ANAM4 results may be useful for identifying cognitive impairment even when comparing to normative data alone.21

---

FIGURE 4. Number and percentage of participants meeting “clearly impaired” criteria on the Simple Reaction Time test (throughput) at baseline and at 13 weeks, defined as scoring at least 2 standard deviations below the mean (mean = 100, standard deviation = 15).

FIGURE 5. Number and percentage of participants meeting “clearly impaired” criteria on the Procedural Reaction Time test (throughput) at baseline and at 13 weeks, defined as scoring at least 2 standard deviations below the mean (mean = 100, standard deviation = 15).
All HOPPS participants met the inclusion/exclusion criteria for enrollment into the clinical trial, which included a history of mTBI and ongoing postconcussive symptoms. Although intended as an outcome measure, the ANAM results support that the study participants had abnormalities consistent with postconcussive symptoms. In this study, the SRT and PRT results support that the HOPPS participants had cognitive impairment from their mTBI that may reflect the injury and disability the HOPPS project intended to study. The results are strengthened by the prospective nature of this clinical trial and the rigorous oversight from regulatory bodies, as the study was conducted under an Investigational New Drug Application with the U.S. Food and Drug Administration.

With the upcoming completion of another study in the program, additional data may become available to help guide the application of the ANAM in clinical research.

In conclusion, there was no evidence that the observed SRT and PRT changes across time differed among the HBO2 sham, and local care groups, which is consistent with the larger findings of this clinical trial. Although perhaps too sensitive to external factors to be considered a primary outcome measure in interventional trials, the ANAM SRT and PRT subtests may be appropriate secondary outcome measures in select interventional studies, particularly in patient populations where deficits on these subtests are expected. Additional validation studies with a larger population sample would be required for this outcome measure to become well-accepted in clinical trials.

ACKNOWLEDGMENTS

We acknowledge the Naval Health Research Center, as well as the Army Contracting Command for contracting support throughout this trial. This study was funded by the Defense Health Program and managed by the U.S. Army Medical Materiel Development Activity.

REFERENCES

Neurosensory Assessments of Concussion

Mark Tommerdahl, PhD*†; Robert G. Dennis, PhD*†; Eric M. Francisco, PhD†;
Jameson K. Holden, PhD†; Richard Nguyen, PhD†; Oleg V. Favorov, PhD*

ABSTRACT The purpose of this research was to determine if cortical metrics—a unique set of sensory-based assessment tools—could be used to characterize and differentiate concussed individuals from nonconcussed individuals. Cortical metrics take advantage of the somatotopic relationship between skin and cortex, and the protocols are designed to evoke interactions between adjacent cortical regions to investigate fundamental mechanisms that mediate cortical-cortical interactions. Student athletes, aged 18 to 22 years, were recruited into the study through an athletic training center that made determinations of postconcussion return-to-play status. Sensory-based performance tasks utilizing vibrotactile stimuli applied to tips of the index and middle fingers were administered to test an individual’s amplitude discrimination, temporal order judgment, and duration discrimination capacity in the presence and absence of illusion-inducing conditioning stimuli. Comparison of the performances in the presence and absence of conditioning stimuli demonstrated differences between concussed and nonconcussed individuals. Additionally, mathematically combining results from the measures yields a unique central nervous system (CNS) profile that describes an individual’s information processing capacity. A comparison was made of CNS profiles of concussed vs. nonconcussed individuals and demonstrated with 99% confidence that the two populations are statistically distinct. The study established solid proof-of-concept that cortical metrics have significant potential as a quantitative biomarker of CNS status.

INTRODUCTION

Currently, there is no standard, reliable, cost-effective paradigm or methodology for assessing the degree to which the central nervous system (CNS) is impacted by neurological disorders. One of these disorders or systemic central alterations due to trauma is concussion, or mild traumatic brain injury (mTBI). Although awareness of concussion and mTBI is significantly growing in the general public, there is still no standardized, quantitative, biologically based methodology that is effective for assessing the impact of mild neuro-trauma. Current existing methods and products for this need are expensive, extremely slow, and in many cases fail to definitively and quantitatively diagnose the problem. For example, medical imaging technologies—though they are able to discern differences in subjects with traumatic brain injury —show few or no differences for mTBI or concussion, are costly (about $1K per scan), are not portable, and are not practical for getting a quick assessment. No modern medical imaging techniques are as sensitive to subtle alterations in cortical information processing as those detected by sensory percept. While it is unlikely that there will be any medical imaging technologies able to provide such high resolution in the near future, it is even more improbable that such a technology could be widely distributed.

One of the greatest issues with concussion, or mTBI, is determination of return-to-duty status for the military or return-to-play status for athletes at multiple levels of competition (secondary school, college/university, and professional level). Because injury from secondary concussions can be much more serious, if not fatal, during the critical postconcussion recovery period, it is imperative that methods for this determination be developed. Several years ago, we proposed to design and fabricate a noninvasive, portable, sensory-based diagnostic system using state-of-the-art technology to investigate cortical information processing. Sensory perceptual protocols were designed based on our findings from in vivo studies of cerebral cortical dynamics in nonhuman primates (and thus called cortical dynamic metrics or “cortical metrics”). These proved successful in that a number of specific protocols appeared to be very sensitive to detecting differences between subjects with compromised neurological conditions and healthy controls. Multiple proof-of-concept studies have independently demonstrated that a number of these newly developed metrics are sensitive to systemic cortical alterations.1–16

The somatosensory system is uniquely suited for the design of a diagnostic system for overall cortical health for a number of reasons. First, the somatotopic organization of the somatosensory system provides an ideal template for evoking cortical–cortical interactions in adjacent or near-adjacent cortical regions. Second, ambient environmental noise in the system can be easily controlled (i.e., it is less likely that a patient will be exposed to distracting tactile input than auditory or visual input). Third, the somatosensory system is the only sensory system that is highly integrated with the pain system, and this is often an important aspect of a patient’s diagnosis. Fourth, a key concept in the model is that alterations in sensory percept occur in parallel with alterations in systemic cortical alterations, and “sampling” from the center of the brain (where the somatosensory cortex is located) is more analogous to obtaining a noninvasive biopsy of the cerebral cortex than any other sensory modality.
In this study, we obtained cortical metrics from both concussed and nonconcussed individuals, and subsequently, comparisons of the results were obtained that demonstrated that concussion had impacted the metrics significantly.

**METHODS**

A portable, noninvasive tactile stimulator was designed and fabricated to deliver stimuli to adjacent finger tips (previously described in Holden et al17 [Fig. 1]). Taking advantage of the somatotopic relationship between skin and cortex, biologically based hypothesis-driven protocols were designed to evoke interactions between adjacent cortical regions and investigate fundamental mechanistic changes that occur in cortical-cortical interactions. The measured changes in sensory percept can be easily and rapidly obtained (1 to 3 minutes per test) in a manner similar to reading an eye chart, and the battery of tests described below takes approximately 20 minutes to administer. In this report, we describe three sets of paired metrics, which are relatively simple sensory perceptual measures obtained in the presence and absence of a conditioning stimulus. Because the conditioning stimuli result in healthy controls performing worse, we define these conditioning stimuli as confounding or illusion-inducing. Descriptions of the paired tests administered are described in the section below after the general procedure section.

**Subjects**

Data were collected from 89 college students (67 male, 22 female, mean age = 20.1 years, and SD = 1.2 years), of which 31 experienced a sports-related concussion (15 played football, 7 basketball, 7 soccer, and 2 lacrosse). All concussed athletes were diagnosed with mTBI in the form of a concussion by a certified athletic trainer and the team physician. Data were collected from 89 college students (67 male, 22 female, mean age = 20.1 years, and SD = 1.2 years), of which 31 experienced a sports-related concussion (15 played football, 7 basketball, 7 soccer, and 2 lacrosse). All concussed athletes were diagnosed with mTBI in the form of a concussion by a certified athletic trainer and the team physician with the help of the Sport Concussion Assessment Tool 2 (SCAT-2) and had no prior history of concussion or any other diagnosed medical conditions. The assessments reported were obtained in 1 to 3 days postconcussion. The experimental procedures were reviewed and approved in advance by an institutional review board.

**General Procedure**

During the experimental session, the subjects were situated with the left arm on an armrest attached to the head unit of a portable four-site vibrotactile stimulator. Mechanical stimulation was applied on the glabrous tips of the second (index, D2) and/or the third (middle, D3) fingers of the left hand. An automated procedure guided subjects through a series of questions (answered via computer mouse) related to what the subjects perceived on D2 and D3. In each of the procedures described below, a simple tracking procedure that utilized a two-alternative forced choice (2AFC) paradigm was used to determine an individual’s difference limen (DL). The tracking procedures for each of the protocols queried the individual as to which of two stimuli were larger (amplitude discrimination), which of two stimuli came first (temporal order judgment [TOJ]), or which of two stimuli lasted longer (duration discrimination), and differences between the two stimuli delivered were made smaller when subjects answered correctly.

Visual cueing was provided via a computer monitor during the experimental runs. Specifically, an on-screen light panel indicated when the subject was to respond. An audiometer was used to make sure that no auditory cues were emitted from the stimulator during delivery of the stimuli. Practice trials were performed before each test to allow the subjects to become familiar with the test, and correct responses on three consecutive training trials were required before commencing with the data acquisition portion of the test. The subject was not given performance feedback or knowledge of the results during data acquisition.

**Paired Cortical Metrics No. 1: Amplitude Discrimination Capacity in the Presence and Absence of Confounding Conditioning Stimuli**

**Baseline Metric**

Amplitude discriminative capacity is defined as the minimal difference in amplitudes of two mechanical sinusoidal vibratory stimuli at which an individual can successfully identify the stimulus of larger magnitude. Two stimuli were delivered simultaneously to D2 and D3, and discrimination capacity was assessed using a previously described 2AFC tracking protocol.1,11,12,14–16,18,19 The standard stimulus was set at 200 μm and the test stimulus was initially 400 μm. This difference was subsequently decreased or increased as a result of subject response (decreased for correct answers and increased for incorrect responses). Which of the two fingers received the standard stimulus and which finger received the test stimulus was chosen randomly on each trial.

**FIGURE 1.** Four-site vibrotactile simulator. Each of the four probe tips was positioned by rotating the four independently positioned drums to maximize contact between finger pads and the simulator tips.
Illusory Conditioning

The amplitude discrimination procedure described above was repeated in the presence of a vibrotactile conditioning stimulus delivered 1 second before the presentation of the pair of tests and standard stimuli (Fig. 2). The result of such a protocol modification is that the DL is typically significantly elevated due to a healthy subject’s ability to adapt to the stimulus.1,7,10,11,14,16

Paired Cortical Metrics No. 2: TOJ in the Presence and Absence of Confounding Stimulation

Baseline Test

To evaluate TOJ, two sequential taps were delivered, one to each digit tip, with an initial interstimulus interval of 150 ms. The interstimulus interval was subsequently reduced as a result of subject response as defined by a 2AFC protocol. The finger that received the first of the two pulses was chosen randomly on each trial. Subjects were queried as to which finger was tapped first.

Illusion-inducing Conditioning

TOJ was assessed in the presence of simultaneously delivered synchronized 25 Hz conditioning stimulation before the TOJ task. In healthy controls, this synchronized conditioning typically significantly impacts TOJ, but it does not impact TOJ in some neurologically compromised individuals.3,13,20

Paired Cortical Metrics No. 3: Duration Discrimination Capacity in the Presence and Absence of an Illusory Confound

Baseline Metric

Duration discriminative capacity is the minimal difference in durations of two stimuli at which an individual can successfully identify the stimulus of larger duration. Sequential stimuli were delivered to D2 and D3. Discrimination capacity was assessed using a 2AFC tracking protocol, and subjects were queried as to which of the two digits received the longer stimulus duration. The standard stimulus lasted 500 ms and the initial test stimulus lasted 750 ms. The finger and order of the stimuli were chosen at random on each trial. The duration of the test stimulus was reduced when subject responses were correct and increased when responses were incorrect.

Illusion-Inducing Confound

Duration discrimination capacity was assessed in the presence of an increased standard amplitude. Increasing the amplitude results in a neurophysiological response that is longer in duration21,22 and would predictably make it more difficult for healthy controls to correctly discriminate duration.

Data Analysis

Statistical significance of the difference of the means between the concussed and healthy control samples was assessed separately for each of the six cortical metrics using a paired t test. In addition, using the approach of quantitative sensory testing—which treats the performance of a human subject on a battery of psychophysical tests as a multidimensional “sensory profile” of that subject, potentially reflecting the functional status of his/her CNS23,24—quantitative performance of each subject in this study on six cortical metrics tests was treated as the “CNS profile” of that subject, localizing him/her in a 6-dimensional cortical metrics space. The cortical metrics space is an abstract space in which each coordinate axis corresponds to one of the cortical metrics. Since different metrics vary on different scales, to make different axes of the cortical metrics space comparable to each other, each metric contributing to the CNS profile was autoscaled by subtracting its mean (measured over the entire studied subject population) and dividing by its standard deviation. Hotelling’s T-squared test of the difference between the multivariate means of different populations25 was used to compute the statistical significance of the difference in the locations in the cortical metrics space of the centers of the concussed and healthy control samples. Finally, to graphically visualize the spatial relationship between the clusterings of the concussed and healthy control subjects in the cortical metrics space, the 6-dimensional space and all the subject-representing data points in it were projected, using the Partial Least Squares Discriminant Analysis.
(PLS-DA) algorithm,26 onto a 2-dimensional plane oriented such as to maximize the separation between the concussed and healthy control distributions.

RESULTS

**Paired Cortical Metrics No. 1: Amplitude Discrimination Capacity in the Presence and Absence of Confounding Conditioning Stimuli Demonstrates That Concussed Individuals Adapt Less Than Nonconcussed Individuals**

Control data were consistent with amplitude discriminative capacity measures that previously demonstrated robustness across the age spectrum.16 Figure 3 shows that concussed subjects performed worse on the amplitude discrimination task than did healthy controls (DL of controls 30.1 ± 1.3 μm vs. concussed 42.1 ± 5.9 μm for a 200 μm standard).

With the addition of a confounding conditioning stimulus, amplitude discriminative capacity is typically worse across the age spectrum16 and the results in Figure 3 are consistent with that previous finding for control values (DL increased from 30.1 ± 1.3 μm to 63 ± 2.2 μm with confound). However, concussed subjects did not perform significantly differently postconditioning (DL increased from 42.1 ± 5.9 μm to 46 ± 5.4 μm).

**Paired Cortical Metrics No. 2**

Typically, healthy individuals have a TOJ capacity on the order of 30 to 40 ms, and in the presence of an illusory conditioning stimulus healthy controls perform significantly worse on the same TOJ task.13,20 Figure 4 shows that healthy control data in this study are consistent with that finding (DL increases from 36.4 ± 2.8 ms to 95.2 ± 4.3 ms), and concussed subjects do not appear to deviate significantly from healthy controls on the baseline TOJ metric. However, concussed subjects did not perform worse in the presence of the “illusion-inducing confound” (DL for concussed subjects was 40.1 ± 7.6 ms without conditioning vs. 42.5 ± 7.3 ms with conditioning).

**Paired Cortical Metrics No. 3**

Comparison of healthy controls and concussed subjects (Fig. 5) suggests that while there is little or no difference between duration discriminative capacity of the two subject groups (DL for controls 64.6 ± 3.7 ms vs. 75.2 ± 5.4 ms for concussed individuals), the discriminative capacity of healthy controls is impacted by the illusion-inducing confound (DL for controls increased to 124.7 ± 15.2 ms) while the confound does not appear to impact the discriminative...
capacity of concussed subjects significantly (DL increased to 77.7 ± 10.3 ms).

**Multivariate Analysis Demonstrates Different Profiles for Concussed vs. Nonconcussed Individuals**

Treating the performance of any given subject on multiple cortical metrics tests as a multidimensional metrics vector (or CNS profile) in an abstract space, each axis of which corresponds to one of the test metrics, we can compare the spatial distributions of such vectors in the concussed vs. healthy control groups. To visualize these two group distributions, they were projected onto a 2-dimensional plane, shown in Figure 6, using PLS-DA algorithm. In Figure 6, control individuals are shown as black dots and concussed individuals are shown as asterisks, revealing that these two groups form distinct, only partially overlapping clusters. While a few concussed individuals are mixed in among the control individuals—thus indicating that their performance on the cortical metrics tests was indistinguishable, as a whole, from the control population—the majority of concussed subjects were clearly displaced relative to the control distribution. Hotelling’s T-squared statistic indicates with greater than 99% confidence that these two populations have different centers.

Figure 7 suggests that the distance between the performance vector of a given individual and the center of the healthy control distribution might be indicative of the concussion impact. The plot in Figure 7 was constructed by computing the average concussion symptom score SCAT-2 for 9 different subsets of concussed subjects, each subset farther away from the center of the control distribution. This plot shows that concussed subjects with more distant performance vectors tended to have higher SCAT-2 scores.

**DISCUSSION**

For the past several years, we have been developing protocols that utilize illusion-inducing confounds that alter the perception of a sensory stimulus. For example, delivery of a repetitive vibrotactile conditioning stimulus to one of two skin sites before an amplitude discrimination task results in degradation of performance in healthy controls. However, a number of neurologically compromised subjects have demonstrated that this conditioning stimulus—or the illusion-inducing confound—does not impact their performance. In other words, some subject populations (e.g., individuals with autism, alcoholism, multiple types of chronic pain, and concussion) do not adapt to the conditioning stimulus, and because the illusion-inducing conditioning stimulus has little or no impact, they actually “outperform” healthy controls on the postconditioning amplitude discriminative task. Another example of an illusion-inducing conditioning stimulus is one in which healthy controls perform worse (but neurologically compromised subjects do not) on a TOJ task in the presence of synchronized, but not asynchronized, conditioning stimuli. Duration discrimination, or the ability to accurately determine which of two stimuli has a longer temporal duration, is impacted in an illusory manner by increasing the intensity of one of the stimuli. This illusory condition apparently has less of an impact on individuals who are concussed.

It should be emphasized that the measures described in this report do not simply reflect alterations in tactile perception, but rather differences in cortical information processing capacity. The lack of a difference in amplitude discrimination
with vs. without the illusion-inducing confound reflects a systemic cortical alteration and a decrease in the individual’s capacity to adapt. In other words, plasticity has been reduced, and the alteration in the somatosensory-based task is a reflection of a systemic cortical alteration. The lack of a change in the TOJ metric in the presence of the confound also reflects a systemic cortical alteration—cortical ensembles are no longer coordinated in their response to the tactile conditioning, and the TOJ cortical metric reflects an alteration in functional connectivity. Similarly, the lack of an impact of the confound on duration discrimination reflects a systemic alteration in neuron-glial interactions, possibly due to neuroinflammation that occurs with concussion.

The potential utility of this work is highly significant. A simple, fast, noninvasive, and cost-effective means for assessing the impact of concussion on CNS health that could be utilized by health care providers would have a far-reaching impact. To date, there are no standardized, quantitative measures that are biologically based for assessing concussion. The advantage of the proposed methodology is that it will be low-cost, easy to use, and effective at both providing information about a patient that would enable a diagnostician to make a more informed decision about diagnosis or treatment, and providing a means for assessing treatment efficacy.

ACKNOWLEDGMENT

This work was partially supported by the Office of Naval Research and through an award made to Cortical Metrics, LLC as one of the winners of the General Electric Company/National Football League Head Health Challenge.

REFERENCES

An Interim LAeq8 Criterion for Impulse Noise Injury

Brissi Zagadou, PhD*; Philemon Chan, PhD*; Kevin Ho, MS*

ABSTRACT
Objectives: We present a method to account for the effects of the hearing protection devices (HPDs) for use with the 8 hours equivalent A-weighted energy (LAeq8) criterion. The method involves the calculation of the LAeq8 equivalent unprotected free-field dose (LAeq8EUFF), which is obtained by using the insertion loss (IL) data of the HPD together with free-field pressure measurements. Methods: The method was validated against the historical U.S. Army Medical Research and Materiel Command walk-up study data with volunteers exposed to simulated large weapon noise wearing a range of HPDs. The IL data were obtained using standard acoustical test fixtures fitted with the matching HPDs in replicated field tests and using shock tubes at conditions comparable to the actual exposure intensities. Logistic regression calculations were performed to correlate the LAeq8EUFF values against the walk-up study outcomes to determine the L(95,95) threshold for the protection of 95% of the population with 95% of the time. Results: Data comparison shows that L(95,95) is 89 dBA, which is slightly higher than the 85 dBA criterion but falls in the 80 to 90 dBA range as used by various NATO nations. Conclusions: Therefore, considering the limitation of the walk-up dataset, it is conservative to adopt the 85 dBA threshold for general application.

INTRODUCTION
Hearing loss (HL) is one of the most disabling conditions for both the military and civilians. According to the U.S. Department of Veterans Affairs, 21% of nearly 7 million cases contain the pathological outcomes of auditory damage. The World Health Organization has also pointed out the significance of HL with a staggering 5.3% of the world population suffering from a variety of causes including noise exposures. Noise-induced HL remains one of the top medical medical problems. Soldiers are routinely exposed to high-level impulse noise from large weapons during training exercises or on the battlefield that causes severe and permanent hearing damage. Impulse noise injury can be more effectively prevented if an adequate medical standard can be provided to set the appropriate limit on exposure levels. This task can be achieved if a representative injury dataset from impulse noise exposures is available for modeling.

As part of the Blast Overpressure Project, the U.S. Army Medical Research and Materiel Command conducted a test series at the blast test site at Kirkland Air Force Base in Albuquerque, New Mexico, commonly known as the “Walk-up” or “Albuquerque” study. The historical human walk-up study collected valuable data but did not provide a new standard. The data include the waveforms of the simulated large weapon noise for several distances (1, 3, and 5 m) from the noise source and the resulting human temporary threshold shift (TTS) data for five types of hearing protection devices (HPDs) worn during the study. The HPDs included the regular (unmodified) and modified earmuffs, the French, Rucker, and Perforated earplugs. The use of the modified muff was to simulate improper fitting of the earmuff that was considered a common (unintentional) occurrence.

Using the modified earmuff data alone, Chan et al performed statistical analysis that confirmed the deficiency of various noise standards. The data predicted that the current standards overestimate the injury threshold by about 10 dB. The data also showed that the 8 hours equivalent A-weighted energy (LAeq8) model provides a better fit compared to the peak-duration based standards. The analysis, however, was based on the crude approximation of 15 dB attenuation for the modified earmuffs due to the lack of insertion loss (IL) data.

The American Institute of Biological Sciences recommended the use of the LAeq8 model as an interim standard until a robust biomechanically based auditory standard is found. This choice was motivated by the extensive research conducted on the LAeq8 model and its popular use across nations. A system acquisition standard was recently adopted by the military, MIL-STD-1474E, which, however, is not a medical standard. MIL-STD-1474E includes a variant of the LAeq8 that has not been validated and the current auditory hazard assessment algorithm for the human that is still under rigorous investigation for validation and improvement.

In order to use the LAeq8 standard as a general procedure, a method to account for HPDs is needed. The LAeq8 standard limits the daily unprotected noise exposure to 85 dBA, but this threshold was established based on free-field pressure measured at the ear position without HPDs and in continuous noise conditions. For intense impulse noise for operational application, HPDs are required. In this case, the use of the free-field pressure data is not appropriate for correlation with injury because the free-field data does not account for the effects of the HPDs, which can become nonlinear at high-exposure levels. For this reason, the use of a constant
attenuation is not adequate. A procedure is, therefore, needed to account for the effects of HPDs and to set the threshold.

In this article, we present an injury correlation based on the LAeq8 method that includes a test procedure to incorporate the HPDs using IL data derived from acoustical test fixture (ATF) measurements. The LAeq8-based correlation method is validated against the human walk-up study data.

**METHODS**

The method consists of (1) establishing an ATF method to characterize the HPDs, (2) incorporating the HPD characteristics to calculate an exposure dose based on the LAeq8 model, and (3) performing statistical analysis to correlate the dose to the human injury data to predict the L(95,95) threshold for the protection of 95% of the population, 95% of the time against auditory TTS exceeding 25 dB.

**ATF Method for HPDs Characterization**

*Shock Tube Evaluation of the ATF*

An ATF suitable for high-impulse noise was selected and its response was verified against benchmark data. The G.R.A.S. 45CB (G.R.A.S. Sound & Vibration, Twinsburg, Ohio) was selected for this study based on microphone specs, the maximum level permissible of 174 dB at the eardrum, and the ability to handle both earplugs and earmuffs. The G.R.A.S. 45CB ATF is fully ANSI S12.42 compliant. During the tests, the pinna/ear canal material was heated to 37°C in accordance to the ANSI S12.42 requirement. The transfer function of the open-ear (TFOE) was measured using shock tube experiments to evaluate the ATFs for impulse noise exposure (Fig. 1). The results were compared to the human TFOE data from Shaw and Vaillancourt, obtained with continuous noise exposures.

*Replication of Field Tests to Collect ATF Eardrum Data*

The human walk-up tests were replicated to collect the ATF eardrum data at matching conditions. To quantitatively compare the original and the replicated free-field waveforms, matching criteria based on the waveform characteristics were defined. The criteria were that when the peak level, A-duration, impulse, and A-weighted sound exposure level (SELA) of the replicated waveform lie simultaneously within one standard deviation of the original blast overpressure (BOP) waveform, the two waveforms are considered matched. Thus, a series of calibration tests were performed to search for the distance from the blast center that would satisfy the matching criteria for each test condition designated by the (nominal) distance and level. With the respective optimal distances determined, the ATF fitted with HPDs and the pencil gauges were placed at these distances from the blast, and the eardrum data were collected. Six to seven shots per condition were measured for each distance and blast level. Figures 2 and 3 show the test setup for the 1, 3, and 5 m tests, respectively. The recovered HPDs are shown in Figure 4, and Figure 5 shows the ATF with the recovered earmuff and other protective gear worn in the walk-up study.

**FIGURE 1.** Shock tube test setup for acoustical test fixture (ATF) evaluation.

**FIGURE 2.** One and three meter setup. Two acoustical test fixtures (ATFs) fitted with hearing protective devices, helmet, and goggle are shown with the right ear facing the explosion, which originates from underground guided by a center tube. Pencil gauges shown are placed at the level of the ATF’s ear at the same distance from the blast.

**FIGURE 3.** Five meter test setup. Two acoustical test fixtures fitted with the hearing protective device, helmet, and goggle are shown paired with pencil gauges. The explosive charge is hung above the ground between three tall cylindrical protection tubes.
Characterization of HPDs

The HPD was characterized by its IL, defined as the difference between the pressures recorded at the eardrum without and with the HPD for the same free-field pressure:

\[
\text{IL} = P_{\text{ED}} - P_{\text{ED}}^0 = P_{\text{FF}} - P_{\text{ED}}^0 + \text{TFOE}
\]

where \( P_{\text{ED}}^0 \) denotes the pressure measured under the HPD at the location of the eardrum.

The IL was calculated using the TFOE and the ATF eardrum data, as in Equation 1, for all five HPDs used during the human walk-up study. For the free-field peak pressure level greater than 180 dB as observed above Level 4 in the walk-up study, eardrum measurements could not be made using the ATF due to microphone saturation. The ILs for the high-intensity levels were obtained by extrapolating from the ILs at lower levels.

**LAeq8 Dose Calculation With HPD**

Using the IL, the LAeq8 equivalent unprotected free field (LAeq8\text{\textsubscript{EUFF}}) dose with the HPD worn is defined as:

\[
\text{LAeq8}_{\text{EUFF}} = 10 \log_{10} \left( \sum_{k=1}^{n} \left( \frac{10^{\text{L}_{\text{FF}}(k) - \text{IL}(k)} - 1}{10} + \text{F}(k) \right) \right) - 10 \log_{10}(T_{\text{8h}}) + 10 \log_{10}(N)
\]

where the summation is based on 1/3-octave band frequency \((k)\) intervals, F\((k)\) is the A-weighting correction factor for the 1/3-octave band frequency interval, \(T_{\text{8h}}\) is the 8-hour conversion factor, and the last term represents the traditional dosage accumulation rule with \(N\) being the number of shots.

**LAeq8 Correlation With Human Walk-Up Test Data**

The logistic regression fit to the BOP injury data with the LAeq8\text{\textsubscript{EUFF}} dose as the predictor variable was used to validate the LAeq8 method. The calculations were carried out for the BOP HPD series to determine the threshold dosage for 25 dB TTS. In the present analysis, all human BOP data collected in the free field were used, which included the data from the modified and unmodified earmuff and earplug test.
series. Data from the 1, 3, and 5 m series were pooled together for LAeq8 correlation with the injury data.

The statistical analysis procedure is identical to that in Chan et al.³ The population average logistic regression model⁸ is used to correlate the human injury data with LAeq8EUFF to establish the threshold for TTS₂ ≥ 25 dB at 2 minutes after exposure. The predictor variable x is the LAeq8EUFF dose and the response variable y is 1 when injury (TTS ≥ 25 dB) occurs and 0 when no injury occurs. The logit function g(x) is thus given in terms of the conditional probability of injury given a LAeq8EUFF dose, \( \pi (y = 1 | x) \) as:

\[
g(x) = \log \left( \frac{\pi (y = 1 | x)}{1 - \pi (y = 1 | x)} \right) = a + b \times x
\]  

The logistic regression model coefficients a and b were estimated by fitting the model with the pooled human injury response and the associated LAeq8EUFF dose for all test cases using the STATA software (StataCorp, Stata: Release 11. Statistical Software, 2009, StataCorp LP: College Station, Texas).

The \((1 - \alpha)100\%\) confidence interval (CI) for the estimated probability of injury is calculated from the mean estimates and covariance matrix Cov_{ab} (i, j) of the regression coefficients and by assuming normal distribution of the error of the regression. The formulas for the CIs for the estimated probability of injury, logit, and standard error of the logit are summarized in Equation 4 to 6.

\[
\hat{\pi}_\pm = \frac{1}{1 + e^{-\hat{g}_\pm}}
\]
\[
\hat{g}_\pm = \left( \hat{a} + \hat{b} \right) x \pm \delta g
\]
\[
\delta g = z_{1 - \alpha / 2} \\
\times (Cov_{ab}(1,1) + 2 \times Cov_{ab}(1,2) + Cov_{ab}(2,2) \times x^2)^{1/2}
\]  

In Equation 6, \( z_* \) satisfies the cumulative standard normal distribution function, \( \phi(z_*) = x \) and \( z_{0.975} = 1.96 \) corresponds to 95\% confidence in the injury prediction. The LAeq8EUFF dose threshold for the protection of 95\% of the population with 95\% confidence, \( L(95,95) \) is given by the upper bound of the CI with probability of injury of 5\%, i.e., \( \pi_+ = 0.05 \).

RESULTS

G.R.A.S ATF Evaluation

The G.R.A.S ATF TFOE measurements show comparable results with the TFOE measured from human subjects. Figures 6 and 7 show the TFOE comparison result of ATF vs. human for normal and grazing noise exposures, respectively. Figure 6 shows that the TFOE generated from shock tube testing at normal incidence at 163 dB peak pressure level (PPL) is fairly close to the historical curve from Shaw and Vaillancourt⁷ obtained from human subjects at low sound levels. Similarly, Figure 7 shows that at grazing incidence, the TFOE obtained from shock tube test at 164 dB PPL compares closely to that from humans. The similarity between the TFOE obtained from the ATF and humans appear to be independent of direction for the highest intensity of \(~164\) dB PPL tested, but slightly dependent on the impulse intensity level.

ATF Eardrum Data Comparison

Figure 8 shows the results of the statistical comparison between the ATF eardrum and the undermuff waveforms from the walk-up tests for the peak pressure and the SELA as a function of the blast level. The peak and SELA increase linearly for the ATF eardrum data, and somewhat less linearly for the human undermuff data toward the high levels. The difference between the mean peak and SELA values is \(~5\) to \(~7\) dB, which represents the amplification from the ear canal...
entrance to the eardrum. It is noted that the ATF eardrum data show very little variability compared to the undermuff data.

IL Comparison With Real-Ear Attenuation at Threshold (REAT)
The IL results are shown in Figures 9 and 10 for the earmuffs and earplugs, respectively, with comparison to the REAT data. In Figure 9, the IL data of the earmuffs somewhat overlap with the REAT data. However, depending on the level, differences between the IL and the REAT as high as 10 dB can be observed. The overall trend indicates that the IL of the earmuffs is closer to the REAT than that for the earplugs. As shown in Figure 10, there are indeed large differences between the IL of the earmuffs and the REAT. For the French no. 1 plug, the difference is as high as 25 dB; and for the Rucker and the perforated plug, the differences are less pronounced, but at least 10 dB. The results indicate that REAT should not be used as IL.

Interim LAeq8 Validation
Figure 11 shows the result of the logistic regression calculation with injury data from all HPDs when the IL derived from the historical TFOE are used in the LAeq8EUFF dose calculation. The L(95,95) threshold predicted is 89 dBA as shown by the large, open square on the upper bound of the CI. This value is 4 dB higher than the standard threshold of 85 dBA shown by the vertical bar. The data points shown represent the mean failure rates based on 10-bin data grouping, and each short horizontal bar is the standard deviation of the dose for each data bin. The data shows that all injuries occur to the right of the vertical bar, i.e., above the 85 dBA threshold.

DISCUSSION
The TFOE obtained from the historical human tests was used in the IL calculation. As shown in the evaluation results, the TFOE compared fairly well with that obtained from the shock tube measurements with an ATF. In addition, the effects on the LAeq8EUFF dose calculation of the differences in the TFOEs were verified and produced a dose variation of less than 1 dB. Therefore, the TFOE measured either with an ATF or human can be used for the LAeq8 dose calculation.

Three factors that may affect the predictions are the effects of the head orientation, bone conduction (BC) and the ILs for high-intensity levels. These factors are discussed in turn.

The orientation effects on IL measurements and prediction should be considered. However, although these effects on IL measurements can be readily assessed using an ATF with the HPD, it is difficult to evaluate the significance on the threshold prediction. The main reason is that the available injury data do not cover all orientations. Animal studies may help guide the inclusion of the orientation effects.

The effects of BC on the IL and LAeq8 prediction were considered to be minimal and BC correction was not made. Since the ATF does not model the BC effect, an adjustment of the IL measured with the ATF is commonly made based on the BC limits. For impulse noise, however, this sort of correction appears not to change the ILs significantly.9

A verification of the extrapolated IL results is desirable. A high-pressure microphone sensor installed in the ATF may be used for this purpose. Our results show that the extrapolated IL values for the high-intensity follow the data trend well, and we do not expect the measured IL data to deviate significantly from their current values. Therefore we do not expect the verification results to significantly affect the current LAeq8 predictions.

The L(95,95) threshold prediction by Chan et al with their LAeq8 analysis is ~25 dB higher than that found in the present analysis. As previously noted, the former analysis did not include the earplugs, which caused the most injuries, nor did it include the unmodified earmuff data, which represent the largest dataset and contain the least injuries. In addition, the analysis by Chan et al did not use adequate attenuation data for the earmuff. In the present analysis, using the traditional dose accumulation according to
the 10log(N) rule as used in Chan et al, all BOP HPD data were pooled together to predict the injury threshold. The IL for all HPDs used was measured under real life exposure conditions using an ATF. All these factors explain the difference between the threshold predictions.

Although the use of an ATF constitutes the most practical method for characterization of HPDs, it is recognized that the method does not account for human factors such as ergonomics, HPD fit and fitting methodology and thus prevents systematic generalization of the ATF results to real world performance. Fitting the HPD to the ATF is equivalent to the investigator fitting the HPD with human subjects during testing, where all the necessary precautions are taken by the investigator to ensure proper fit of the HPD. In a real world, however, the subjects wear the HPD themselves, and fitting inconsistency is thus unavoidable. HPD misfit is the predominant factor that can cause the ILs as measured with the ATF to deviate from their performance during operations. The earmuffs can easily be moved during operations, whereas the insertion depth of the earplug can vary from one human subject to another. The variability in these human factors can have a pronounced influence on HPD. On the other hand, no one method for measuring IL is perfectly accurate. The used of REAT is not adequate for high-level impulse, even when corrected for physiological noise masking and BC; this is probably true at least for the earplugs used in the BOP walk-up study as shown in the results.

It is recognized that the human BOP walk-up study data is still limited. The data simulate a limited class of blasts although the inclusion of additional weapon noise types would certainly improve the predictions. Although representing both earmuffs and earplugs, double protection is not simulated, and the HPDs used in the study may not represent modern HPD types. Modern earplugs or earmuffs appear to show different attenuation characteristics even among the same category when measured at relatively low levels.


**FIGURE 9.** Insertion Loss vs. Real-Ear-Attenuation-at-Threshold (REAT) for earmuffs. (A) 1 m test with modified earmuff, (B) 3 m test with modified earmuff, (C) 5 m test with modified earmuff, and (D) 5 m test with unmodified earmuff.
However, the response of these HPDs as measured with the same ATF but at high-level impulse conditions are not known. Another limitation is that the injury data result from normally incident blast conditions only. These conditions are clearly different from the operational conditions. For these various reasons, the use of the LAeq8-based correlation method that is not ambitiously relaxed with the risk of causing injury such as the LAeq8 method presented in this article is appropriate.

CONCLUSION

An LAeq8 standard is presented that includes the use of the ATF to account for the effects of the HPDs by measuring attenuation under realistic exposure conditions. The head orientation effects are captured by the ATF and the intensity of the exposure can be matched. The ATF method shows that the 85 dBA threshold is valid for large weapon noise as demonstrated using the most complete and scientifically relevant injury dataset available to date. The method provides a major step forward to harmonize with environmental noise and unprotected rifle noise criteria and should be validated against small arms.

ACKNOWLEDGMENT

This work was funded through the U.S. Army Medical Research and Material Command (MRMC) under contract W81XWH-11-D-0011.
REFERENCES


Impulse Noise Injury Model

Philemon Chan, PhD*; Kevin Ho, MS*; Allen F. Ryan, PhD†

ABSTRACT  Objectives: The new Auditory 4.0 model has been developed for the assessment of auditory outcomes, expressed as temporary threshold shift (TTS) and permanent threshold shift (PTS), from exposures to impulse noise for unprotected ears, including the prediction of TTS recovery. Methods: Auditory 4.0 is an empirical model, constructed from test data collected from chinchillas exposed to impulse noise in the laboratory. Injury outcomes are defined as TTS and PTS, and Auditory 4.0 provides the full range of TTS and PTS dose–response curves with the risk factor constructed from A-weighted sound exposure level. Human data from large weapons noise exposure was also used to guide the development of the recovery model. Results: Guided by data, a 28-dBA shift was applied to the dose–response curves to account for the scaling from chinchillas to humans. Historical data from rifle noise tests were used to validate the dose–response curves. New chinchilla tests were performed to collect recovery data to construct the TTS recovery model. Conclusions: Auditory 4.0 is the only model known to date that provides the full TTS and PTS dose–response curves, including a TTS recovery model. The model shows good agreement with historical data.

INTRODUCTION

Hearing loss remains one of the top military medical problems, but there is still no dose–response model for the assessment of impulse noise injury and recovery. According to the 2011 report by the U.S. Government Accountability Office, the annual disability payment to veterans suffering from hearing impairment exceeded $1.1 billion in 2009, and recommendations were given to improve hearing conservation for the Department of Defense and Veteran Affairs.¹

More than 250,000 service members have reported hearing loss following redeployment from the Gulf War conflicts and 60% of veterans returning from Iraq and Afghanistan come home with hearing loss and tinnitus. According to the Veteran Affairs in 2009, there were over 700,000 veterans receiving service-connected tinnitus disability, and this number is projected to increase to 1.5 million by 2014, translating into annual tinnitus disability-related payments of $2.26 billion.² Current trends suggest that the incidence of tinnitus and hearing loss is increasing 13% to 18% annually.³

Impulse noise from blast explosions and weapon firings are shock waves that can cause traumatic injuries to the auditory system. Auditory injury outcomes are assessed as temporary threshold shift (TTS) and permanent threshold shift (PTS). For impulse noise, TTS is usually assessed at 2 minutes after exposure (TTS₂) and undergoes recovery with time. At low TTS, full recovery could occur within a few hours, but at high TTS, full recovery may not be possible, resulting in PTS. MIL-STD-1474D is the current military impulse noise standard in the United States, which is only an occupational standard used for protection against threshold of blast injuries from weapon-firing exercises. For occupational protection, TTS₂ of 25 dB is taken as the threshold against PTS. The European Union generally adopts the LAeq8 standard, which limits the accumulated dose of A-weighted energy in an 8-hour day not to exceed 85 dBA. The critical limitation of both MIL-STD-1474D and LAeq8 is that they do not provide dose–response curves for predicting the risks of the full range of TTS and PTS beyond threshold (TTS₂ = 25 dB). Furthermore, there is no model available that predicts TTS recovery after impulse noise exposure.

To provide a full capability for the assessment of hearing loss from impulse noise, a dose–response model including the time course of near-term TTS recovery is needed. The current occupational standards are inadequate for application to operational conditions, where the insults are likely beyond threshold levels. To develop dose–response curves, injury data for the full range of TTS and PTS are needed, but human test data at high injury levels are practically nonexistent. Furthermore, since most previous work in impulse noise was performed to support the occupational protection standards, TTS recovery data were usually collected at fairly large time intervals (between hours and days) for long-term tracking of the risk of PTS. TTS recovery data at close time intervals are needed.

The objective of this work is to develop an impulse noise injury model that will predict the risk of the full range of TTS and PTS injuries to unprotected ears from impulse noise, including dose accumulation from multiple shots,⁴ and the near-term TTS recovery time course.⁵ The new model has been packaged as software labeled as Auditory 4.0 (L-3 Applied Technologies, San Diego, California).

METHODS

Auditory 4.0 is an empirical model. The dose–response curves were constructed from chinchilla data and scaled to...
humans. Chinchilla data are used because the auditory system of chinchillas is a close surrogate to that of humans. Figure 1 shows the close comparison of the audibility spectra between humans and chinchillas. The injury threshold of chinchillas is lower than that of humans\textsuperscript{4–6} and it is assumed that this difference can be adjusted by scaling the dose–response curves from chinchillas to humans with no effect on the TTS recovery time course. It is a common approach in biomechanics to scale injury data between species because of the realism of the limitation of human data, especially at high injury levels, and comparison to as much human data as possible will be carried out to validate the results as will be shown later.

**Historical Data**

The dose–response curves were constructed from ordered logistic regression of the historical chinchilla data with guidance from human data. The details of the methodology were previously documented by Chan and Ho, with the model refined to include an improved dose accumulation algorithm (Chan P, Ho K: Auditory 3.0 Impulse Noise Injury Model. 2010; Technical Report under Contract No. DD22009DO2. In review by the Joint Non-Lethal Weapons Directorate). A brief summary of the data used is given below.

**Chinchilla Data**

The chinchilla data used were collected from the historical Blast Overpressure Project (BOP) conducted by the U.S. Army Medical Research and Materiel Command (USAMRMC) that is probably the largest data set known to date, covering a wide range of impulse noise, with 905 subjects total\textsuperscript{7,8}. Table I shows the summary of the tests performed and Table II shows the types of stimuli used. Tests were conducted in open-field and enclosure conditions in the laboratory, including tests with multiple shots. Both audiometric and histological data were collected. The TTS and PTS data averaged across 1, 2, and 4 kHz were used for developing the dose–response curves for the present work. The use of averaged data across these three frequencies is based on the guidance from the recommendation of the National Institute for Occupational Safety and Health and the availability of historical human data. The National Institute for Occupational Safety and Health recommended the averaging of data at 1, 2, 3, and 4 kHz for preventing excess risk, which is supported by findings from Phaneuf et al\textsuperscript{9} as a good metric for assessing hearing disability. Even though the historical chinchilla data only contained outcomes at 1, 2, and 4 kHz, it was considered the averaging of data across these three frequencies would be an adequate representation of the outcome data for dose–response correlation. It is known that most of the frequency range of human communication is from 1 to 4 kHz.

To maximize statistical power, the data across test series were pooled together for analysis. TTS and PTS data were each binned in 5-dB intervals for statistical analysis. For the historical USAMRMC data that were used to develop the injury dose–response curves, only the right ear was used as the test ear. The pooling of the data from open-field and

![FIGURE 1. Comparison of audibility curves (SPL = sound pressure level).](image)

<table>
<thead>
<tr>
<th>Study</th>
<th>Animals with Audiometry Data</th>
<th>Animals with Histology Data</th>
<th>Open Field or Enclosure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional Shock Tube, Nonreverberant</td>
<td>109</td>
<td>109</td>
<td>Open Field</td>
</tr>
<tr>
<td>Fast-Acting Valve (5&quot;), Nonreverberant</td>
<td>105</td>
<td>105</td>
<td>Open Field</td>
</tr>
<tr>
<td>Fast-Acting Valve (3.5&quot;), Nonreverberant</td>
<td>105</td>
<td>105</td>
<td>Open Field</td>
</tr>
<tr>
<td>Spark Gap, Nonreverberant</td>
<td>104</td>
<td>104</td>
<td>Open Field</td>
</tr>
<tr>
<td>Conventional Shock Tube, Reverberant</td>
<td>135</td>
<td>135</td>
<td>Enclosure</td>
</tr>
<tr>
<td>Fast-Acting Valve (3.5&quot;), Reverberant</td>
<td>136</td>
<td>136</td>
<td>Enclosure</td>
</tr>
<tr>
<td>Narrow-Band Impact</td>
<td>130</td>
<td>130</td>
<td>Open Field</td>
</tr>
<tr>
<td>290C Driver: 146 and 138 dB Peak Sound Pressure Level</td>
<td>12</td>
<td>12</td>
<td>Open Field</td>
</tr>
<tr>
<td>290C Driver, High Peak Wave</td>
<td>36</td>
<td>36</td>
<td>Open Field</td>
</tr>
<tr>
<td>290C Driver, Low Peak Wave</td>
<td>18</td>
<td>18</td>
<td>Open Field</td>
</tr>
<tr>
<td>290C Driver, 131 dB Peak SPL</td>
<td>5</td>
<td>5</td>
<td>Open Field</td>
</tr>
<tr>
<td>U.S. Army Aeromedical Research Laboratory Conventional Shock Tube, Nonreverberant (Unpublished)</td>
<td>10</td>
<td>26</td>
<td>Open Field</td>
</tr>
</tbody>
</table>
enclosure tests assumes the risk factor will be general enough to apply to both conditions and that will be evaluated with data comparison. The A-weighted sound exposure level (SELA) was used to construct the risk factor accounting for dose accumulation from multiple shots. The recent peer panel review of impulse noise models by the American Treaty Organization occupational impulse noise standards was performed to analyze the Walk-up data that provided a best-fit model for large weapon noise with hearing protectors. The assumption is that the TTS recovery time course would be affected by the use of the hearing protectors. The objective of the Walk-up study was to evaluate the adequacy of the United States and North American Treaty Organization occupational impulse noise criteria against large weapon noise. Therefore, the study was designed to only limit auditory failures to low TTS. Tests were performed at 1, 3, and 5 m from the charge with volunteers wearing different earmuffs or earplugs. Previous work was performed to analyze the Walk-up data that provided a best-fit model for large weapon noise with hearing protectors. A brief summary of the test data used for the present work is shown in Table III.

The Albuquerque data were used to support the development of the TTS recovery model. Even though hearing protectors were worn by the volunteers, it is not expected that the TTS recovery time course would be affected by the use of the hearing protectors. The assumption is that the TTS recovery time course is independent of how the TTS was incurred.

**Chinchilla Test to Collect Recovery Data**

New chinchilla tests were performed to collect data to construct the TTS recovery model. Tests were conducted at the University of California at San Diego (UCSD). Impulse noise was generated using a small shock tube. The main objective of the new test series was to collect near-term TTS data. Threshold shift outcomes were measured using the auditory brainstem response (ABR) method.

**Shock Tube Noise Generation**

A small 2-inch diameter shock tube was used for generating the impulse noise needed for animal testing (Fig. 2). The shock tube consists of a high-pressure driver chamber at one end and an expansion section at the other. The high-pressure driver section is separated from the expansion section by a thin layer of diaphragm that self-ruptures at specific

**Human Rifle Noise Data**

The historical human data collected from rifle noise exposures was from the Committee on Hearing, Bioacoustics, and Biomechanics (CHABA), consisting of data from various sources that essentially formed the basis for the historical development of occupational impulse noise standards in the United States. No noise signature data could be found for the CHABA data. However, the critical metrics, such as peak sound pressure level (SPL) and SELA could be recovered from the published technical reports. The CHABA data were recently reanalyzed by Smoorenburg, showing that the difference between peak pressure level and SELA for a wide range of rifle noise data is about 37 dB. Some of the SELA data needed were obtained using this conversion. Smoorenburg showed that a 4-dB increase in SELA for normal incidence to the ear could be justified from field data analysis, and this rule is adopted for this study. The human rifle noise data were used to validate and refine the dose–response model.

**BOP Human Data**

As part of the USAMRMC BOP research, human data from volunteers exposed to large weapon noise were collected from the study known as the Albuquerque (Walk-up) Study. Because of the high intensity levels, volunteers wore hearing protectors. The objective of the Walk-up study was to evaluate the adequacy of the United States and North American Treaty Organization occupational impulse noise criteria against large weapon noise. Therefore, the study was designed to only limit auditory failures to low TTS. Tests were performed at 1, 3, and 5 m from the charge with volunteers wearing different earmuffs or earplugs. Previous work was performed to analyze the Walk-up data that provided a best-fit model for large weapon noise with hearing protectors. A brief summary of the test data used for the present work is shown in Table III.

**TABLE II. Types of Stimulus Used in Chinchilla Tests**

<table>
<thead>
<tr>
<th>Stimulus Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1–3</td>
<td>Conventional Shock Tube, Nonreverberant</td>
</tr>
<tr>
<td>4–6</td>
<td>Fast-Acting Valve (5″), Nonreverberant</td>
</tr>
<tr>
<td>7–9</td>
<td>Fast-Acting Valve (3.5″), Nonreverberant</td>
</tr>
<tr>
<td>10–12</td>
<td>Spark Gap, Nonreverberant</td>
</tr>
<tr>
<td>13–15</td>
<td>Conventional Shock Tube, Reverberant</td>
</tr>
<tr>
<td>16–18</td>
<td>Fast-Acting Valve (3.5″), Reverberant</td>
</tr>
<tr>
<td>19–20</td>
<td>260 Hz CF Narrow-Band Impact</td>
</tr>
<tr>
<td>21–23</td>
<td>775 Hz CF Narrow-Band Impact</td>
</tr>
<tr>
<td>24–27</td>
<td>1,025 Hz CF Narrow-Band Impact</td>
</tr>
<tr>
<td>28–30</td>
<td>1,350 Hz CF Narrow-Band Impact</td>
</tr>
<tr>
<td>31–34</td>
<td>2,450 Hz CF Narrow-Band Impact</td>
</tr>
<tr>
<td>35–38</td>
<td>3,550 Hz CF Narrow-Band Impact</td>
</tr>
<tr>
<td>39–40</td>
<td>2,075 Hz CF Narrow-Band Impact</td>
</tr>
<tr>
<td>41–42</td>
<td>146 dB Peak SPL and 138 dB Peak SPL</td>
</tr>
<tr>
<td>43, 45, and 47</td>
<td>29OC Driver, High Peak Wave, USAARL Report 86–7</td>
</tr>
<tr>
<td>44, 46, and 48</td>
<td>29OC Driver, Low Peak Wave, USAARL Report 86–7</td>
</tr>
<tr>
<td>49</td>
<td>29OC Driver, 131 dB Peak SPL, 100X, USAARL Report 85–3</td>
</tr>
<tr>
<td>50</td>
<td>U.S. Army Aeromedical Research Laboratory Conventional Shock Tube, Nonreverberant (Unpublished)</td>
</tr>
</tbody>
</table>

**TABLE III. Summary of Blast Overpressure Project Albuquerque Data**

<table>
<thead>
<tr>
<th>Study</th>
<th>Total Subjects</th>
<th>Total TTS Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 m Modified Muff</td>
<td>66</td>
<td>2,141</td>
</tr>
<tr>
<td>3 m Modified Muff</td>
<td>68</td>
<td>2,400</td>
</tr>
<tr>
<td>5 m Modified Muff</td>
<td>59</td>
<td>1,907</td>
</tr>
<tr>
<td>5 m Unmodified Muff</td>
<td>62</td>
<td>1,870</td>
</tr>
<tr>
<td>3 m Perforated Ear Plug</td>
<td>19</td>
<td>614</td>
</tr>
<tr>
<td>Total</td>
<td>274</td>
<td>8,932</td>
</tr>
</tbody>
</table>

**Shock Tube Noise Generation**

A small 2-inch diameter shock tube was used for generating the impulse noise needed for animal testing (Fig. 2). The shock tube consists of a high-pressure driver chamber at one end and an expansion section at the other. The high-pressure driver section is separated from the expansion section by a thin layer of diaphragm that self-ruptures at specific
pressure levels depending on the thickness. The driver section is pressurized by helium. Pressure waveforms were measured using impulse noise microphones (Fig. 3). A sample noise waveform data trace generated by the shock tube is shown in Figure 4.

Animal exposures were conducted inside a small noise-insulated test chamber. The shock tube was mounted inside the test chamber with the helium supply provided through a feed from a high-pressure tank outside of the test chamber. The data acquisition system for recording the sound pressure data was placed outside of the test chamber with connecting cables to the pressure sensors mounted inside the test chamber.

Animal Test Preparation
Before the initiation of the project, an application to perform the experiments was submitted to the UCSD Institutional Animal Care and Use Committee (IACUC) for approval. The UCSD IACUC initially was reluctant to approve the use of injectable anesthetics for sound exposure and recording, preferring gas anesthesia. However, recent publications have indicated that gas anesthesia can reduce sensitivity to noise-induced hearing loss, which has not been shown for injectable anesthetics.16 Once this was explained to the Committee, a compromise was reached in which injectable anesthesia would be used for blast exposure and ABR recordings on the first day, lasting 3 hours, followed by gas anesthesia on subsequent days. After approval by UCSD, the IACUC approval was sent to Joint Non-Lethal Weapons Directorate for final approval.

ABR Measurement
Stimulus delivery to the animal ear via a Tucker-Davis TDT ABR workstation (Tucker-Davis Technologies, Alachua, Florida) was calibrated and evaluated to ensure that it could generate the sound levels required to document threshold shifts of the desired magnitude without distortion or damage to the speaker. It was determined that threshold shifts up to the 60 to 70 dB range could readily be detected with the existing electrostatic speakers used for sound delivery, as determined by the range of initial thresholds of the animals and the maximum output of the speaker. Because body temperature is known to affect both sensitivity to noise and ABR thresholds, a thermal blanket was installed, with power controlled by a rectal thermistor with the ability to control body temperature with an error of <0.1°C. Previous work had shown that ABR thresholds reflect those measured with behavioral techniques, with a slight loss of sensitivity, and that threshold shifts measured with either method are equivalent.17

Before any blast exposures, a group of naïve chinchillas was tested in the ABR workstation to assess the suitability of stimulus delivery and recording equipment for this species. There was initial difficulty with using a closed system delivery of sound stimulation to the ear canal of the chinchilla as proposed. The external ear canal of this species has a sharp bend at the interface between the cartilaginous
portion and the bony portion. Insertion of a speculum into the canal with sufficient depth to produce a seal caused the ear canal to collapse, and the speculum was blocked by the canal wall. To resolve this issue, it was decided to change to an open-field sound delivery system, with the speaker positioned at a fixed distance (4 cm) above the ear canal, and the opposite ear blocked by a foam earplug. Thresholds determined with open-field sound delivery were much more stable than those obtained with closed field, with test–retest values showing variation of less than 5 dB. The use of a free-field configuration also greatly increased the speed at which animals could be tested after blast. On the basis of these pilot studies, the open-field procedure was used to assess the thresholds of both ears of each experimental chinchilla before and after blast. It was reasoned that, since all blast exposures were to be bilateral, the earplug would provide sufficient isolation of the test ear to obtain thresholds that could be reliably attributed to that ear (tests of the 2 ears after blast supported this reasoning). To test this, we assessed ABR thresholds from the side of the contralateral (plugged) ear and found wave I thresholds to be 40 to 50 dB above those for the stimulated (unplugged ear). No animals showed this degree of difference between the two ears after exposure, indicating the adequacy of this procedure.

For ABR recording, the active electrode was placed vertically to the ear canal of the tested ear, the reference electrode was placed on the vertex, and the ground electrode was placed on the leg. Each ear was tested separately. Needle electrodes (Tucker-Davis Technologies) designed for animal experimentation were inserted subcutaneously. The signal was filtered at 0.5 to 3 kHz and gated to the acoustic stimulus. The response to 512 stimulus presentations was averaged to generate the ABR waveform. The ABR stimulus consisted of tone bursts (25 msec in length, with a ramp time of 2.5 msec, presented at 20/sec). Thresholds were tested using a descending stimulus method, beginning at 90 dB SPL and descending in 5 dB steps until the waveform had clearly diminished into the background of the recording, after which one additional step was taken. Threshold was assigned as halfway between the lowest intensity, at which a waveform could be distinguished subjectively by two independent observers, and the next lower intensity. This method is commonly used for threshold determination in animal studies (Erkman et al, 1996; Housley et al, 2013).18,19

Thresholds of chinchillas before blast showed a median (interquartile range) value of 12.5 (7.5–17.5) dB SPL at 1.0 kHz, 2.5 (2.5–7.5) dB SPL at 2.0 kHz, and 2.5 (2.5–2.5) dB SPL at 4.0 kHz. Mean threshold (±SD) values were 15.5 (±6.2) dB SPL at 1.0 kHz, 6.3 (±5.7) dB SPL at 2.0 kHz, and 3.4 (±2.7) dB SPL at 4.0 kHz.

**Blast Exposure**

The animals were exposed at various distances to the shock tube to cover a range of intensity levels. To generate high TTS levels, it was found that there was a risk of tympanic membrane rupture. Therefore, multiple shots at low-intensity levels were used to generate high TTS instead of using a single high impulse. Occasionally, animals experienced bilateral tympanic membrane rupture, and in many cases one membrane was ruptured so that threshold data could only be collected from one ear. In all cases, animals with tympanic membrane rupture were discarded. The ears of each animal were treated independently. It was noted that the responses of the ears of an animal to blast could vary quite widely, presumably due to differences in blast exposure of the two ears. This variation exceeded what we might expect from individual differences (Ryan and Bone, 1978),20 which was the rationale for treating the ears separately.

Immediately following blast exposure, ABR threshold data were collected at 1, 2, and 4 kHz in each ear that exhibited an intact tympanic membrane, over the following schedule whenever possible as soon as the animal could be placed in the recording setup. For the first ear, this occurred 2 to 10 minutes after blast, and for the second ear, this occurred 15 to 30 minutes after blast. Threshold shift recovery tracking data were taken as follows:

—2 to 30 minutes after blast
—1, 2, and 3 hours after blast
—1, 3, 7, and 14 days after blast

Data from both ears collected from the new tests were pooled to develop the recovery model. TTS data up to 50 dB were collected from a total of 38 subjects with no tympanic membrane rupture, while a total of 98 animals were used. Animals were assigned to groups based upon the initial degree of threshold shift, averaged across the three test frequencies. Data were collected for groups of animals with an average TTS of 10 to 20 dB, 20 to 30 dB, 30 to 40 dB, and 40 to 50 dB. Very few animals were observed with threshold shifts exceeding 50 dB. Sample ABR recordings of threshold shifts before and after blast exposure are shown in Figure 5.

**RESULTS**

**Dose–Response Model (Auditory 4.0)**

The dose–response curves were developed using ordered logistic regression of SELA against the historical chinchilla TTS and PTS data grouped in 5-dB bins accounting for dose accumulation for multiple shots. Previous analysis of the BOP human volunteer data by Chan et al15 suggests that dose accumulation should follow a 3.44 log N rule for multi- shot exposures where N is the number of shots. Findings by Patterson also reported that dose accumulation tends to be weaker for lower N.21 On the basis of these findings and further comparison with the historical human rifle noise injury data,12 the dose accumulation rule was adopted as 3.44 log N for N ≤ 25 and 10 log N for N > 25. It is understood that the dose accumulation rule is only applicable to
impulses of equal intensity at constant presentation rates of about 1 minute but at least greater than 1 second apart similar to the way the data were collected. SELA is calculated over 1-second duration according to standard practice.

Hence, based on logistic regression analysis, the dose–response curve is expressed as the probability of injury ($p$) for each type and level of failure outcome (e.g., TTS$_2$ > 40 dB) as a correlation with the risk factor function $L$:

$$ P = \frac{\exp(L)}{1 + \exp(L)} $$

where $L = \alpha (\text{SELA} + 3.44 \log N) + \beta$ for $N \leq 25$

and

$$ L = \alpha (\text{SELA} + 3.44 \log(25) + 10 \log(N/25)) + \beta \text{ for } N > 25 $$

where $\alpha$ and $\beta$ are the regression coefficients obtained from statistical calculations. The correlations obtained from the chinchilla data were then shifted by 28 dBA representing scaling from chinchillas to humans. This 28-dBA shift was guided by best fitting the dose–response curves with the historical CHABA rifle noise human data as shown in Figure 6, and was found to be in close agreements with other findings that suggested about a 20-dBA threshold difference between humans and chinchillas.

The recovery model was constructed using the chinchilla data with guidance from the BOP human data. The BOP human data provided trend guidance as well as supplemental data for the low TTS range. The chinchilla recovery data were found to merge smoothly with the human data on the low end.

**BOP Human Recovery Data**

Findings from the BOP human recovery data trend are first given. The total number of subjects and TTS records derived is shown in Table III, with 274 subjects and 8,932 audiogram records. Analysis of the data shows the TTS recovery follows a log-linear time course even though the slope is different for different TTS levels. Figures 10 and 11 illustrate the log-linear TTS recovery trend from two test series. For all dose–response curves because of the use of ordered logistic regression with the pooled data.

The dose–response model shows favorable comparison with the CHABA data as shown in Figure 6. The data comparison plots show the mean data correlation and the 95% confidence intervals to indicate the range of data variability. Compared to the 10% and 25% injury group, the 95% confidence interval for the 50% injury group is much wider because of a smaller number of data points available for this higher injury rate (Fig. 6).

The 25-dB TTS dose–response curve is also found to agree well with the LAeq8 ≤ 85 dBA standard. Assuming equal impulses are fired at 1/sec over 8 hours, the present model predicts the 5% failure rate occurs at 86.9 dBA and the ±95% confidence interval is from 83.3 to 89.8 dBA. Therefore, the LAeq8 = 85 dBA criterion is well within the confidence interval for protection of 95% of the general population.
The TTS recovery is correlated for each TTS bin, which is a log-linear function with time in minutes as:

\[ TTS = a \ln \text{time} + b \]

The coefficient \( a \) represents the TTS recovery (decay) rate and \( b \) represents the TTS at 1 minute. Coefficients \( a \) and \( b \) are different for different TTS groups, and the ±1σ (SD) error bars for each data point are presented for each test series. As can be seen in the TTS recovery plots, the majority of the TTS injury outcomes are less than 10 dB, and the wide error bars for the highest TTS bin (>10 dB) are due to the small number of data samples available (Figs. 10 and 11). Furthermore, it can be observed that the recovery rate, coefficient \( a \), is stronger (more negative) at higher TTS level, with the highest TTS group showing the strongest TTS decay rate with time. Other research findings also support a log-linear TTS recovery behavior.5

When all the BOP data were combined together, it was found that the TTS recovery slope can be fitted as a single
polynomial function with TTS$_2$, as shown in Figure 12 and expressed as:

$$a = -0.0215 \ (\text{TTS}_2)^2 - 0.0122 \ (\text{TTS}_2)$$

Figure 12 shows that the magnitude of the TTS decay rate is stronger with higher TTS$_2$. Again, the wider data scatter at the higher TTS levels is due to smaller number of data samples available; nevertheless, the overall correlation coefficient, $R^2$, has a value of 0.926, which is considered good.

**Chinchilla Recovery Data and Model**

The newly collected chinchilla TTS recovery data are shown in Figure 13 as labeled by test series in the legend. The TTS data were binned for construction of the recovery model. Using a bi-log-linear approach with a change of slope at the 3-hour time point, the recovery slopes were determined statistically for each TTS group. For recovery within 3 hours, the recovery curves for four selected TTS groups are shown in Figure 14. The combined recovery slope model merging the chinchilla and human data is shown in Figure 15, which shows a smooth transition from the human (low TTS) to the chinchilla (high TTS) bins, with the recovery slope being essentially constant for TTS$_2 > 50$ dB. For recovery time beyond 3 hours, the recovery slope correlation is shown in Figure 16.

The complete recovery curves for all TTS groups are shown in Figure 17. As shown, the results include two kinds of extrapolations. The first is for TTS from 2 minutes back to 30 seconds after exposure. The second is for the recovery curves for TTS$_2$ exceeding 50 dB. Both extrapolations show reasonable results. Furthermore, a qualitative check of the recovery model was performed against the German human recovery data, which show a strong likelihood of full recovery within 24 hours if TTS$_2$ is less than 25 dB. On the basis of the recovery model constructed (Fig. 17), the time for full recovery for TTS$_2 = 20$ dB is about 1,386 minutes, which is about 19 hours, and for 25 dB, about 43 hours. Hence, this recovery range is in qualitative agreement with the German threshold of recovery within 24 hours for prevention of PTS based on human data, considering the limitation of the data points from the German study. The recovery model results shown in Figure 17 indicate TTS outcomes starting at 30 seconds after exposure based on extrapolations from the outcomes at 2 minutes. As shown for low TTS conditions (<25 dB), the TTS at 30 seconds can be from 2 to 3 dB higher than that at 2 minutes. For TTS from 25 to 50 dB, TTS at 30 seconds can be 3 to 5 dB higher than that at 2 minutes.

In summary, based on the model developed, the recovery window for 25-dB TTS$_2$ is within 43 hours. For 50-dB TTS$_2$, the subject will recover to about 27-dB TTS by 24 hours and will reach full recovery in about 38 days.

**DISCUSSION**

The new model, Auditory 4.0, has been developed to predict the full range of TTS and PTS as well as the recovery of TTS with time for exposures to impulse noise with no
FIGURE 13. Chinchilla temporary threshold shift (TTS) recovery data.

FIGURE 14. Chinchilla temporary threshold shift (TTS) recovery from 2 minutes to 3 hours.

Impulse Noise Injury Model
hearing protection. It is the only model known to date to provide these capabilities. Auditory 4.0 shows good comparison with historical rifle noise data for TTS reaching 50 dB (Fig. 6), and the recovery curves blend in well with the BOP human data (Fig. 12). If hearing protectors are worn, a method to calculate an equivalent free-field dose is needed although the established dose–response curves should remain valid.

For the TTS recovery model, a limitation was the inability to measure thresholds prior to 2 minutes after exposure. This was due to the time required to move each animal from the blast exposure arrangement to the recording arrangement, to insert the recording electrodes and position the stimulus delivery speaker, and especially to record ABR traces at several intensities. In the best case, this time was minimized to 2 to 3 minutes after blast for the first ear tested and 10 minutes for the second ear. Consequently, no data were collected from the first 2 minutes after blast. It should be noted that threshold shifts assessed over the first 60 minutes after blast showed only small changes for most of the subjects. The threshold response to bilateral blast is generally stable over the first hour after exposure, both for a single blast and for multiple, successive blasts. This is consistent with previous behavioral data on continuous noise exposure in the Mongolian gerbil. It is most likely that thresholds recorded 2 to 5 minutes after blast should not be too different from those that existed between time 0 and 2 minutes after blast; however, this remains an assumption.

It is important to note that the dose accumulation rule developed is limited to exposures to equal impulses at constant presentation rates of at least greater than 1 second apart. At present, there is no model for dose accumulation for a complex mixture of impulses from unequal intensity and irregular presentation rates. Study in this area is very limited because most of the previous work has focused on developing damage risk criteria for occupational training involving regular firings of equal impulses.

**ACKNOWLEDGMENTS**

This work was funded by JNLWD under contract no. N66001-11-C-0022.

**REFERENCES**


A Comparative Case Study of Risk, Resiliency, and Coping Among Injured National Guard

Lisa A. Gorman, PhD*; Angela J. Huebner, PhD*; Mara K. Hirschfeld, MS†; Sudha Sankar, MS‡; Adrian J. Blow, PhD‡; Danielle Guty, BS*; Michelle Kees, PhD§; Joel S. Ketner, MS‡

ABSTRACT An injury during deployment disrupts family and life functioning. The purpose of the present study was to provide an in-depth examination of three injured National Guard soldiers showing how differential experiences of navigating multiple systems to obtain treatment for injury resulted in different adjustment trajectories for these soldiers and their families. A comparative case study examined three families where a soldier’s injury was a central theme of family adjustment. Qualitative data were drawn from interviews conducted conjointly with both the soldier and spouse to provide an in-depth perspective of adjustment, meaning, and resource utilization patterns. In addition, survey data were collected at three time points in the deployment cycle (predeployment, 90 days post, and 1 year). These data were integrated into the case analysis, including mental health, marital relationship, treatment history, and characteristics of resilience. Study findings suggest that a delay in diagnosis, wait time for treatment, and the lack of comprehensive formal and financial support for a soldier following nonhostile injury lead to a pileup of stressors that are detrimental to the soldier’s physical and mental health, financial stability, and family well-being. Further study is needed to understand how these system level issues impede resilience among National Guard families.

INTRODUCTION

The purpose of the present study was to provide an in-depth examination of three physically injured National Guard (NG) soldiers, and to describe how the navigation of injury treatment contributes to soldier and family adaptation following a deployment to Afghanistan. Data were drawn from interviews conducted conjointly with both service members and their spouses. Survey data collected before deployment and at two additional times within the first year of reintegration illustrate different adjustment trajectories.

A self-reported injury by the service member predicts higher levels of post-traumatic stress, depressive symptoms, and parenting stress 45 to 90 days postdeployment.1 Among the combat injured, family disruption following injury was related to high child distress but the severity of the injury on its own was not.2 For the family, what happens during the reintegration phase of deployment can determine whether stress reactions are mitigated or exacerbated.3 Additional stressors,2,4 the availability of formal and informal supports, and meaning making are important factors in the reintegration process.4

Injuries incurred during deployment—combat or noncombat related—can add additional stress to the already complicated process of reintegration. Combat-related injuries may result in amputation, burns, severe soft tissue and orthopedic injury, and traumatic brain injury,5 whereas noncombat-related injuries tend to be fractures, inflammation/pain, and dislocation caused by sports/physical training, fall/jumps, and motor vehicle-related incidents.6 There is a growing body of evidence that suggest an injury increases the risk that the service member will also develop post-traumatic stress disorder (PTSD).5,7–9 Most of this research has focused on combat-related injuries while far less is known about the adjustment trajectory of service members returning with nonhostile injuries.

Given the fluidity through which NG soldiers move, between mobilization day (M-Day one weekend per month), active duty, and veteran status, their access to health care benefits can be complex.7 A “line of duty” (LOD) injury determination status states that those who incur or aggravate an injury, illness, or disease in the LOD are entitled to treatment10 at an approved military treatment facility and along with pay and allowances.11 If not already reported, a noncombat injury can be reported at the demobilization when the soldiers complete a battery of health screenings and questionnaires. Without an official LOD, the burden falls on the soldier to prove the injury was incurred during military service. Without this designation, receipt of benefits such as Veteran’s Administration (VA) health care, and disability compensation is also jeopardized. There are no known studies that examine the personal or family adjustment trajectory of both combat and nonhostile injured NG members’ in relation to navigation of systems during the reintegration process.

The present study employed a comparative case study methodology12 to explore the impact of differential experiences of system navigation on the adjustment trajectory of injured NG soldiers and their families. This study fills a gap in the literature by using qualitative data to expand the meaning construct of the family stress model and explain the influences of health systems on family resiliency processes. The
Resiliency Model of Family Stress, Adjustment and Adaptation\(^4\) served as a guide for assessment and interview questions. This model assumes a relational perspective of family adjustment with recursive effects such that overall family adaptation (X) is dependent on the interplay of deployment and injury severity (A), pileup of demands (AA), family resources including utilization of services (BB), and meaning or family perspective of their situation (CC) within the context of dealing with the injury.

**METHODS**

A comparative case study methodology was employed using cross-case comparison and within case analysis.\(^1\) This method allows for empirical inquiry and in-depth investigation of multiple sources and variables, which captures the complexity of real-life context of family and system interaction. Comparative case study intentionally selects a small number of cases that differ on outcome variable of interest. The small number of cases allow for a more in-depth probe of processes that may be related to the different outcomes. As employed in this study, the comparative case study approach allowed us to contribute to the limited literature specifically exploring the impact of deployment injury on family adaptation from the perspective of a service member and spouse. In this way, this method gives us the strongest means of drawing inference of cases for theory development.\(^1\) The study was approved by all partnering institutional review boards governing the use of human subjects.

**Participants and Procedures**

Data for the comparative case study were drawn from a larger ongoing mixed-method longitudinal study that followed a battalion of soldiers who deployed to Afghanistan. Soldiers and family who self-identified as resilient during their reintegration event could volunteer to participate in interviews in addition to completing survey data. Unique identification codes were used to match qualitative data with survey data. Because we were interested in family processes that predict resiliency, individuals with suicidal ideation and hazardous alcohol use were excluded from the interview pool. In-depth qualitative interviews were conducted with a target sample of 35 families representing demographics of the larger sample. Only couples in the qualitative interviews reporting an injury as a contributing factor to their reintegration process were eligible for inclusion in this comparative case analysis. We made every attempt to match the cases as closely as possible on variables that could also impact overall adjustment. Table I shows the comparison of cases with their cohort of injured (\(n = 77\)) and noninjured (\(n = 568\)) soldiers.

**Data Collection**

Surveys were collected approximately 90 days before deployment, at reintegration events 45 to 90 days after they returned home, and 1 year after reintegration. Surveys measured family adjustment using the Revised Dyadic Adjustment Scale\(^1\) and the Parental Stress Scale.\(^1\) To assess the psychological health of soldiers we used the PTDS checklist,\(^1\) the Patient Health Questionnaire,\(^1\) and the Generalized Anxiety Disorder 7-item scale.\(^1\) Pileup of demands were assessed using a 21-item checklist for life events occurring in the prior year. In addition to the in-depth interview, appraisal of their situation was measured using the Perceived Stress Scale-4\(^1\) and Satisfaction with Life Scale.\(^1\)

The in-depth family interviews were conducted 6 to 9 months postdeployment and averaged 90 minutes in length. Each interview was conducted by a two-person (male/female) team with one licensed therapist and an individual with military experience. In the semi-structured interviews, families responded to questions about family adjustment, supports that contributed significantly to their experience, and the family appraisal of their situation. Field notes of major themes and observations were created following the interviews, which were taped, transcribed, and reviewed by the interviewer for accuracy.

**Data Analysis**

Qualitative data were organized using Atlas.ti software (Scientific Software Development, Berlin, Germany).\(^1\) The coding team employed theoretical thematic analysis\(^2\) to identify patterns or interactions related to the constructs in our theoretical model. Consistent with theoretical thematic analysis, factors from the Resiliency model of Family Stress, Adjustment and Adaptation (i.e., family adaptation [X], deployment and injury severity [A], pileup of demands [AA], family resources [BB], and family meaning making [CC]) were used to guide initial coding. To this end, transcripts were initially coded independently and then codes (e.g., ABCX) and their application were compared, discussed, and consolidated into broader themes within each factor. Further the scored survey measures from pre, post, and 1 year follow-up were charted, mapped, interpreted, and incorporated into the analyses to explore the potential interaction between systems of support, family appraisal, pileup of demands, individual, and family outcomes.

**RESULTS**

Table I shows a comparison of outcomes for each case throughout the deployment cycle. A number of overall themes, concepts, and relationships emerged from the within-case analysis and cross-case comparisons. Factors contributing to a positive reintegration trajectory following service-related injury included prior deployment experience, timely medical and behavioral health treatment, financial stability in particular uninterrupted income through the community-based warrior transition unit (CBWTU), formal and informal supports from a community that understands their experiences, and personal grit of the spouse. In comparison, not having a LOD triggered a pileup of demands including a delay in VA health care treatment and disability compensation that exacerbated their problems leading to poorer family adjustment. Key factors of
the deployment and reintegration process were collected at 4 time point from multiple sources. The case comparisons of that data are illustrated in Figure 1 showing how injury intersects with other life-course events and how pathways to adjustment may be altered by system level barriers and supports. The trajectories are described in greater detail providing background information and quotes from the soldiers and spouses.

**Case 1: Mixed Adjustment Trajectory**

Prior deployment experience: reintegration from the first deployment was reported as difficult. According to the soldier, “When I came home from Iraq I put her through hell. I was drinking and doing other stuff and staying out late . . . I promised her when I came back from Afghanistan that I wouldn’t do that to her.” Both vowed to make the second deployment experience different (CC).

When soldier returned to Walter Reed for treatment, spouse was able to join him for the lengthy rehabilitation process. Supports (BB) were central to sustaining family. According to spouse, “I was just very fortunate with my job and the family and my parents took our dogs and somebody else took care of our house and somebody mowed our lawn and coordinated all of those services that you don’t really think about and take for granted.” They spoke positively about the support they received from nonprofits that donate to the wounded warriors. In addition, the commanding officer’s wife reached out to the spouse in support.

The couple also talked about their frustrations in navigating the formal medical system: “I don’t know exactly what we needed but I feel like a lot of the stuff we were left to do that to something and then they send you somewhere

### TABLE I. Cases Compared With Injured and Noninjured Cohort at Postdeployment (T2)

<table>
<thead>
<tr>
<th>Measurement Scores</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Injured Cohort T2 (n = 77)</th>
<th>Noninjured Cohort T2 (n = 568)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTSD (PCL&lt;sup&gt;a&lt;/sup&gt;)</td>
<td>53</td>
<td>47</td>
<td>23</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Depression (PHQ 9&lt;sup&gt;b&lt;/sup&gt;)</td>
<td>17</td>
<td>13</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Anxiety (GAD 7&lt;sup&gt;c&lt;/sup&gt;)</td>
<td>19</td>
<td>11</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**MH, mental health; T1, Time 1 survey completed before deployment; T2, Time 2 survey completed approximately 90 days following battalion demobilization, and T3, Time 3 survey completed approximately 1 year later; PCL, PTSD Checklist; PHQ, Patient Health Questionnaire; GAD, Generalized Anxiety Disorder; SWLS, Satisfaction with Life Scale; RDAS, Revised Dyadic Adjustment Scale; PSS4, Perceived Stress Scale-4; PSS, Parental Stress Scale. *Missing data. **PCL scores ≥50 is likely PTSD. 4PHQ 9 scores of 5, 10, 15, and 20 represent cut points for mild, moderate, and severe depression, respectively. 5GAD 7 scores of 5, 10, and 15 represent cut points for mild, moderate, and severe anxiety, respectively. 6SWLS scores 26 to 30 = satisfied, 21 to 25 = slightly satisfied, 5 to 9 = extremely dissatisfied. 7PSS4 higher scores indicate higher levels of parenting stress.**

---

**Case 2: Matched Adjustment Trajectory**

Prior deployment experience: both soldiers reported the transition to reintegration was challenging. According to the soldier, “I think I was in shock and didn’t really think about it but it was really hard to come back from the war and adjust to civilian life.” Both reported a desire to make the second deployment experience different (CC).
else ... we just kind of ended up giving up so they did offer programs, they did offer evening counseling sessions for couples ... but we didn’t really bother with a lot of just because of our experiences so far weren’t very helpful.”

Spouse credited her training as a mental health counselor in helping her cope. Both cited spirituality/religion as an important coping resource. Although the rehabilitation was described as difficult, the spouse was an advocate for her soldier, calming him, and keeping track of what needed to be done. According to soldier, “She [spouse] was my angel ...” The spouse also said, “I knew that my role in our relationship was to be the rock through this whole thing.”

With respect to overall family adaptation (X), results seem mixed. From a relationship perspective, the couple
assessments reflect high marital satisfaction and nondistressed adjustment postdeployment consistent with the in-depth interview. The spouse said, “I think it (second deployment) definitely made our marriage stronger not weaker and we really found out some things about each other in the midst of it all.” One year later, the soldier reports less marital satisfaction and more distress compared to the spouse. Though his symptom level of depression, anxiety, and PTSD improved over time, the soldier continues to struggle: “The thing I deal with the most is the TBI just because my memory, my irritation and my anger and what not . . . I have some anxiety pills . . . which help a lot.” The spouse also said, “He had no history of anxiety, depression or any other kind of mental health [issue] prior to this. I have known him for a long time and it was like a switch that was thrown because now he has anxiety.”

Case 2: Positive Adjustment Trajectory

Soldier said that he was injured (nonhostile) during his first deployment but did not report it because he was eager to return home to his family. He assumed he would be able to access treatment but ran into considerable difficulty: (regarding the first deployment injury) “… I am just going to let the VA take care of this when I get there. And as it turns out, the wheels of justice turn very slowly at the VA so in the year and a half that I was home between the deployments, I managed to get an MRI and some physical therapy. I never even got to talk to a surgeon.” (A) Because of his previous experience, soldier completed an official LOD injury report, stating: “I had made up my mind overseas that I was not coming off active duty orders until I was fixed—even if I had to stay . . .”

Medical treatment (BB: formal support) extended his deployment for 2 months, bringing his total time away from family to 14 months. The spouse and kids had phone access and traveled to visit on weekends diminishing some informal support (BB—familial). The spouse admitted not utilizing formal supports (BB) because meetings and events were too far away for her to get to. She reported informal support (BB) from her family, most of whom live in the same neighborhood and have prior military service.

The couple noted the difficulty of separation, but they also shared how it helped ease the transition back into family life. A unit buddy with the soldier during the rehabilitation process was an important source of informal support (BB). According to soldier, “I mean as sucky as that was not to be able to come home, it was probably really good as well because it gave me time to adjust from the daily life in Afghanistan to be more civilized. . . . one of the guys I deployed with was there with me [in hospital] and we would go out and see movies and go out to dinner so it gave me that decompression time that I didn’t have the first go round.” When asked about accessing military benefits after this deployment, soldier responded, “They have been spot on with them . . . as far as benefits, they have been very good. I haven’t missed a paycheck so I am still on Title 10 order.”

The couple seem to share an outlook on life and service that connects them (CC). In commenting on his future job prospects, soldier said, “There has to be somewhere for somebody with my skills to do something that makes a difference and that is the big thing to me . . . . I don’t have to change the world but I want to do something that makes a difference.” Spouse reflected, “What is important to me is change so I don’t look at things so much as obstacles, I look at is as being willing to adapt to what is going on and accept that other things can be just as important . . . Look at what is important to you today . . . . That is how I live every day.”

Overall family adjustment seemed positive. The couple talked about having learned from the first deployment how to reintegrate more successfully. Spouse talked about being less timid in her communication, more direct and firm. According to soldier, “I feel better now than I did before the first deployment . . . So for whatever reason, this deployment was really good for my marriage . . .” Both the soldier and the spouse assessed on the dyadic adjustment scale show significant improvement from pre to postdeployment.

Case 3: Poor Adjustment Trajectory

The soldier did not complete a LOD at demobilization but offered no explanation for why he did not do this. At the time of the interview the injury had not healed and he was on pain medication. Following deployment, the soldier went back to former employer but injury interfered with ability to continue in position. He took a part-time position for less pay and subsequently experienced a pileup of demands including loss of health care insurance and other financial stress (AA).

In terms of resources (BB), the spouse noted that other formal supports like the Armory’s Family Assistance Center were very helpful in providing rent money when the couple was struggling and their children were able to get health care through a government subsidized program (BB). The soldier said he was receiving disability benefits through his civilian employer while he waits for VA disability. His frustration is evident: “. . . it is the VA itself that sucks. They take forever to do anything . . . We applied [to the VA for benefits] in December so we’re on month four of the waitlist which is like 16–18 months . . . That is to find out if you have been denied or approved for it. And then if you are denied you can appeal and you put your appeal in and it takes another 16–18 months.”

Spouse elaborated her concerns about the level of support from the VA: “It would definitely help if the VA wasn’t so slow at doing things and they could actually get the records [of soldier’s service] . . . instead of just prescribing narcotics all the time . . . . He is going to end up in a rehab facility for being addicted to them if you just keep prescribing more and more on top of one another . . . .”
When asked how soldier was coping, the spouse said, “the VA not helping him is really getting to him. That is when his PTSD really kicks in and he gets so frustrated and so anxiety ridden over not being able to provide for his family that it is just irritating him and that doesn’t help at all.” The family narrative is consistent with the increase in PTSD symptoms from early postdeployment to 1 year later. When asked how they were functioning as a couple, spouse said: “We have our moments and we tend to argue, but I don’t know how to explain it. Especially now. The biggest thing is his PTSD. Now I see the changes—he doesn’t necessarily see the changes but I definitely do. His mother does … I think a lot of people don’t understand why he is the way he is now because they don’t know.” The soldier also noted changes in his interactions with others. The couple’s assessments suggest that the spouse experienced a decline in life and relationship satisfaction earlier than the soldier and by the 1 year postdeployment survey the couple were going through a divorce.

DISCUSSION
Over the course of the study, the couple representing a nonhostile injury (Case 2) receiving treatment and compensation through the CBWTU showed the most resilience across all domains including dyadic adjustment, parenting stress, and life satisfaction. This couple had the benefits that come with older age, higher income, rank, more years in their relationship, and older children. He also had the benefit of a previous service-related injury where he learned the value of a LOD for receiving care through the CBWTU that integrated primary care, mental health, and social services intended to reduce barriers.22 Like 30% of veterans receiving VA medical care in the Sayer study,23 Case 3 experienced marital conflict and anger control problems following deployment. Lower family income/resources, no prior deployment experience, young children, and intersection with life-course events may be confounding issues and opportunities for targeted intervention. The ability to access health care and disability benefits in a timely manner seemed to be critical junctures in the reintegration process and the additive stressors further complicated family finances and marital strain leading to marital separation, as well as increased symptoms of anxiety and PTSD. Both cases of a nonhostile injury shed light on the unique challenges NG families face navigating systems of care without a LOD. Though the Case 2 couple faced delay in treatment following the soldier’s first deployment, the spouse’s income could support the family and likely buffered some of the financial stress as well as access issues associated with injury treatment.

A deployment-related injury is an unexpected event in the life course of a soldier, yet the detrimental psychological and financial affect seemed ameliorated by formal and informal supports. Though Case 1 experienced a combat injury of greater severity, the formal and informal supports seemed to buffer the effect on family outcomes and well-being. Case 2 had experienced CBWTU during reintegration from his most recent deployment and VA during reintegration from a previous deployment. His experiences were stark in contrast and illustrate a challenge for NG early in the reintegration cycle that is not faced by their active duty counterparts who have uninterrupted pay and access to health care at military treatment facility. Severity of the injury with extended treatment and chronic symptoms affects the trajectory of soldier and family well-being. In addition, a delay in diagnosis, wait time for treatment, lack of comprehensive formal, and financial support may be associated with a pileup of demands and need further investigation. This comparative case study suggests that families with a greater pileup of demands exhibit poorer health and family outcomes.

Of note in this comparative study is that each case is different. This is in contrast to programs and services offered to military personnel that may treat all individuals the same. Each case in our study had a married soldier, who experienced a war time physical injury. Each soldier had a spouse as a part of the deployment. However, each family had a different trajectory postinjury that was dissimilar. Some of these changes can be ascribed to individual characteristics of the soldier, others to military and civilian supports and resources. Although others to pre-existing marital dynamics, and the ability of the couple to work through the event together. What stands out among the case comparisons is how different each trajectory looks, and how maximization of supports and minimizations of frustrations and barriers to recovery can ameliorate the pileup of stressful events.

The in-depth case comparison was limited to three families from a Midwest NG unit, which limits the generalizations to a narrow sample of NG families contingent on region of the country and barriers to access health and social services within that region.12 In addition, we acknowledge that we were particularly interested in factors associated with navigation of injury. There may be other factors not captured in our study that also contribute to difference in adjustment. Despite these noted limitations, the comparative case analysis begins to provide insight into some of the reintegration challenges and complex interaction effects unique to NG families of injured soldiers. The deeper investigation of three cases within the constructs of the Resilience model illustrates the additive effects of multiple stressors. The comparative case study may serve as a way to identify potential causal variables to focus on in future research and larger quantitative studies of injury trajectory.

CONCLUSION
This study increases our understanding of risk, resilience, and coping among NG families when a soldier is injured during deployment. Study findings regarding intersection of normative life events and trajectory of soldier and family well-being are consistent with other conceptual models.53 This study builds on the qualitative study of New York veterans that found the systems of care that serves them is complicated and
difficult to navigate. This study sheds light on the family’s perceptions of services and how a delay in diagnosis, wait time for treatment, lack of comprehensive formal, and financial support following the soldier injury interacts with process of risk and resilience as families tackle subsequent pileup of stressors. The additive effects of multiple stressors and barriers point toward poorer soldier and family adjustment within the first year of reintegration and greater life-course disruption. Young soldiers, first-time deployers, and spouses may benefit from education regarding the necessity of LOD and remaining on active duty military status for nonhostile injuries. This study raises significant concern about an unknown number of veterans who do not meet the VA priority ranking to receive services and are now spiraling toward mental and financial instability as well as family disruption and crisis. Further study is needed to understand how system level issues, such as wait time for treatment of nonhostile injuries, may impede resilience. Immediate actions could do much to ameliorate risk and build resilience and coping strategies among injured veterans and their families.

ACKNOWLEDGMENT
This work was supported by the Office of the Assistant Secretary of Defense for Health Affairs through the Psychological Health/Traumatic Brain Injury Research Program Award No. W81XWH-12-1-0419 and W81XWH-12-1-0419.

REFERENCES
Physical Training Outcome Predictions With Biomechanics, Part I: Army Physical Fitness Test Modeling

Bryant L. Sih, PhD; Charles H. Negus, PhD

ABSTRACT  Objectives: The U.S. Army Basic Combat Training (BCT) is the first step in preparing soldier trainees for the physical demands of the military. Unfortunately, a substantial number of trainees fail BCT due to failure on the final Army Physical Fitness Test (also known as the “end of cycle” APFT). Current epidemiological studies have used statistics to identify several risk factors for poor APFT performance, but these studies have had limited utility for guiding regimen design to maximize APFT outcome. This is because such studies focus on intrinsic risks to APFT failure and do not utilize detailed BCT activity data to build models which offer guidance for optimizing the training regimen to improve graduation rates. Methods: In this study, a phenomenological run performance model that accounts for physiological changes and fatigue due to training was applied to recruits undergoing U.S. Army BCT using high resolution (minute-by-minute) activity data. Results: The phenomenological model was better at predicting both the final as well as intermediate APFTs ($R^2$ range = 0.55–0.59) compared to linear regression models (LRMs) that used the same intrinsic input variables ($R^2$ range = 0.36–0.50). Conclusions: Unlike a statistical approach, a phenomenological model accounts for physiological changes and, therefore, has the potential to not only identify trainees at risk of failing BCT on novel training regimens, but offer guidance to regimen planners on how to change the regimen for maximizing physical performance. This paper is Part I of a 2-part series on physical training outcome predictions.

INTRODUCTION

The U.S. military is under pressure to reduce costs and personnel while maintaining capabilities and man power.1,2 Military readiness includes having properly trained and fit personnel to meet operational physical requirements. Initial military training programs, such as the U.S. Army’s Basic Combat Training (BCT) prepare a civilian recruit for the rigors of military life. Standardized physical fitness tests are used to assess both BCT baseline and final fitness. Unfortunately, some BCT trainees are unable to reach the necessary conditioning level to pass the final physical fitness test, which results in further training or attrition. For the U.S. Army Physical Fitness Test (APFT), this failure rate has hovered at about 2 to 3% for more than a decade3 although this percentage may be rising as incoming trainees have had progressively higher weight and body mass index.4 Low-fit trainees are also more likely to be injured,5,6 potentially affecting their quality of life as well as causing additional medical and financial strain on military resources.7

To address this problem, a few epidemiological (i.e., statistical) studies in the peer-reviewed literature that have identified intrinsic characteristics of trainees who fail to reach the necessary conditioning to pass the final BCT physical fitness test. Most studies appear in government technical reports which consistently have found risks that include anthropometric traits such as older age, female gender, being overweight, and high body mass index4,8,9 as well as health history measures (low initial fitness6 and prior smoking, prior injury, and sedentary activity history9). However, epidemiological studies identify risk factors without consideration of either the physiological mechanisms or the training regimen involved and so, while important, are merely correlative and so are only applicable to the training regimen from which they were derived. Thus, they do not offer insight in how physical training in BCT can be improved to optimize outcomes for high-attrition risk trainees.

Attrition due to an inadequate score on the end of cycle APFT or other fitness tests is not solely due to intrinsic risk factors. The BCT training regimen is designed to improve strength and endurance while minimizing injury. Those studies which included the extrinsic risk factors did not, surprisingly, confirm the influence of regimen design on fitness outcomes.13 Due to the difficulty of collecting training data in a military environment, these studies were limited in physical activity detail. Nevertheless, changes to the training regimen have been correlated with outcomes. For example, marching distance has been seen to affect trainee fitness14 and remedial training regimens for those deemed too unfit to begin BCT have shown promising, though inconsistent, results when implemented.10,11,15 None of these studies, however, were able to incorporate exercise details such as intensity and duration in a physiologically meaningful way, limiting their ability to make predictions on performance gains from novel training regimens.

The purpose of this study was to adapt a validated physiologically based phenomenological model of fitness evolution...
in runners to a military environment that accounts for both intrinsic trainee traits (i.e., anthropometry) and activity regimen. We hypothesize that phenomenological fitness prediction modeling that captures the complex and nonlinear biomechanical and physiological mechanisms that lead to fitness changes as well as detailed activity regimen will be better than purely statistical models in predicting fitness outcome for novel training regimens. A phenomenological model will additionally require less data for parameter calibration. Running performance—specifically, the 2-mile run time required by the U.S. Army APFT—was chosen as the fitness metric since running is widely utilized in all branches of the military. In Part II of this series, we use biomechanical and physiological models to estimate overuse injury risk using detailed activity data and trainee traits.  

METHODS

The phenomenological model chosen for adoption to the U.S. Army BCT training cohorts was originally developed by Banister et al., has undergone previous validation and peer review, and has shown promise in predicting mildly fatigued performance level during BCT and the final 2-mile APFT run time. (A description of the data collected on the U.S. Army BCT trainees follows model development.) The “Banister” model is based on the effect of training on two competing components to performance: the positive effect of increased fitness and the negative effect of fatigue. Predicted performance level with time \( P(t) \) is calculated by the difference between the two components:

\[
P(t) = P^* + k_1 g(t) - k_2 h(t)
\]

where \( P^* \) is the initial performance level, \( g(t) \) is the fitness component, \( h(t) \) is the fatigue component, and \( k_1 \) and \( k_2 \) are dimensionless weighting factors. Both \( g(t) \) and \( h(t) \) are recurrently affected by training load \( w_i \):

\[
g(t) = \sum_{j=0}^{t-1} e^{-(t-j)/\tau_1} w_i
\]

\[
h(t) = \sum_{j=0}^{t-1} e^{-(t-j)/\tau_2} w_i
\]

where \( i \) is a given day of training, and \( \tau_1 \) and \( \tau_2 \) are decay time constants. Thus, the change in both fitness and fatigue depend on the training \( w_i \) and how long ago previous training was done (via \( \tau_1 \) and \( \tau_2 \)). The values of \( k_1, k_2, \tau_1 \), and \( \tau_2 \) are chosen such that \( k_1 < k_2 \) and \( \tau_1 > \tau_2 \). In this configuration, a daily bout of exercise \( w_i \) results in an increase in both fitness \( g(t) \) and fatigue \( h(t) \), with fatigue dominating in the short term due to a larger \( k_2 \). As time passes, the effect of fatigue diminishes faster than fitness due to the differences in time-constant values, and overall performance increases until the effect of fitness has also become negligible.

The cumulative, time-decaying effect of all previous exercise bouts result in the current performance level \( P(t) \).

Because it is impractical to directly measure oxygen consumption in a nonlaboratory setting, Morton et al. used a heart rate (HR) ratio along with duration to quantify \( w_i \):

\[
w_i = \sum D \left( \frac{HR_{ex} - HR_{rest}}{HR_{max} - HR_{rest}} \right) e^{-\left(\frac{HR_{max} - HR_{rest}}{HR_{max} - HR_{rest}}\right)}
\]

where \( D \) is the duration of the exercise (minute), \( HR_{ex} \) is the average HR during exercise, \( HR_{rest} \) is the resting HR, \( HR_{max} \) is the maximal HR, and \( w_i \) is the sum of all exercises during that day. The exponent term weights \( w_i \) for high intensity training, and \( e \) varies for males (1.92) and females (1.67), which are values based on blood lactate concentration differences. Note that the units of \( w_i \) are “weighted” minutes and are arbitrarily defined as training impulse or “trimp” units. Thus, \( g(t), h(t), P^*, \) and \( P(t) \) are also in units of trimp. A higher trimp value indicates a higher level of fitness.

In order to convert a predicted performance level \( P \) at time instance \( t \) into an actual performance result such as run time, Morton et al. proposed a formulation based on the tendencies of world track records to improve exponentially:

\[
y = L + ae^{-P(t)/b} \quad \text{or} \quad P(t) = b \ln \left( \frac{a}{y - L} \right)
\]

where \( y \) is the actual run time, \( L \) is an ultimate limit, \( a \) is an amplitude parameter, and \( b \) is a time parameter. Values for \( L, a, \) and \( b \) depend on the performance exercise (e.g., 2-mile run in the case of U.S. Army BCT). \( L \) is set as the theoretical best human effort. The parameters \( a \) and \( b \) are derived by arbitrarily assigning values of \( P(t) = 1,000 \) for a world record performance and \( P(t) = 0 \) for the lowest possible performance by an able-bodied person and solving for \( a \) and \( b \).

Equation (5), along with actual run time measures, is used to determine the parameters \( k_1, k_2, \tau_1, \) and \( \tau_2 \) through optimization so that \( w_i \) scales appropriately for the given run task. As mentioned above, this model was able to predict multiple run time performances for two recreational level runners over a 90-day training and cessation protocol \(( R^2 = 0.71 \text{ and } 0.96) \). Additional details on model derivation and parameter optimization can be found in Morton et al.

For this study, since neither HR nor oxygen consumption were directly measured on the BCT trainees, regression-based estimates of resting \(( VO_{2rest} \)) maximal \(( VO_{2max} \)), and exercise \(( VO_{2ex} \)) oxygen consumption were used as replacements for HR in Equation (4) to calculate individualized \( w_i \). For \( VO_{2rest} \), the Harris–Benedict equations were used:

\[
VO_{2rest}^{\text{male}} = 66.473 + (13.751 \times w_t) + (5.0033 \times h_t) - (6.755 \times a)
\]

\[
VO_{2rest}^{\text{female}} = 655.0955 + (9.463 \times w_t) + (1.8496 \times h_t) - (4.6756 \times a)
\]
where VO₂rest is in kcal/day, ht is height (m), wt is weight (kg), and age is in years. This was converted to L/min assuming 4.825 kcal/L of O₂. Maximal oxygen consumption was estimated from Daniel’s Formulation, which uses the time to complete a run distance at maximal effort. In the formulation, oxygen consumption in mL/kg/min is estimated from:

\[ VO_2 = -4.60 + 0.182258\times + 0.000104v^2 \]  

where \( v \) is run velocity (m/min) and the corresponding percent of VO₂max is estimated from:

\[ \%VO_{2\max} = 0.8 + 0.1894393e^{-0.012778} + 0.1989558e^{-0.1932605} \]  

using the time \( t \) (minute) to complete the run. VO₂max is thus calculated as \( VO_2/\\%VO_{2\max} \). We assumed that each trainee’s initial 1-mile run was maximal and used Daniel’s Formulation along with the recorded run time and body mass to estimate their VO₂max in L/min.

VO₂ex (L/min) was estimated based on PAtracker acquired intensity levels for running (described below) along with total weight (body mass plus any reported extra weight) using the gender-based VO₂ estimates shown in Table I.

In order to apply the model to U.S. Army BCT, \( P^*, k_1, k_2, \tau_1, \) and \( \tau_2 \) of Equations (1) through (3) need to be specified. The initial performance level \( P^* \) was estimated from a recruit’s 1-mile run time \( t_{\text{run}} \) (minute), scaled up to the APFT standard 2-mile run using a validated regression equation derived by Riegel and converted to trim units using Equation (5). We assume that \( k_1, k_2, \tau_1, \) and \( \tau_2 \) characterize the average response of the underlying physiological mechanisms of human performance and are the same for all individuals (i.e., constants) regardless of gender or fitness level and that it is the workload \( w_i \) that differs between trainees. The time constants \( \tau_1 \) and \( \tau_2 \) are assumed to be 45 and 11 days, respectively, which are the average values found by Morton et al for the two casual runners. In addition, we modified \( P(t) \) with a “laziness” factor \( f \) to account for trainees who did not give a maximal APFT run effort if they were sufficiently fit to pass the test easily. The “laziness” factor reduces the portion of a recruit’s predicted performance level \( P(t) \) that is greater than that which is needed to pass the APFT by a fixed percentage. This factor, along with \( k_1 \) and \( k_2 \) of the Banister model were optimized using the Groupcal data subset described below. Because a few trainees had marked increases or decreases in performance compared to their initial starting condition, a biweight rather than standard least squares fit was used for optimization, reducing the influence these outliers had on the model fit.

Since gender differences are accounted for via VO₂rest, VO₂ex, and variable \( c \), males and females were combined in the optimization process so the \( k_1, k_2, \) and \( f \) parameters can be applied to both genders for prediction. Also, the \( L, a, \) and \( b \) values used to convert \( P(t) \) for the prediction of U.S. Army APFT 2-mile run times can be found in Table II.

Linear regression models (LRMs) were also created to test the hypothesis that a statistical approach will have a greater drop in run time prediction accuracy than the Banister model on novel datasets. Since the Bannister model inputs were composed of initial run time \( t_{\text{run}} \) (minute) as well as trainee age (years), weight (kg), height (m), and gender (to estimate VO₂rest), stepwise linear regression was used to identify which of these inputs were statistically significant (\( p < 0.05 \)) and create LRMstep. A plot of residuals was used to visually identify outliers, which were removed to maintain a normal distribution of residuals and separate male and female models were developed to maintain homogeneity of variance. Homogeneity and a linear relationship between the response and linear predictor were verified visually (residual vs. fitted plot). In addition, because a stepwise linear regression would likely identify and discard nonsignificant inputs, over-fitted LRMsteps with all inputs (\( t_{\text{run}}, \text{age}, \text{weight}, \text{and height} \)) for each gender was also created (LRMfull) since all inputs were used in the Banister model.

Data were collected on 563 trainees undergoing 9 weeks of U.S. Army BCT at Fort Jackson, South Carolina, in 2010. Measures included basic anthropometry (height, weight, age, and sex), times to complete an initial 1-mile run \( t_{\text{run}} \) and three APFT 2-mile runs (two intermediate and one final). The initial 1-mile run was completed prior to training; the 2-mile runs were completed around days 12, 32, and 48. Trainees were required to complete the final APFT 2-mile run in a minimum time that is based on age and gender.

After removing subjects with incomplete baseline information, data from 347 males (20.7 ± 3.8 years; 77.9 ± 12.9 kg, 176 ± 7 cm) and 108 females (20.5 ± 3.5 years; 62.0 ± 8.5 kg, 162 ± 7 cm) were used for this study. The initial APFT runtimes mean ± SD were virtually identical between those included in this study and those that were not, supporting our assumption that the probability of

<table>
<thead>
<tr>
<th>Intensity Level</th>
<th>Male (mL/kg/min)</th>
<th>Female (mL/kg/min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light(10–30)</td>
<td>19.1</td>
<td>16.2</td>
</tr>
<tr>
<td>Moderate(40–60)</td>
<td>26.8</td>
<td>23.5</td>
</tr>
<tr>
<td>High(70–100)</td>
<td>34.5</td>
<td>30.8</td>
</tr>
</tbody>
</table>

The values are based on mid-range values of different levels of physical activity reported in McArdle et al.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>10 Minutes</td>
</tr>
<tr>
<td>a</td>
<td>21 Minutes</td>
</tr>
<tr>
<td>b</td>
<td>697 Trimp</td>
</tr>
</tbody>
</table>
missing data was not related to the value of the variable. Of the 455 trainees, 64 (45 males and 19 females) had all intermediate and final APFT run times recorded. These were selected for Banister, LRM\textsubscript{step}, and LRM\textsubscript{full} model development (the “calibration group”: Group\textsubscript{cal}) since the Banister model calibration described above is based on changes in run fitness during an entire training regimen. The remaining 391 with one or more missing APFT run times had a dual use in testing model predictability: final APFT runtimes were used for validation (the “validation group”: Group\textsubscript{val}), and the intermediate (nonfinal) APFT results served as our novel training regimens (Group\textsubscript{nov}) for all three models since intermediate APFTs occurred at a substantially earlier time in the training regimen. Mean ± SD of the model inputs (height, weight, age, gender and, importantly, initial APFT runtimes) were nearly identical between Groups. Recruit data were provided by the U.S. Army Research Institute of Environmental Medicine (USARIEM) and was collected in accordance with their institutional review board as well as the protection of human subjects (Army Regulation 70-25: Use of Volunteers as Subjects of Research. Washington, DC: US Department of the Army; January 25, 1990; US National Archives and Records Administration. Code of Federal Regulations. Title 45. Public Welfare. 2009).

To quantify the U.S. Army BCT regimen needed to calculate each trainee’s training load \( w_i \) of Equation (4), USARIEM observers followed the trainees in six Companies, recording group activity parameters including activity type, duration, intensity, and load carried for the entire 9 weeks of BCT on a minute-by-minute basis using an Android smartphone app (PAtracker), a custom in-house application developed for the purpose.\(^{24}\) All observers were trained on the app before data collection began. Each Company had an average of 75.8 ± 17.7 trainees. The PAtracker recorded activity type (stationary, menial tasks, walk, calisthenics, cadence march, combatives, run, obstacles/climbing, crawl, and lift/carry), activity intensity (light, moderate, and high), and activity time duration. For this study, it was assumed that a Company’s PAtracker duration and intensity for walking, running, calisthenics, and other activities were representative of all the trainees in that Company. However, because \( w_i \) is based on VO\textsubscript{2rest}, VO\textsubscript{2max}, and VO\textsubscript{2ex}, which are also dependent on trainee gender, anthropometry and initial run time (i.e., initial fitness), the same PAtracker intensity results in individualized \( w_i \)’s. For example, we define a “fit” individual as one whose initial runtime is in the top 10%, an “unfit” person as having a runtime in the lowest 10% and “average” as those in the middle 10%. Thus, two males of the same age, height, and weight with very different initial runtimes (i.e., a “fit” and “unfit” individual) undergoing the same training regimen will have the same VO\textsubscript{2rest} and VO\textsubscript{2ex} based on Equation (6) and Table I, but the “unfit” individual will have a lower VO\textsubscript{2max} estimate (Equations 7–8). This results in the ratio of (VO\textsubscript{2ex}−VO\textsubscript{2rest})/(VO\textsubscript{2max}−VO\textsubscript{2rest}) used in Equation (4) and the correspond-

**RESULTS**

Changes in performance level \( P \) with time during BCT show the contribution of different characteristics of the Banister model (Fig. 1). A decrease in \( P \) is seen during approximately the first 14 days of training due to the fatigue component \( h(t) \), which has a larger weighing factor \( k_2 \) but shorter time constant \( \tau_2 \) than the fitness component. This phenomenon also results in the “jagged” \( P(t) \) profile—a large exercise bout results in a rapid decline in \( P(t) \) due to fatigue but an overall increase in fitness over the following weeks. Also, to more clearly illustrate how training has affected individuals differently through the calculation of training load \( w_i \), and how the same \( k_1, k_2, \) and \( f \) values results in a range of individualized predictions, the predicted run time of three representative trainees from the same Company (approximately “fit”, “unfit”, and “average” based on their initial 1-mile run time) are shown in Figure 1.

Importantly, trainees with low initial fitness are predicted to have relatively large drops in run time for a given amount of \( w_p \), whereas highly fit trainees see little predicted change and, in some cases, a loss of performance because the regimen is insufficient to maintain their fitness level (Fig. 2).
Optimization of Banister model parameters using Groupcal yielded the $k_1$, $k_2$, and $f$ values shown in Table III ($R^2 = 0.57$). Applying the model to Groupcal resulted in similar $R^2$'s of 0.55 when predicting final APFT only and 0.59 when predicting Groupnov intermediate APFTs' recorded during BCT (Fig. 3).

The $LRM_{\text{step}}$ derived from Groupcal for each gender was found to be:

$$
\text{Male: } 30.429 - (0.89964 \times t_{\text{run}}) - (0.24409 \times wt) \\
+ 0.014334(t_{\text{run}} \times wt) \\
(9)
$$

Since the Banister model accounts for gender through VO2 estimates, the predictions from $LRM_{\text{step}}$ (and $LRM_{\text{full}}$) for males and females were combined prior to the calculation of $R^2$. For $LRM_{\text{step}}$, Groupcal $R^2$ was 0.69. When validated against Groupval for final APFT run time, $R^2 = 0.47$ and for all intermediate APFT runtimes (Groupnov), $R^2 = 0.36$. See Figure 4.

The $LRM_{\text{full}}$ that used all the same inputs and Groupcal data set as the Banister model was found to be:

$$
\text{Male: } 11.017 + (0.18249 \times t_{\text{run}}) + (0.015038 \times age) \\
+ (0.010432 \times wt) - (0.27597 \times ht) \\
(10)
$$

Gender-combined $R^2$ values for $LRM_{\text{full}}$ applied to the Groupcal, Groupcal, and Groupnov datasets were 0.74, 0.50, and 0.40, respectively (Fig. 5).

In both Groupcal and Groupnov validation cases, the Banister model, which includes individualized training

**TABLE III.** Optimized Parameter Values for Equation (1) and the “Lazy” Factor $f$ Found by Using a Biweight Least Squares Fit of 2-Mile Run Time Performance for Groupcal. Average Parameter Values for Equation (1) Reported by Morton et al\textsuperscript{17} Are Provided for Comparison. All Parameters are Unitless

<table>
<thead>
<tr>
<th></th>
<th>$k_1$</th>
<th>$k_2$</th>
<th>$f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Model</td>
<td>0.99</td>
<td>1.91</td>
<td>0.65</td>
</tr>
<tr>
<td>Morton et al\textsuperscript{18}</td>
<td>1.0</td>
<td>1.9</td>
<td></td>
</tr>
</tbody>
</table>

figure also highlights the expected observation that unfit trainees had higher variability, with large changes in run time performance between APFTs since these trainees have more room for improvement.

**FIGURE 2.** Model predicted and actual 2-mile run times vs. training day of example representative unfit, average, and fit trainees undergoing the same Basic Combat Training regimen.

**FIGURE 3.** Predicted vs. actual APFT run time for the “Banister” model. Left are Groupcal trainees used to optimize the model parameters. Middle are the results when the model was applied to Groupcal trainees' final APFT. Right is when the model was applied to intermediate APFTs of Groupnov with 3 outliers not shown (but used to determine $R^2$). “—” is the identity line (i.e., a perfect model prediction). Note consistent accuracy across both calibration and validation analyses, but less accuracy in predicting trainees with slower run times.
regimen workload (\( w_i \)), performed better than the LRM\(_{\text{step}} \) and LRM\(_{\text{full}} \).

**DISCUSSION**

These results suggest that by accounting for training effects, the model originally proposed by Banister et al\(^{17} \) and adapted to U.S. Army BCT, is able to more accurately predict final APFT run time than statistical models that use initial fitness and basic anthropometry measures. The statistical models (LRM\(_{\text{step}} \) and LRM\(_{\text{full}} \)) are very good at predicting APFT run time from the data from which they were derived but have a substantial drop in accuracy when applied to novel data (Figs. 4 and 5). On Group\(_{\text{val}} \), the Banister model \( R^2 \) was better at predicting final APFT run time than LRM\(_{\text{step}} \) (0.55 vs. 0.47, Figs. 3 and 4) and similar to LRM\(_{\text{full}} \) (0.55 vs. 0.50, Figs. 3 and 5). However, we find the Banister model is a substantially better predictor of APFT results for intermediate APFT results (Group\(_{\text{nov}} \)) compared to both LRM\(_{\text{step}} \) and LRM\(_{\text{full}} \) (\( R^2 \) of 0.59 vs. 0.36 and 0.40, respectively). This demonstrates the extensibility of a physiology-based model to training scenarios other than the regimen on which it was calibrated.

As with all models, errors in data collection and poor assumptions will affect accuracy. Models with more inputs...
have more error sources than simple ones. In the case of the Banister model, since it was not feasible to collect either oxygen consumption or HR on an individual-by-individual basis, these were estimated from observer-recorded perceived intensity of an entire training Company. This introduces 3 sources of error. First, “intensity” is a subjective observation and thus susceptible to observer bias. A second source of error is that each trainee’s actual intensity will differ from the Company mean depending on his genetics, motivation, exercise technique, conditioning, injury status, etc. Finally, the relationship between intensity and oxygen consumption is itself empirical. Also, although trainees were monitored for 18 out of the 24 hours, any activities that occurred during the nonobserved hours (i.e., in barracks) were not included in the training workload $w_i$.

Unmeasured factors that may affect subject fitness such as reduced training due to injury, sleep, menstrual cycle, diet, hydration, and psychology are also potential sources of error. Since these factors ultimately affect the physiological systems that relate to fitness, if measured some could potentially be used to individualize constants $k_1$, $k_2$, $\tau_1$, and $\tau_2$ to increase accuracy.

To minimize all of these effects universally, we optimized the three model parameters ($k_1$, $k_2$, and $f$) using the calibration data set even though average values for $k_1$ and $k_2$ were available in the literature. (Recall $f$ is a “lazy” factor to account for trainees who did not have to perform a maximal effort to obtain a passing APFT score.) Nevertheless, when both $k_1$ and $k_2$ (along with $f$) were optimized for U.S. Army BCT, $k_1$ and $k_2$ values were nearly identical to those found by Morton et al17 for recreational runners undergoing a very different training protocol (4 weeks of intense training followed by 6 weeks of no training) for about the same duration (Table III). This supports the assumption that the model parameters are capturing the effect of the underlying physiological mechanisms of fitness and fatigue during this time frame and illustrates the fact that fewer subjects are typically needed for calibration of a mechanistic model. However, additional studies of shorter and longer duration training regimens with different levels of intensity are needed to fully validate these parameter values as “universal” for all scenarios. In fact, when we used the “non-optimized” parameters $k_1$ and $k_2$ from the literature, we found very little decrease in prognostic accuracy. Thus, the only factor that was “tuned” specifically for this model was the lazy factor $f$. By optimizing that one parameter, we were still able—relative to both LRMs—to (1) make similar or better prognostic predictions for the end-of-cycle final APFT, (2) make far better predictions of midcycle intermediate performance, and (3) develop a model that can be extended to novel regimens since it accounts for physical activity.

The model presented herein does not seem extensible to the two other components in the APFT: the 2-minute maximal effort sit-up test and the 2-minute maximal effort push-up test. We tried to adapt the Bannister model to these tests, but could not produce predictions which were better than linear regressions from initial baseline scores. This is not surprising since the biomechanics and physiology of running—for which Bannister was designed—are different than push-ups and sit-ups. Still, we feel that the promising results seen with run time predictions imply that similar predictability could be achieved with biomechanics-based models, which are specific to those exercises.

It is worth noting that since the PTracker activity monitoring app was developed in 2008, there has been an explosion in personal fitness monitoring devices and smartphone apps. The most ubiquitous of such devices rely on Global Positioning Satellite (GPS) signal tracking to monitor running distance and speed. These two pieces of information are directly related to duration and intensity, which were found to be most determinant of fitness adaptations in our study. Thus, we feel that these GPS-based activity monitors would offer a higher resolution measurement quantification of activity than the observer-based technique used in this model. Perhaps more importantly, the prognostic run time model used here could be linked directly to the GPS stream in a smartphone app, which updates fitness and fatigue predictions in near real time for a trainee or coach to monitor.

In summary, this is the first known application of an established phenomenologically based physical performance model on a military training regimen. We demonstrate that the model and regimen quantification techniques described here are a viable method of predicting changes in run time with training. Because the model provides a predicted time history of fitness, $P(t)$, it has the potential to not only identify trainees at risk of failing the final APFT a priori, but also give guidance as to how the training regimen might be designed to better increase recruit fitness. In a companion manuscript, we address the prediction of the other major source of attrition in Army BCT—overuse injuries.16

ACKNOWLEDGMENTS

We are indebted to Dr. Ed Zambraski and the Military Performance Division at U.S. Army Research Institute of Environmental Medicine for their commitment to the countless days collecting and organizing the most comprehensive set of BCT data to date. Funded by USAMRMC (award number: W81XWH-14-C-0106).

REFERENCES


Physical Training Outcome Predictions With Biomechanics, Part II: Overuse Injury Modeling

Charles H. Negus, PhD; Bryant L. Sih, PhD

ABSTRACT  Objectives: In Part II of a two-part series, we develop a phenomenological model of a negative outcome of U.S. Army Basic Combat Training that affects a large proportion of trainees. Previous models have been epidemiological in nature and have focused on trainee risk factors such as previous injury, gender, and initial fitness. This approach is limited due to difficulties extrapolating results to other cohorts. In addition, training regimen is often neglected, limiting accuracy when applied to novel scenarios. Methods: The prognostic Training Adaptation Injury Model (TAIM) developed accounts for both individual characteristics as well as regimen by integrating validated sub-models of physiological and biomechanical principles known to be important for tibial stress fracture. Results: We find that when used to predict any type of overuse injury, the TAIM is most accurate when the effect of training activities on both overall fitness as well as muscle fatigue during activities is accounted for area under the receiver-operator curve of 0.65. This compares favorably with statistical-based models that do not account for training regimen (area under the receiver-operator curve $\approx 0.56$. Conclusions: The TAIM has the potential to both identify trainees at overuse injury risk as well as make recommendations on regimen changes to reduce that risk.

INTRODUCTION

Physical training during U.S. Army Basic Combat Training (BCT) is designed to improve the physical fitness of trainees so they can meet the physical demands of military operations. Although this 9-week training regimen will ideally enhance the fitness level of the recruit, it can also lead to negative outcomes including musculoskeletal injury and failure to pass the Army Physical Fitness Test (APFT). In Part I,\cite{Negus2016} we described the development of a fitness model to predict a 2-mile run time during BCT. In this article, we describe a biomechanics-based model to predict overuse injury likelihood in U.S. Army BCT.

Historical averages suggest that approximately 25% of men and 50% of women will incur one or more training-related injuries during BCT.\cite{Morgan2009,Young2010,Anderson2013,Mease2015} These are epidemic proportions that have, and will continue to be, a significant drain on military readiness, budgets, and medical services.\cite{Sih2015}

Overuse injuries, as defined by the military, can include minor muscle strains, contusions, tendinopathy, bursitis, muscle or tendon tears, joint sprains or dislocations, and bone stress fractures and stress reactions.\cite{Sih2015} Injury mitigation strategies have succeeded in recent years in reducing the prevalence of training-related overuse injuries,\cite{Sih2015} but they are still a leading cause of attrition and medical costs.\cite{Sih2015} As a result, models that predict likelihood of overuse injury in a military training setting have been developed to identify trainees who may face a predisposition to injury.

Previous prognostic overuse injury models have focused on statistical analysis of epidemiological risk factors.\cite{Sih2015,Mease2015} These risk factors include previous injury, gender, and initial fitness. The application of a statistical approach is limited; however, because of difficulties extrapolating results to training regimens and cohorts different from those that the models are built from. Recent modeling research has recognized that recruit-specific biomechanics such as range of motion limits or asymmetries may predispose one to injury. Functional Movement Screening has been used to predict injury potential in Coast Guard midshipman\cite{Sih2015}; in a cohort of 874 men enrolled in Marine Corps officer candidate training, researchers looked at baseline PFT, self-reported exercise and previous injury history, and Functional Movement Screening scores,\cite{Sih2015} but did not account for training regimen.

To better understand the impact of individual exercises on overall fitness, biomechanical and physiological (B-P) modeling has been used to study a wide variety of exercises including running,\cite{Sih2015,Sih2015} cycling,\cite{Sih2015} and swimming.\cite{Sih2015} These models focus mostly on understanding the musculoskeletal loads such as joint forces, muscle forces, bone stresses, and limb kinematics during exercises, and how they adapt over time. Models of acute and overuse injury mechanisms have been used to estimate the tissue damage resulting from repetitive loading\cite{Sih2015} but simulation of physiological responses to training such as muscle strength and fatigue has been limited. Prognostic models that incorporate B-P factors have the benefit of accounting for the extrinsic risk to injury posed by a specific training regimen, and not simply risk factors intrinsic to a trainee. In doing so, they have potential applicability to any training regimen that can be adequately characterized.

If the inclusion of a training regimen is the strength of a B-P model, it is also what makes model development a
challenge. The mechanistic nature of the model is such that some knowledge of subject-specific physiological characteristics is needed beyond what is typically captured in personal information surveys. More difficult is the need for some characterization of the day-to-day, and possibly minute by minute, activities, which a trainee is experiencing. Without this information, quantities such as internal forces or metabolic cost cannot be estimated with the accuracy needed to make a useful outcome prediction. So the challenge with a mechanistic model is to get the degree of model granularity correct: enough detail to capture injury etiologies, but not so much as to make data collection intractable.

To our knowledge, no such model has been attempted. The purpose of this project is to demonstrate the feasibility of a prognostic overuse injury model that incorporates both trainee characteristics as well as regimen activity. We hypothesize that B-P models that include training activity type, duration, and intensity on a minute-by-minute basis is sufficiently granular to improve predictive accuracy.

METHODS

Data Collection
As described in Part I, trainee and high resolution training regimen data were collected on 563 trainees undergoing 9 weeks of U.S. Army BCT at Fort Jackson, North Carolina, in 2010. Measures include basic anthropometry (height, weight, age, and gender) and initial APFT run times. In addition, medical records were reviewed and injuries classified as overuse were noted. After removing subjects with incomplete information (missing any initial fitness scores or baseline data), data from 347 males (20.7 ± 3.8 years; 77.9 ± 12.9 kg, 176 ± 7 cm) and 108 females (20.5 ± 3.5 years; 62.0 ± 8.5 kg, 162 ± 7 cm) were recorded. There was no significant difference between the mean initial run times of the excluded recruits and those used in the model development and there was no reason to believe that the probability of missing data was related to the value of the variable, so we therefore assumed that the remaining subjects (347 males and 108 females) are simple random subsamples of each group. Ninety male and 64 female overuse injuries were observed using the injury classification developed in Knappik.

To quantify the regimen, U.S. Army Research Institute of Environmental Medicine observers followed the trainees in 6 companies, recording activity parameters including type, duration, intensity, and load carried for the entire 9 weeks of BCT on a minute-by-minute basis. Observers were trained and used a custom Android-based PAtracker software to record activity parameters. Each Company averaged 75.8 ± 17.7 trainees. For this study, it was assumed that a company’s PAtracker data, for walking, running, calisthenics, and other activities, were representative of all the trainees in that company. To our knowledge, this is the most detailed account of U.S. Army BCT currently available and may be the highest resolution military training regimen data of any U.S. military branch to date.

An additional, limited dataset from U.S. Army BCT at Fort Sill, Oklahoma, in 2011 was also collected, which lacked initial APFT results. All data were provided by U.S. Army Research Institute of Environmental Medicine and was collected in accordance with their institutional review board as well as the protection of human subjects (Army Regulation 70-25: Use of Volunteers as Subjects of Research. Washington, DC: U.S. Department of the Army; January 25, 1990; U.S. National Archives and Records Administration. Code of Federal Regulations. Title 45. Public Welfare. 2009).

Model Development
We present the development of a Training Adaptation Injury Model (TAIM), a B-P model that integrates physiological- and biomechanical-based submodels found in the literature and whose combined pathways are intended to capture the primary mechanisms that lead to overuse injury. Specifically, our hypothesis is that injury etiologies leading to stress type overuse injuries in the tibia will be indicative of propensity for mechanically induced overuse in other components of the lower musculoskeletal system. This hypothesis is in keeping with previous research correlating prior injury with future risk (see Kaufman and Marti). In this way, our model uses change in tibia tissue density as a proxy for overuse injury risk in other bones, joints, and ligaments. In other words, a trainee who is predicted to need large increases in tibial density to avoid stress fracture will also be at high risk for other overuse injuries which have a mechanical pathogenesis.

Ultimately, prediction of overuse injury likelihood is based on a logistic regression of two variables: tibia density before BCT $\rho_0$ and the maximum predicted change in tibia density $\Delta \rho_{\text{max}}$ over the first 2 weeks of training, the period in which stress fracture symptoms most frequently occur. Initial density $\rho_0$ is estimated from regressions based on subject anthropometry and as such is training regimen invariant. The $\Delta \rho_{\text{max}}$ is based on per-activity bone loading (including effects of trainee fitness level and muscle fatigue during each activity), bone material properties, and functional remodeling effects. If the model predicts large changes in tibial density in the first 2 weeks of training, we assume that the demands of the training regimen are too great to be accommodated by a trainee’s musculoskeletal system before the onset of injury. This premise is based on research from the authors which has found that a functional deficit in the tibia may contribute to fracture susceptibility under intensive loading conditions.

Component 1: Initial Density, $\rho_0$
To estimate $\rho_0$, we used age, height, and weight in a multivariate linear regression developed from British army infantry training trainees (Table I). The combined cohort had an average starting density of, $\rho_0 \text{Mean} = 1,166.7 \pm 18.0335 \text{ mg/cm}^3$ with a range (minimum/maximum) of 1,136.5/1,222.3 mg/cm³. Since this component of the injury model is independent of
training regimen, it only captures intrinsic injury predispositions. Adaptations to training are captured by Component 2, change in bone density.

**Component 2: Maximum Change in Density, Δρ_{max}**

The second TAIM predictor quantifies whether large changes in tibial stiffness are needed to maintain homeostasis of in vivo stresses to meet the physical demands in the early weeks of training.

Changes to bone density Δρ are described by an iterative cycle of equations driven by lower limb loading. But lower limb loading itself depends not just on a trainee’s activity but current level of muscle fatigue and overall fitness. Thus, capturing bone density changes Δρ because of BCT activities required the linkage of four submodels describing (1) trainee “fitness” changes during BCT, (2) per-activity “muscle fatigue”, (3) net internal “physiologic loading” felt by the tibia per-activity, and (4) “bone integrity” changes because of functional adaptation. Each of these models is an adaptation of previously published models and are shown schematically in Figure 2. The “bone integrity” submodel captures the dynamic processes of density adaptation needed to maintain homeostasis of internal strains. These strains occur as a result of lower limb forces arising during physical activity. While it is relatively straightforward to measure external forces from activity, the “physiological load” (the stress acting on the tibia) must be estimated using biomechanics and/or direct invasive measures and depends on the activity being done as well as tissue geometry and the state of muscle fatigue. The tibia geometry is estimated from regressions of more easily acquirable data such as basic anthropometry. “Muscle fatigue” quantifies how the physiological load changes as muscles become tired during an activity and is dependent on “fitness”, which itself depends on the initial fitness condition of the trainee and the training regimen performed (activity).

---

**TABLE I.** Multivariate Linear Regression Equations at 38% of Tibial Length for Cortical Density (mg/cm³) and area (mm²) for Males and Females From Age (Years), Ht (cm), and BW (kg). Based on Data From Casey et al.²⁶

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>$p$</td>
<td></td>
</tr>
<tr>
<td>Density</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.340607·Age − 0.28955·Ht − 0.27164·BW + 1162.066</td>
<td>0.266761</td>
<td>1.15E − 14</td>
</tr>
<tr>
<td>F</td>
<td>2.38094·Age + 0.016039·Ht − 0.50498·BW + 1170.633</td>
<td>0.180103</td>
<td>1.83E − 12</td>
</tr>
<tr>
<td>CtAr</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>−0.82661·Age − 0.08919·Ht + 1.828133·BW + 209.8625</td>
<td>0.115263</td>
<td>2.66E − 06</td>
</tr>
<tr>
<td>F</td>
<td>0.894749·Age + 1.200174·Ht + 1.394122·BW − 48.7745</td>
<td>0.310369</td>
<td>1.09E − 22</td>
</tr>
</tbody>
</table>
**Fitness Level**

The purpose of the fitness submodel is to estimate trainee fitness at the start of each training day. This is accomplished using the “Banister” model, whose adaptation to U.S. Army BCT is described in detail in Part I.\(^1\) This model is based on the effect of training on two competing components to performance—the positive effect of increased fitness and the negative effect of fatigue\(^2\):

\[
P(t) = P^* + k_1 \sum_{t=0}^{t-1} e^{-(t-i)/\tau_1} w_i - k_2 \sum_{t=0}^{t-1} e^{-(t-i)/\tau_2} w_i
\]

where \(P(t)\) is the level of fitness at time \(t\), \(P^*\) is the initial performance level, the second term is the fitness component, the third term is the fatigue component, \(k_1\) and \(k_2\) are dimensionless weighting factors, \(w_i\) is training load, \(i\) is a given day of training, and \(\tau_1\) and \(\tau_2\) are decay time constants. In this formulation, the change in both fitness and fatigue depend on the training \(w_i\) and how long ago previous training was done (via \(i\), \(\tau_1\), and \(\tau_2\)). An exercise’s contribution to the training load \(w_i\) is the product of the duration and \(\text{VO}_2\)-based intensity. The total training load \(w_i\) for Day \(i\) is the daily sum:

\[
w_i = \sum D \left( \frac{\text{VO}_2\text{ex}}{\text{VO}_2\text{max}} - \frac{\text{VO}_2\text{rest}}{\text{VO}_2\text{max}} \right) e^{\left( \frac{\text{VO}_2\text{ex} - \text{VO}_2\text{rest}}{\text{VO}_2\text{max} - \text{VO}_2\text{rest}} \right)}
\]

where \(D\) is the duration of an exercise (min), \(\text{VO}_2\text{ex}\) is the intensity-based estimated oxygen consumption during exercise, \(\text{VO}_2\text{rest}\) is at resting, and \(\text{VO}_2\text{max}\) is the maximum. The exponent term weights \(w_i\) for high intensity training, and \(c\) varies for males (1.92) and females (1.67), which are values based on blood lactate concentration. Note that the units of \(w_i\) are “weighted” minutes and are arbitrarily defined as “training impulse” or “trimp” units. Thus, \(P^*\) and \(P(t)\) are also in units of trimp and a higher trimp value indicates a higher level of fitness. Additional details on model origin can...
Muscle fatigue plays a significant role in overuse injury. Bio-
mechanical analyses suggest that muscle strength is important
for controlling stresses on the bone through stabilization
mechanics.29,30 Changes in kinematics are also known to occur with
running fatigability and load carriage.31,32 In addition, fatigability
depends on a trainee’s current fitness at the time of activity
which can be estimated using the fitness submodel described
above. Thus, the purpose of the muscle fatigue component is
to account for how the muscles of the body fatigues during
each day’s activities based on the fitness level of the individual.
Fatigue amount is then used to scale the physiological load
felt by the tibia.

In our previous work,14 we described an expansion of
a fatigue model presented by Liu et al.33 and formulated
an algorithm that allows the equations governing muscle
fatigue to be solved if the desired force profile is known.
The model is based on the interaction between four possible
states of muscle motor units: unactive unfatigued
(MUC), active unfatigued (MA), active fatigued (MFU), and
unactive fatigued (MFU). M0 is the total available motor
units, which is always the sum of the motor units in each
state. Only active-unfatigued motor units (MA) generate
force. The equations that define how motor units change
states utilize F and R parameters to control the rate at
which unfatigued motor units fatigue and fatigued motor
units recover, respectively. Fatigue occurs when a muscle
(or activity) requires more unfatigued (MA + MUC) motor
units than available.

The major feature of the muscle fatigue model (MFM) is
that the equations can determine fatigue state (and subsequent
decline in performance) if the desired amount of active-
unfatigued motor units MAD for a given activity is specified.

This is in contrast to motor unit activation via brain function,
which is computationally expensive to determine. The model
attempts to match MAD (i.e., MA = MAD), which results in
two possible cases:

\[
M_A = \begin{cases} 
\text{Case A: } & \text{if } M_A < M_A + M_{UC}, \text{ then } \frac{dM_A}{dt} = \frac{dM_{AD}}{dt} \\
\text{Case B: } & \text{if } M_A \geq M_A \text{ and } M_{UC} = 0, \text{ then } \frac{dM_A}{dt} = R \times M_F - F \times M_A 
\end{cases}
\] (3)

In addition, the total amount of fatigued motor units can be
calculated from:

\[
\frac{dM_F}{dt} = F \times M_A - R \times M_F 
\] (4)

The derivation of Equations (3) and (4) can be found in
Sih et al.14

Because overuse injuries can occur in different parts of
the body and with different tissues, we use VO2, a whole-
body measure, as the basis for M0. M0 can be measured from
an instantaneous maximal effort when all motor units are
assumed to be in the MA state (M0 = M_L). The earliest mea-
sured maximal run (initial 1-mile run for the U.S. Army) was
used as input to a regression equation14 to estimate theoretical
instantaneous maximal speed. The metabolic cost relative to
resting VO2 (METs) for that speed (mph) was determined
from a second regression (METs = 1.5222 × speed + 0.152,
R² = 0.97)$^{35}$ which was then divided by VO2max to normalize
M0 estimates as a percentage of VO2max.

MAD is also specified as a percent of VO2max, using the pre-
viously described VO2 in Equation (2) divided by VO2max.

To account for the effect of training (fitness submodel)
MFM parameters, scaling factors were used to adjust M0,
F, and R at the start of each day’s activities (see $^\circ$ and $^\circ$
in Fig. 2). Since fitness typically increased from an initial
value of 2,000 to 10,000 trimp or more during BCT,1 we
assumed that values above 4,000 indicated improved fitness
and imparted a percentage change in \( M_0, F, \) and \( R \) with maximum change occurring at 6,000 trimp or above. Based on a previous analysis of the effect of training,\(^{14}\) we specified a linear change of \( M_0, F, \) and \( R \) baselines with zero change at 4,000 trimp or below and a maximum of 109% for \( M_0, \) 25% for \( F, \) and 200% for \( R \) when 6,000 trimp or above was reached.

**Physiological Load**

Basic physics indicates that axial stress in a long bone depends on the physiological force and bone area. While physiologic loads are continuously changing, our model is based on responding to peak ground reaction forces (GRF’s), estimated from total body weight (including any external weight such as a backpack) and activity. Scaling factors are used to adapt GRF’s to internal physiological force and adjust for fatigue:

\[
\sigma = \frac{(M_{\text{body}} + M_{\text{ext}})g}{CtAr} \times SF_{\text{act}} \times SF_{\text{bio}} \times SF_{\text{fat}} \quad (5)
\]

where \( M_{\text{body}} \) and \( M_{\text{ext}} \) are the mass (kg) of the trainee and any external load, respectively, and are multiplied with gravity \( g \) (m/s\(^2\)) to give force (N). This force is scaled by an activity-dependent factor, \( SF_{\text{act}} \) (Table II). The force is further scaled by \( SF_{\text{bio}} \), which amplifies external GRF to an internal physiological force from the affect of muscle moment arms. We used a value of 10 for \( SF_{\text{bio}} \) based on observed values during one-foot hopping\(^{36}\) and which we assumed would be an upper bound of internal tibial loading during field exercises. The fatigue scale factor \( SF_{\text{fat}} \) further modifies the force based on input from the MFM. To adjust the loading forces to account for fatigue (\( SF_{\text{fat}} \)), we turned to cadaveric studies that quantified changes in bone strain due to muscle fatigue.\(^{30}\)

At the most extreme, increases of up to 202% with fasciotomy (i.e., complete fatigue) were seen, which we assumed only occurred at \( M_f \) levels of 95% \( M_0 \) or more. Force is then divided by cross-sectional area \( CtAr \) (mm\(^2\)) of the tibia to give stress \( \sigma \) in Equation (5).

Finally, to estimate cortical area of the tibia \( CtAr \) we use tibial trait regressions derived from peripheral quantitative computed tomography images from the same studies used for the density regressions.\(^{26}\) These regressions correlated basic anthropometrics (which are collected for every trainee) to bone geometry and tissue density. The regressions for male and female cortical area at 38% of tibial length can be found in Table I.

**Bone Integrity**

The bone integrity model is intended to describe a trainee’s ongoing ability to withstand a stress fracture (also called a fatigue fracture). A bone’s ability to withstand repeated loading at levels less than its ultimate strength depends on multiple factors, including the material property, the loading conditions, and the ability of the bone to remove cortical tissue with concentrations of microcracks and replace it with new bone.

It is this remodeling process that is captured in the TAIM by a functional adaptation model\(^{37}\) whose rate of density changes \( dp/dt \) are governed by tibial strain \( \psi \). Referring to the bone integrity submodel in Figure 2, stress \( \sigma \) from the physiological load submodel is converted to \( \psi \) by dividing by elastic modulus \( E \) (constitutive law). The \( dp/dt \) is governed by a bone remodeling rule, which is of the same type used previously.\(^{38,39}\) And \( \rho \) modifies \( E \) in the structure-function relationship via a regression\(^{25}\): 

\[
E = 0.03021 \cdot \rho - 18.35 \quad (6)
\]

The key mathematical concept of the damage-repair model is the remodeling rule (Fig. 3), which adjusts \( \rho \) based on the strain magnitude \( \psi \). If stimuli (in vivo strains) are less than what is needed to maintain bone mass, density decreases. If stimuli are excessive, bone density increases which acts to stiffen the bone organ and reduce future strains. Activities that produce strains in a “lazy zone” elicit no changes in density. The rate of change of density on either side of the lazy zone were set to match observed maximal changes in density reported in the literature.\(^{40}\) These values \( r_1 \) and \( r_2 \) were set at 0.28E-3 and 0.342E-3 mg/cm\(^2\)-min. \( \psi_{\text{EQ}} \), and \( S \) of the target homeostatic strain and lazy zone width (the range of strain values which maintain density) are 2,000 and 200 \( \mu \)strain, respectively, which were chosen based on a survey of in vivo strains during physical activity.\(^{41}\) This target strain is at the higher end of equilibrium strains in the literature and is intended to be conservative so that only trainees with chronically high changes in density will be

---

**TABLE II.** Ground Reaction Force Scaling Factors (\( SF_{\text{act}} \)) that are Multiplied by Body Weight and External Load to Get Total Force. Values are Commonly Seen in Biomechanical Analysis of These Movements

<table>
<thead>
<tr>
<th>Stationary</th>
<th>Walk</th>
<th>Cadence March</th>
<th>Calisthenics</th>
<th>Run</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>1.2</td>
<td>1.2</td>
<td>2.0</td>
<td>2.5</td>
</tr>
</tbody>
</table>

---

**FIGURE 3.** The remodeling rule uses a step linear ramp function to define the rate of change for bone density \( \rho \) at different loading strains (stimulus values) \( \psi \). A strain “lazy zone” centered about \( \psi_{\text{EQ}} \) with a width of 2S is defined as an area that elicits no change in density. The values \( r_1 \) and \( r_2 \) set \( dp/dt \) outside the lazy zone. Values for \( \psi_{\text{EQ}}, S, r_1, \) and \( r_2 \) can be found in the text.
flagged for injury potential. Maximum and minimum densities were set at 2,000 and 1,000 mg/cm³, respectively. Both bounds are set to give some buffer around physiologic maximum and minimum volumetric densities.

Accuracy Assessment
Having an established method to estimate initial bone density $\rho_0$ (component 1) and a set of submodels based on the literature to estimate the change in density in tibial bone during training so that $\Delta \rho_{\text{max}}$ can be calculated (Component 2), they were implemented in MatLab R2014b and logistic regression was used to determine a combined injury risk for the Fort Jackson dataset. We report the area under the receiver-operator curve (ROC AUC) to objectively determine the accuracy of the models to predict overuse injury. AUC compares the change in precision and recall over a range of model output thresholds. A higher AUC is considered better since higher precision and recall is desired and an AUC of 0.5 is no better than random chance.

RESULTS
To quantify the gain in overuse injury predictability by accounting for the training regimen, four different analyses are presented. The first model uses only Component 1 as an injury predictor (Model $\rho_0$), using different thresholds of initial density to see if those with higher initial volumetric density are at risk. The second isolates the training regimen by correlating overuse injury to $\Delta \rho_{\text{max}}$ while ignoring fitness and muscle strength adaptations by bypassing these submodels of Component 2 (Model $\rho_0$). The third model (Model $\Delta \rho_{\text{musc}}$) introduces these submodels in predicting maximum change in $\rho$. The fourth and final model (TAIM) is logistic regression of components 1 and 2 (Model $\rho_0$ and Model $\Delta \rho_{\text{musc}}$).

When applied to the Fort Jackson dataset, Model $\rho_0$ (Fig. 4) had an AUC of 0.61. Model $\Delta \rho$ was not predictive (AUC = 0.49), whereas the inclusion of the muscle submodels (Model $\Delta \rho_{\text{musc}}$) improved AUC to 0.56 (Fig. 5). TAIM (Fig. 6) had the highest AUC of 0.65 (positive likelihood ratio: 1.758, negative likelihood ratio: 0.568). We were unable to apply the Model $\Delta \rho$, Model $\Delta \rho_{\text{musc}}$, and TAIM models to the Fort Sill dataset because of missing initial fitness test results. However, Model $\rho_0$ had an AUC of 0.60 with Fort Sill, which is a similar predictability as with the Fort Jackson dataset. The regression coefficients for the TAIM are given in Table III.

To compare the results of the TAIM model against a purely statistical model, we developed a logistic regression involving age, gender, height, weight, and initial 1-mile run time (Table IV) using the same subjects as with the TAIM. This five parameter model produced a ROC curve with AUC of 0.67, slightly higher than the TAIM (Fig. 7).

DISCUSSION
This article represents the first application of established physiological models to mathematically represent the mechanisms that lead to changes in tissue so that overuse injury risk can be assessed in U.S. Army BCT. The model accounts
for the effects of anthropometric traits (height, age, weight, and gender), initial fitness, and also training regimen to make an assessment. After considering the isolated effects of initial tibial density (Model $\rho_0$), maximum density change (Model $\Delta \rho_{\text{musc}}$) predicted from training regimen alone, and maximum density change predicted by inclusion of fitness adaptations and muscle fatigue effects (Model $\Delta \rho_{\text{musc}}$), we found that inclusion of both (Model $\rho_0$) and Model $\Delta \rho_{\text{musc}}$ were necessary to maximize the ROC AUC; each submodel incrementally improved the model as indicated by the TAIM, which has the highest ROC AUC. The overall model is modular in nature to accommodate higher resolution measurements of any component. For example, while initial volumetric cortical density is currently surmised from a regression analysis, if this value was available from computed tomography, it would provide a better prediction. And since the pathway leading to overuse injury is similar for all soft tissues, it may be possible to use this technique to predict a specific overuse injury such as a joint strain by tailoring the submodels for that specific injury.

When capturing the effect of training, we compared predicted density changes with and without the influence of soldier fitness adaptation and muscle fatigue (Model $\Delta \rho_{\text{musc}}$ vs. Model $\Delta \rho$). In short, we saw no improvement in overuse injury predictability until changes in muscle strength and fatigue were included. This supports the hypotheses put forth by others that muscle fatigue plays an important role in controlling the skeletal loading conditions and preventing overuse injuries.  

This may also explain the lack of consensus in the literature on the effect of training regimen. For example, while it is well established that higher running mileage increases injury risk, the relationship to more specific measures such as intensity, duration, and rest are not as consistent. The contribution of the fitness and muscle fatigue submodel in the TAIM suggests that injury rates are dependent on fitness and fatigue, which are time dependent, non-linear functions affected by intensity, duration, and rest. Thus, it is not surprising that straightforward relationships between these measures were not found.
A major premise of our model is that cortical bone adaptations are a good proxy for overall skeletal health. It should be made clear that actual functional adaptation in adult bone—a phenomenon whose first observation is commonly credited to Julius Wolff44—is a slow process. Astronauts in extended spaceflight have been observed to lose 2% of bone mineral density per month in microgravity and over 10% over 180 days,45 but on Earth, changes to physical activity regimens are typically seen over a timescale of half a year, if measurable at all. Our cortical functional adaptation cycle aims to simulate the extent to which a tibia would have to stiffen to accommodate the increased loads placed upon it. Since peristeal diameter is essentially fixed by adulthood, the only mechanism that adult bone has to increase organ stiffness is compensation by cortical density.46 Our model for density evolution assumes that those trainees who are predicted to compensate for inadequate tibia stiffness during training with large increases in density should be considered to be at a higher risk of injury since their actual remodeling rate in vivo would be far too slow to meet these suddenly increased loading demands.

Compared to statistical based results found in the literature, the TAIM AUC of 0.65 compares favorably. The five-parameter model logistic regression had a slightly higher AUC (0.67 vs. 0.65) but required optimizing the regression coefficients of five intrinsic recruit traits, compared to the TAIM which had a single intrinsic regression coefficient. A previously reported statistical model that did not include training regimen effects reported an AUC of 0.55 and 0.57 for females and males, respectively,10 which is similar to the 0.61 AUC of Modelρ0.

Although AUC is widely used by clinicians for evaluating prognostic tests and models, its application to overuse injury modeling is not perfect for two reasons. First, “overuse injury” is not a binary outcome in the same manner as other diseases. The label can cover a variety of conditions with unique etiologies. Even tibial stress fractures—the archetypal overuse injury—can be given multiple diagnoses (e.g., “stress reaction,” “stress fracture,” and “shin splints”) because of variability in presentation of symptoms and the timing of the clinical visit. Second, different types of overuse injury (e.g., stress fracture, ligament strain, fasciitis, and compartment syndrome) do have varying etiologies. Since the TAIM model used bone density evolution as a “canary in a coal mine” for all lower-limb, mechanical overuse injury, improving AUC further would likely require incorporation of injury-specific submodels.

Even if the present AUC may be too low to make a recruit-specific training decision, the ability to simulate recruit-specific skeletal adaptations with biomechanics permits estimation of injury outcome trends at the cohort level. We believe this is the first time a set of models (along with Part I1) has been developed that uses quantified physical activity as a model input and as such, has the potential to be used as a decision aid to objectively recommend training regimen changes that will maximize fitness while minimizing overuse injuries in BCT.

In summary, overuse injuries from BCT are a significant burden on military readiness and medical costs. We show that, unlike a purely statistical approach, a B-P model such as the one described here has the potential to go beyond just identifying trainees at risk and begin to reduce these burdens by predicting the time course of injury risk so that objective decisions about training protocol changes and interventions can be made.

REFERENCES


MILITARY MEDICINE, Vol. 181, May Supplement 2016 93
Effectiveness of an Injury Prevention Warm-up for Unit Physical Training: A Case Series of Two Flying Squadrons

Capt Joshua D. Shumway, USAF BSC; Capt Danielle N. Anderson, USAF BSC; Lt Col Brett Bishop, USAF BSC

ABSTRACT  Injury prevention has been assessed and studied in professional and collegiate athletic populations, but application to the military setting has been limited. The purpose of this study was to assess the effectiveness of an injury prevention warm-up in two flying squadrons. At the commanders’ request, two Air Force flying squadrons (276 individuals) were provided an injury prevention warm-up of evidence-based exercises, which focused on functional range of motion and dynamic core stability. The routine was performed before unit physical training twice a week. The number of injuries did not significantly decrease after the injury prevention warm-up compared to 12 months before the intervention. However, the amount of time a subject was “grounded,” duty not involving flying, because of a musculoskeletal injury decreased significantly from 146 days per month to 73 days per month ($p = 0.02$). A quick, generic warm-up of evidence-based exercises may decrease the number of limited duty days in a flying population.

INTRODUCTION  Musculoskeletal injuries are the primary source of disability in the U.S. military contributing up to $3.6$ billion of patient care-related costs in 2013 (including direct, pharmacological, and radiological costs as well as evacuation from theater) per the Military Health System Mart (M2) Database. In the past 13 years, musculoskeletal injuries have steadily increased despite the drawdown in Iraq and Afghanistan. From 2001, approximately 30% of the 45,000 aeromedical evacuations from the Area of Operation have been due to overuse or traumatic-, nonbattle-, musculoskeletal-related injuries (e.g., hamstring strains, herniated nucleus pulposus, and ankle sprains). Evaluation and treatment of these injuries greatly impacts not only the Department of Defense budget, but also largely contributes to the Department of Veterans Affairs annual expenditures. This year alone, over $35$ billion was spent by the Department of Veterans Affairs on musculoskeletal injuries, enough to fund the United States Marine Corps’ budget for nearly 18 months. Clearly, this is a problem contributing to thousands of lost man hours, increased health care costs, and ultimately negatively impacting the overall military mission. In 2010, the U.S. Army took notice of this issue and called for change in the current health care model to switch the focus from a health care system, to a system for health focused on prevention rather than treatment for an already procured ailment. With this, the U.S. Army became a leader in the field of preventing musculoskeletal injuries.

A majority of the previously published injury prevention data focuses on screening athletes to assess if they are at a risk for injury. Early studies focused on foot structure, core strength, and power. However, many of these studies were not shown to be statistically reliable in the military population. On the basis of information provided by Teyhen in 2011, whole-body movement and quality of movement testing, to include the Y-Balance Test and the Functional Movement Screen (FMS), proved to be the most valid, reliable, and efficient tools in the military setting. Similarly, there is substantial evidence on the reliability and validity of these specific tests in the athletic population to include the National Football League, National Basketball Association, National Collegiate Athletic Association, as well as firefighters in southern California. This is beneficial as preliminary data suggests if an individual is screened with the FMS and found to be at increased risk for injury, they can perform tailored corrective exercises to decrease risk for future injury. These exercises are tailored to the individual and include common themes of increased functional range of motion, multi-planar/multidirectional movements, and dynamic core stability.

Despite evidence proving that screening individuals can lead to decreased injuries, performing these tests routinely may not always be feasible. Specifically, to perform an accurate screening assessment can take at a minimum 15 minutes per individual. Although this time may be negligible when screening small groups such as sports teams, the impact increases dramatically when screening the population of an entire military installation. For instance, to screen 7,063 active duty members at Travis Air Force Base would take approximately 73 days to accomplish; to screen 324,380 active duty members, Air Force wide would take approximately 9.23 man-years to accomplish. Spending the time and resources to...
screen the entire military populace may not be feasible. Therefore, the objective of this study is to assess if a generalized prephysical training warm-up, that uses the same principles of the FMS corrective exercises, could be used to decrease injury without performing mass individual screenings.

**METHODS**

At the request of the commanders of the 6th and 9th Air Refueling Squadrons (ARS) for an intervention to decrease injuries, the David Grant USAF Medical Center Physical Therapy Department staff developed a multidirectional, multi-planar injury prevention warm-up focusing on functional range of motion and dynamic core stability. Many of the movements mimic or combine commonly prescribed corrective exercises from the FMS and common physical therapy practices (Figs. 1–10).

The warm-up begins with mini lunges and various overhead, side to side reaches. These movement patterns include gentle lower extremity strengthening, lumbar and shoulder mobility, and thoracic extension. Squats accomplished with specific focus on proper form were performed because of the high-quality strengthening for quadriceps, gluteal, and core strengthening.

A single agility drill (t test) was performed as it is easy to reproduce and helps train participants in agility. Although the t test is not proven in itself to decrease injury in agility sports, it has been shown to reduce the overuse of the preferred leg when stopping during sport, which has been linked to increased injury.20

Finally, dynamic planking was used to strengthen the core through the entire range of motion, increase thoracolumbar range of motion and strengthen the scapular stabilizers. It is important to note that the warm-up as a whole was designed to be applicable to a healthy population and not designed to be an all-inclusive strengthening or range of motion program. Rather it was designed to be comparable to a multitude of evidence-based multiaxial, neuromuscular exercises that have been proven to reduce injury21 and commonly given as corrective exercises from FMS to target generalized essential movement patterns.

With commanders’ approval, this warm-up was taught to the Physical Training Leaders (PTL) of the 6th and 9th ARS, which instruction took approximately 20 minutes per squadron. The warm-up was performed before every physical training session, which occurred twice a week. Musculoskeletal injuries and days where member was on Duties Not Involving Flying (DNIF) data were then obtained from the Aeromedical Services Information Management System.

A hands-off approach was taken with the squadrons; the physical therapists did not interfere with the implementation of the warm-up past the first day of instruction. The reason for this approach was two-fold: to assess how much time needs to be invested on the part of medical staff to accomplish the goal of decreasing injuries, and to encourage members to continue to report injury as it has been observed that the presence of medical staff usually decreases injury reporting due to fear of being “grounded”–DNIF. The squadrons eventually stopped performing the warm-up—the 6th ARS after 7 months, and the 9th ARS after only 2 months as a result.
of changes in the PTLs. Data were obtained for the 12 months before and 12 months after the start of the intervention period for each squadron.

The number of DNIF days related to musculoskeletal injury were assessed in the participants before the time of intervention. Means and standard deviations (SDs) were calculated for 1-year preintervention, during the time of the intervention itself, and 1-year postintervention. Although significance is usually not reported in case series, as a result of the change being fairly small, the authors attempted to

FIGURE 2. Lunges with Lumberjack 1. Starting in standing, the subject lunges forward and reaches both hands to the forward leg. When he steps back, he reaches both hands over the opposing shoulder rotating through the hips.

FIGURE 3. Lunges with Lumberjack 2. Starting in standing position, the subject lunges forward and reaches both hands away from the forward leg. When he steps back, he reaches both hands over the opposing shoulder.
validate their findings in more than just averages before and averages after intervention. Therefore, \( p \) values were assessed to help determine the significance of the effect. As this data analysis is of a case series with limited control over the subjects or conditions, the authors were willing to accept a higher level of error and the significance level was set at \( \alpha = 0.10 \).

\( t \) tests were used to determine if there was a significant difference between the preintervention and the time of the intervention and preintervention and 1-year postintervention. Analysis was performed using STATA statistics software version 13.0 (Stata Corp, College Station, TX).

**RESULTS**

As shown in Figures 11 and 12 and Table I, the 6th ARS performed the warm-up independently for 7 months. During this time, there were a steady number of injuries, about two per month. The number of limited duty days leveled out at about 100 per month despite ceasing the warm-up. Although the number of injuries decreased during the period of the warm-up, the results were not significant (2.58 ± 1.83–2.13 ± 1.36, \( p = 0.55 \)). The results for DNIF days were also nonsignificant (143.17 ± 162.65–86.76 ± 69.94, \( p = 0.37 \)).

However, if analyzed from the 12 months before the start of the intervention to the 12 months after intervention was implemented, the number of injuries per month decreased significantly from 2.77 ± 1.88 to 1.62 ± 0.87 (\( p = 0.06 \)). Also, the number of limited duty days per month decreased significantly from 147.9 ± 156.6 to 68.69 ± 46.23 (\( p = 0.09 \)).

In the 9th ARS, Figures 13 and 14, the warm-up was performed for only 2 months. During these 2 months, a sharp decrease in the number of injuries is noted over the period of performance (2 months of mid-September–mid-November 2013). Despite this, as with the 6th ARS, the decrease in the number of injuries were not significant (3.50 ± 2.08–2.50 ± 1.73, \( p = 0.94 \)) nor were the number of limited duty days (156.92 ± 142.78–87.25 ± 60.66, \( p = 0.37 \)). In the 12 months before the start of the intervention to the 12 months after implementation, number of injuries actually increased (2.42 ± 1.88–2.67 ± 1.37), but this was also nonsignificant (\( p = 0.71 \)). However, the number of limited duty days decreased from 156.92 ± 142.785 to 77.42 ± 54.89, which did prove to be statistically significant (\( p = 0.08 \)).

When we combine the two squadrons for a total of 276 airmen, the mean number of limited duty days per month during the 12-month period before the beginning of
intervention was 146.72 ± 149.94, and the 12-month period after start of the intervention was 72.88 ± 49.69, \( p = 0.02 \). The mean number of musculoskeletal injuries per month during the 12-month period before intervention was 2.60 ± 1.85 (SD), and during the 12-month period after intervention was 2.12 ± 1.24 (SD), \( p = 0.29 \). Finally, as both squadrons did not complete the interventions for the same amount of time, we did not analyze the data before vs. during intervention.

**DISCUSSION**

These results demonstrate that providing a generic, multiplanar, multidirectional injury prevention warm-up was associated with a reduced the number of monthly DNIF days in

**FIGURE 6.** Carioca. In a standing position, the subject moves laterally, bringing the trailing foot up to hip level, across anteriorly, back to starting position, and moves the trailing leg across posteriorly.

**FIGURE 7.** Planks with upward/downward dog. Staring in a push-up position, the subject raising the pelvis toward the air (downward dog), and then toward the ground without touching (upward dog).
two flying squadrons. In our sample size of 276 airmen, the warm-up was associated with a decrease in the number of DNIF days by 74 per month or nearly 900 days per year. The same decreased rate in DNIF days could have dramatic effects on duty and deployable availability mobility rates when extrapolated throughout a larger nonflying population.

The goal of the intervention was to decrease the number of injuries. Results indicated that the number of injuries decreased by one per month per squadron after the intervention was implemented. Although this was not statistically significant, it may have clinical significance. For example, the amount of time invested by the medical staff to provide instruction and demonstration to the PTLs was about 40 minutes, which may have saved a possible 900 limited duty days per year for both the squadrons. When applied to all squadrons at Travis Air Force Base, preventing one injury per squadron per month equates to potentially 288 injuries prevented and nearly 1,800 limited duty days saved per month.

The data are presented as both separate and combined squadrons to illustrate the different time periods when the injury prevention program was given to each squadron as well as to show the previous levels of injuries. The results were then combined for 12 months before and 12 months after the intervention. The 6th ARS performed the warm-up independently for 7 months while the 9th performed the program for only 2 months. Despite the 9th ARS’s limited 2-month involvement, it is interesting to note that the

FIGURE 8. Planks with rotation. Starting in a push-up position, the subject moves his body weight to one side and raises his arm up from the ground to form a “T.” He then returns to the starting position and performs the same movement on opposing side.

FIGURE 9. Planks with step over. Starting in a push-up position, the subject moves his body weight to one side, and raises his leg to touch the opposing side. He then returns to the starting position and performs the same movement on opposing side.
number of DNIF days stayed relatively constant yet there was a significant rise in the number of injuries after the intervention stopped (Figs. 13 and 14). This may indicate a decrease in the severity of injuries since implementing the warm-up. It is also interesting to note that during the time period of performing the warm-up, both squadrons had the lowest number of DNIF days per month compared to the past 2 years.

There are many limitations with this study, first and most important is the hands-off approach that was taken. This tactic did not allow the researchers to ensure the PTLs were performing the warm-up correctly or ensure they continued to perform them. However, the fact both squadrons ceased the program after a short amount of time shows us that in future implementations, commanders and medical personnel may need to “check in” with squadrons every 3 to 4 months to ensure they are still performing the exercises and doing them correctly.

Secondly, the nature of a case study does not allow for correlation or cause and effect relationship with the decrease in DNIF days being directly related to the intervention. Several other variables must be taken into account as they were not controlled in this study. Such variables include changes in the culture about seeking medical care in flight medicine, changes in flight surgeons opinions on necessary DNIF days, as well as increase/decrease in deployment ops tempo to name of few. As these confounding variables were not controlled for they must be considered a limitation to
Finally, it is important to note that providing a generic injury prevention warm-up may not be the best method to prevent injury. Nearly all previously published data incorporate screening an individual and providing a targeted approach of corrective exercises to that individual’s impairment.\(^{17-19}\) Although this may prove to be the most effective way to target injuries, it does not seem feasible to screen each individual with such large numbers of military personnel. More research is needed on the effectiveness of task-specific, essential movement pattern training in a variety of populations. This training is currently being performed in U.S. Army Brigade Combat Teams and Special Operators throughout the Department of Defense (e.g., Navy SEALs, Pararescuemen, and Rangers). Studying the effectiveness of having a physical therapist embedded within these units to optimize human performance could prove beneficial. It is our hope that future research is targeted to decreasing injuries in large populations where screening each individual may not be feasible.

**CONCLUSIONS**

In our population, the 5-minute generic warm-up implemented by the squadron’s PTL was associated with a decreased number of limited duty days. Although there was no significant decrease in the number of injuries, the implementation of the warm-up was associated with a decrease in DNIF days. Further, high-quality research is needed to truly assess these findings.

**REFERENCES**


Effectiveness of an Injury Prevention Warm-up for Unit Physical Training


Pulmonary Emboli and Deep Vein Thromboses: Are They Always Part of the Same Disease Spectrum?

Nicole T. Gordon, MD; COL Martin A. Schreiber, MC USAR

ABSTRACT
Background: Pulmonary embolism (PEs) are thought to emanate from deep vein thromboses (DVTs). Government agencies now use thromboembolic events as a quality metric for reimbursement for care. Recent data suggest that PEs and DVTs may represent different pathologic processes. We sought to identify separate risk factors for PEs and DVTs to test whether they are the same disease process. Methods: A retrospective review of the National Trauma Data Bank between 2007 and 2010 was performed. Demographics, complications, comorbidities, and injury data were reviewed for risk factors for patients diagnosed with a PE or DVT. Results: After exclusion criteria were met 521,969 patient entries were analyzed. Of these patients, 4,154 and 1,460 had a DVT or PE, respectively, while 8% (433) of patients had both. PEs and DVTs, had 18 overlapping risk factors, 26 independent risk factors (5 for PEs; 21 for DVTs), and one divergent risk factor. Conclusion: Despite PEs and DVTs having overlapping risk factors, there are significant independent and divergent risk factors for the two diseases, suggesting that they are not always part of the same process. The constellation of risk factors for each disease may help to predict which one patient is predisposed to and draws into question the concept of using them as a quality metric as whether therapeutic anticoagulation is indicated in trauma patients.

INTRODUCTION
Pulmonary embolism (PEs) as a result of dislodgement of clot from deep vein thromboses (DVTs) were originally described by Dr. Rudolph Virchow in 1845. Since his discovery, there has been little change in the thought process surrounding the pathophysiology of venous thromboembolism (VTE). It was not until the advent of computed tomography (CT) scans when questions started to arise regarding how related were the two diagnoses.

In 2009, a retrospective review of trauma patients who underwent CT imaging of the chest and lower extremities in an effort to diagnose patients with a PE and source of the embolism showed little overlap in the number of patients diagnosed with a PE and DVT. The authors hypothesized that PEs could be the result of a primary pulmonary artery thrombosis rather than a dislodged embolus from DVTs. However, no significant different risk factors were identified in patients with only a PE or a DVT.

More recent studies have started to identify independent different risk factors for patients diagnosed with each disease entity. In general, trauma patients diagnosed with a PE tend to have more severe chest injury, whereas patients with a DVT have a more varied injury pattern, and benefit from prophylaxis within 48 hours.

However, criticisms of these studies include the selection process for risk factors and the failure to control for confounding variables. For certain studies, candidate risk factors were limited to known causes of VTE. By limiting a study to only known risks factors, it is possible that other critical modifiable ones could be missed. Furthermore, in addition to limiting risk factors to only known causes of VTE, other studies did not clearly state the selection process for which risk factors to study.

The clinical relevance of questioning whether the two disease processes are always related is two-fold. First, if a patient developed a de novo pulmonary thrombus (DNPT), it draws into question whether therapeutic anticoagulation or an inferior vena cava (IVC) filter is indicated. Combat patients have a higher risk of developing a VTE compared to the civilian. Limiting therapeutic anticoagulation and IVC filters in patients with DNPTs curtails the associated risk with their use, especially in a combat trauma patient that is at risk of bleeding.

In addition, if DNPTs do occur, it challenges the use of VTE as a metric for quality-based reimbursed care. The government initiated Agency for Healthcare Research and Quality considers PEs as the most common preventable cause of hospital death. The Centers for Medicare and Medicaid Services has started to view PEs in certain postoperative patients as a never event and has limited reimbursement for treatment of PEs in these patients.

The objective of this study was to further test whether PEs and DVTs are always related by reviewing all risk factors associated with admission characteristics and all listed comorbidities and complications tracked by the National Trauma Data Bank (NTDB). All discovered risk factors would then be subjected to a multivariate analysis to control for confounding variables. We hypothesize that trauma patients diagnosed with DVTs or PEs will have different risk factors, which may suggest that they are not always the same process.

METHODS
With Institutional Review Board approval, we tested our hypothesis by using the data set provided by the American
College of Surgeons NTDB between 2007 and 2010. These years were used as they consistently reported data in the same fashion. All patients within the databank that had all data points studied were included. Exclusion criteria included patients who were less than 18 years of age and those who expired in the emergency department or within 2 days of admission.

All complications and comorbidities that are tracked within the NTDB were candidate risk factors. These are listed in Table I. To account for quality control of the database as previously described, only patients who were cared for at level I and II trauma centers were studied as they were most likely to report complications and comorbidities. Patients who were cared for by centers that did not report the most common complication (pneumonia) or comorbidity (hypertension) within the years of the NTDB datasets that were used, were also excluded given that if they do not report the most common complication they are likely not to report any in general. In addition, centers that were designated by the NTDB as those that do not submit complication or comorbidity data were excluded from the study.

Patients diagnosed with a PE or DVT was identified. Patients diagnosed with a PE were identified by using the International Classification of Diseases, 9th revision codes for PE (415.1, 415.11, and 415.19) and patients who were diagnosed as having a PE as an NTDB complication. Patients diagnosed with septic PE were not included in the study. Patients diagnosed with a DVT were identified by International Classification of Diseases, 9th revision codes (451.11, 451.19, 451.2, 451.81, 453.2, 453.8, 453.9, 451.83, 541.89, and 451.84) as well as patients who were identified as having a DVT as an NTDB complication.

Variables studied were all NTDB comorbidities and complications tracked within the database, demographics of the patients including age (treated as a continuous variable), gender, race, and mechanism of injury. The highest abbreviated injury scale (AIS) score per region, the Glasgow Coma Scale (GCS) score, and patient vitals on arrival were included in the analysis. All comorbidities and complications that are specifically tracked in the NTDB annually were included as many were considered risk factors for VTEs. The study was not limited to known causes of VTE to decrease the possibility of biasing the study toward known risk factors thereby eliminating other potential ones from being discovered.

Statistical analysis was performed using SPSS (version 21) (IBM Corporation, Armonk, New York). All risk factors were subjected to a univariate analysis with significance being defined as a p value of less than 0.20. Afterward, a multivariate backward logistic regression was used to control for confounding variables. Significance was defined as a p value less than 0.05. All risk factors were studied for each patient population.

### RESULTS

The total NTDB cohort studied included 1.1 million patients. After exclusion criteria were implemented, 521,000 patients remained in the study. Demographics are listed in Table II. The mean age was 42 years, the mean Injury Severity Score was 9, and 64% of the population were males. Blunt injury accounted for the most common mechanism of injury. There were 4,154 patients diagnosed with a DVT, 1,460 patients with a PE, and there were 433 patients diagnosed with both a PE and DVT, indicating an 8% overlap in patients who were diagnosed with both a PE and DVT. PEs and DVTs had 18 overlapping risk factors, 26 independent risk factors (5 for PEs; 21 for DVTs), and one divergent risk factor.

Demographic and clinical characteristics at the time of presentation are listed in Table III. Significant risk factors for thromboembolism were gender, with the female gender having a decreased odds ratio for developing a DVT or PE, when compared to males as a reference. Race was also a risk factor for PE but not DVT.

<table>
<thead>
<tr>
<th>TABLE I. National Trauma Data Bank Studied Comorbidities and Complications</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comorbidities</strong></td>
</tr>
<tr>
<td>No NTDS Comorbidity Present</td>
</tr>
<tr>
<td>Alcoholism</td>
</tr>
<tr>
<td>Ascites Within 30 Days</td>
</tr>
<tr>
<td>Bleeding Disorder</td>
</tr>
<tr>
<td>Chemotherapy for Cancer Within 30 Days</td>
</tr>
<tr>
<td>Congestive Heart Failure</td>
</tr>
<tr>
<td>Congestive Heart Failure</td>
</tr>
<tr>
<td>Current Smoker</td>
</tr>
<tr>
<td>Currently Requiring or on Dialysis</td>
</tr>
<tr>
<td>CVA/Residual Neurological Deficit</td>
</tr>
<tr>
<td>Diabetes Mellitus</td>
</tr>
<tr>
<td>Disseminated Cancer</td>
</tr>
<tr>
<td>Do not Resuscitate Status</td>
</tr>
<tr>
<td>Esophageal Varices</td>
</tr>
<tr>
<td>Functionally Dependent</td>
</tr>
<tr>
<td>Health Status</td>
</tr>
<tr>
<td>History of Angina within Past 1 Month</td>
</tr>
<tr>
<td>History of MI within the Past 6 Months</td>
</tr>
<tr>
<td>History of Revascularization/Amputation for PVD</td>
</tr>
<tr>
<td>Hypertension Requiring Medication</td>
</tr>
<tr>
<td>Impaired Sensorium</td>
</tr>
<tr>
<td>Prematurity</td>
</tr>
<tr>
<td>Obesity</td>
</tr>
<tr>
<td>Respiratory Disease</td>
</tr>
<tr>
<td>Steroid Use</td>
</tr>
<tr>
<td>Unplanned Intubation</td>
</tr>
</tbody>
</table>

NTDB, National Trauma Data Bank; CVA, cerebrovascular accident; MI, myocardial infarction; PVD, peripheral vascular disease; ARDS, acute respiratory distress syndrome; CPR, cardiopulmonary resuscitation; DVT = deep vein thrombosis.
factor with American Indians and Asians having odds ratios of 0.5 and 0.7 (p value <0.02) when compared to white patients for developing a DVT. Black race was not a significant risk factor for developing a VTE compared to whites with an odds ratio of 1 and a p value of 0.99. Increasing age was also associated with developing a VTE with a 1 to 2% increased risk per additional year of life on presentation to the emergency department (p value <0.01).

With regard to vitals on arrival, a patient with hypotension (systolic blood pressure <90 mm Hg) had an odds ratio of 1.18 (confidence interval [CI] = 1.01, 1.39) compared to a normotensive patient. Worsening tachycardia and GCS were also associated with an increased risk of a VTE event. However, a worsening GCS score was more predictive of a DVT than a PE, as is indicated by a near 70% increased risk of developing a DVT compared to a PE in the most severely head injured patients (GCS score of <4). In this patient population, the CIs did not overlap. The odds ratio of developing a DVT was 2.05 (CI = 1.86, 2.26) compared to 1.37 (CI = 1.14, 1.64) in patients diagnosed with a PE, suggesting that the risk is different between the 2 groups in the most severely head injured patients.

Remaining risk factors for each patient population are listed in Table IV. There were 19 and 5 risk factors that were associated with developing a DVT and PE, respectively. Mechanism of injury is listed as separate risk factor for each group, as a different mechanism was either predictive or protective for a PE or DVT. Relative to blunt injury, a penetrating injury was associated with an odds ratio of 1.64 (CI = 1.46, 1.85) for being diagnosed with a DVT, and it was not predictive for developing a PE.

All comorbidity and complications associated with developing a DVT had an increased odds ratio except for patients who were designated with a do not resuscitate code status. This risk factor was associated with a decreased odds ratio of 0.21 (CI = 0.08, 0.58). With regard to the comorbidities of chemotherapy, steroid use, bleeding disorders, dialysis,

### TABLE III. Demographic and Presenting Emergency Department Clinical Characteristics (e.g., Vitals and Glasgow Coma Scale Score) Risk Factors Associated with Developing a Deep Vein Thrombosis, Pulmonary Emboli, or Both. All Risk Factors Were Subjected to a Multivariate Logistic Regression. p ≤ 0.02 for All Risk Factors Except for Hypotension, for Which p = 0.04

<table>
<thead>
<tr>
<th>Nature of Risk Factor</th>
<th>Risk Factor</th>
<th>DVT</th>
<th>95% CI</th>
<th>PE</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>OR</td>
<td>Lower</td>
<td>Upper</td>
<td>OR</td>
</tr>
<tr>
<td>Demographics</td>
<td>Age</td>
<td>1.02</td>
<td>1.01</td>
<td>1.02</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>0.68</td>
<td>0.63</td>
<td>0.73</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>White</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>American</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Indian</td>
<td>0.48</td>
<td>0.28</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>0.69</td>
<td>0.51</td>
<td>0.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Black</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other Race</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systolic Blood Pressure</td>
<td>Normotensive (SBP 80–130 mm Hg)</td>
<td>Ref</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hypotensive (SBP &lt;80 mm Hg)</td>
<td>1.18</td>
<td>1.01</td>
<td>1.39</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hypertensive (SBP &gt;130 mm Hg)</td>
<td>0.89</td>
<td>0.83</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>Heart Rate</td>
<td>Normal Pulse (60–100 bpm)</td>
<td>Ref</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tachycardia (100–120 bpm)</td>
<td>1.23</td>
<td>1.13</td>
<td>1.33</td>
<td>1.25</td>
</tr>
<tr>
<td></td>
<td>Severe Tachycardia (&gt;120 bpm)</td>
<td>1.36</td>
<td>1.23</td>
<td>1.51</td>
<td>1.29</td>
</tr>
<tr>
<td>GCS Score</td>
<td>GCS Score 13–15</td>
<td>Ref</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GCS Score 9–12</td>
<td>1.67</td>
<td>1.44</td>
<td>1.95</td>
<td>1.53</td>
</tr>
<tr>
<td></td>
<td>GCS Score 5–8</td>
<td>2.05</td>
<td>1.76</td>
<td>2.39</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td>GCS Score &lt;4</td>
<td>2.05</td>
<td>1.86</td>
<td>2.26</td>
<td>1.37</td>
</tr>
</tbody>
</table>

DVT, deep vein thrombosis; PE, pulmonary emboli; bpm, beats per minute; OR, odds ratio; SBP, systolic blood pressure; GCS, Glasgow Coma Scale.
### TABLE IV. Risk Factors That Are Associated with Developing a Deep Vein Thrombosis, Pulmonary Emboli, or a DVT or Both Simultaneously. All Risk Factors Have Been Subjected to Multivariate Logistic Regression

<table>
<thead>
<tr>
<th>DVT Risk Factors</th>
<th>Risk Factor</th>
<th>OR</th>
<th>95% CI</th>
<th>OR</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
<td></td>
<td>Upper</td>
</tr>
<tr>
<td>Severe Injury</td>
<td>Severe Head</td>
<td>2.41</td>
<td>2.23-2.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Injury (AIS ≤3)</td>
<td></td>
<td></td>
<td>1.61</td>
<td>1.22-2.11</td>
</tr>
<tr>
<td></td>
<td>Severe Neck</td>
<td></td>
<td></td>
<td>1.23</td>
<td>1.02-2.53</td>
</tr>
<tr>
<td></td>
<td>Injury (AIS ≤3)</td>
<td></td>
<td></td>
<td>1.61</td>
<td>1.31-2.71</td>
</tr>
<tr>
<td></td>
<td>Chemotherapy Within 30 Days</td>
<td>2.31</td>
<td>1.2-4.45</td>
<td>1.49</td>
<td>1.05-2.13</td>
</tr>
<tr>
<td></td>
<td>Steroid Use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bleeding Disorder</td>
<td>1.5</td>
<td>1.31-1.71</td>
<td>1.5</td>
<td>1.05-2.13</td>
</tr>
<tr>
<td></td>
<td>Dialysis Use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CVA/Residual Neurological Deficit</td>
<td>1.39</td>
<td>1.17-1.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Do Not Resuscitate Status</td>
<td></td>
<td></td>
<td>0.21</td>
<td>0.08-0.58</td>
</tr>
<tr>
<td></td>
<td>Complications</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abdominal Compartment Syndrome</td>
<td>3.2</td>
<td>2.21-4.65</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Graft/Prosthesis/Flap Failure</td>
<td>2.76</td>
<td>1.63-4.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Organ/Deep SSI</td>
<td>2.4</td>
<td>1.96-2.93</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Superficial SSI</td>
<td>2.53</td>
<td>2-3.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Decubitus Ulcer</td>
<td>2.17</td>
<td>1.9-2.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Extremity Compartment Syndrome</td>
<td>1.92</td>
<td>1.44-2.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wound Disruption</td>
<td>1.49</td>
<td>1.06-2.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Myocardial Infarction</td>
<td>1.45</td>
<td>1.11-1.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bleeding Complication</td>
<td>1.41</td>
<td>1.12-1.78</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Acute Renal Failure</td>
<td>1.18</td>
<td>1.01-1.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mechanism of Injury</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Blunt</td>
<td>1.64</td>
<td>1.46-1.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Penetrating</td>
<td>0.76</td>
<td>0.53-1.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Burns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comorbidities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe Injury</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Superficial Injury (AIS ≤3)</td>
<td></td>
<td></td>
<td>2.61</td>
<td>1.27-5.35</td>
</tr>
<tr>
<td></td>
<td>Unplanned Intubation</td>
<td></td>
<td></td>
<td>2.02</td>
<td>1.33-3.09</td>
</tr>
<tr>
<td></td>
<td>Cardiac Arrest with CPR</td>
<td></td>
<td></td>
<td>2.45</td>
<td>1.87-3.2</td>
</tr>
<tr>
<td></td>
<td>Decubitus Ulcer</td>
<td></td>
<td></td>
<td>1.41</td>
<td>1.1-1.82</td>
</tr>
<tr>
<td></td>
<td>Mechanism of Injury</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Blunt</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Penetrating</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Burns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Shared Risk Factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe Injury</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Severe Chest Injury (AIS ≤3)</td>
<td>1.74</td>
<td>1.61-1.88</td>
<td>1.27</td>
<td>1.13-1.43</td>
</tr>
<tr>
<td></td>
<td>Severe Abd/Pelvis Injury (AIS ≤3)</td>
<td>1.94</td>
<td>1.76-2.15</td>
<td>1.27</td>
<td>1.13-1.43</td>
</tr>
<tr>
<td></td>
<td>Severe Spine Injury (AIS ≤3)</td>
<td>1.92</td>
<td>1.73-2.13</td>
<td>1.27</td>
<td>1.13-1.43</td>
</tr>
<tr>
<td></td>
<td>Severe LE Injury (AIS ≤3)</td>
<td>2.17</td>
<td>2.02-2.34</td>
<td>1.27</td>
<td>1.13-1.43</td>
</tr>
<tr>
<td></td>
<td>Severe UE Injury (AIS ≤3)</td>
<td>1.74</td>
<td>1.61-1.88</td>
<td>1.27</td>
<td>1.13-1.43</td>
</tr>
<tr>
<td></td>
<td>Comorbidities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hypertension</td>
<td>1.4</td>
<td>1.29-1.52</td>
<td>1.4</td>
<td>1.29-1.52</td>
</tr>
<tr>
<td></td>
<td>Obesity</td>
<td>2.2</td>
<td>1.92-2.53</td>
<td>2.2</td>
<td>1.92-2.53</td>
</tr>
<tr>
<td></td>
<td>Impaired</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sensorium</td>
<td>1.27</td>
<td>1.14-1.43</td>
<td>1.27</td>
<td>1.14-1.43</td>
</tr>
<tr>
<td></td>
<td>Current Smoker</td>
<td>1.28</td>
<td>1.12-1.47</td>
<td>1.28</td>
<td>1.12-1.47</td>
</tr>
<tr>
<td></td>
<td>Alcoholism</td>
<td>1.13</td>
<td>1.02-1.26</td>
<td>1.13</td>
<td>1.02-1.26</td>
</tr>
<tr>
<td></td>
<td>Complications</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pneumonia</td>
<td>2.4</td>
<td>2.2-2.61</td>
<td>1.91</td>
<td>1.41-2.59</td>
</tr>
<tr>
<td></td>
<td>ARDS</td>
<td>1.3</td>
<td>1.17-1.45</td>
<td>1.3</td>
<td>1.17-1.45</td>
</tr>
<tr>
<td></td>
<td>Stroke/CVA</td>
<td>1.91</td>
<td>1.41-2.59</td>
<td>1.31</td>
<td>1.14-1.5</td>
</tr>
<tr>
<td></td>
<td>Systemic Sepsis</td>
<td>1.31</td>
<td>1.14-1.5</td>
<td>1.67</td>
<td>1.33-2.09</td>
</tr>
<tr>
<td>Divergent Risk Factor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Base Deficit</td>
<td>1.3</td>
<td>1.07-1.58</td>
<td>0.63</td>
<td>0.41-0.97</td>
</tr>
</tbody>
</table>

DVT, deep vein thrombosis; PE, pulmonary emboli; AIS, abbreviated injury scale; ARDS, acute respiratory distress syndrome; CVA, cerebrovascular accident; LE, lower extremity; UE, upper extremity; SSI, surgical site infection; CPR, cardiopulmonary resuscitation. \( p \leq 0.01 \) for all risk factors except for steroid use, acute renal failure and bleeding \( (p = 0.04) \), dialysis use \( (p = 0.03) \), wound disruption, and mechanism of injury for PE \( (p = 0.02) \). \( p \leq 0.01 \) for all risk factors except for alcoholism \( (p = 0.02) \).
and history of stroke, these were all associated with an increased risk of DVT and not PE. There were no comorbidities that were associated with developing a PE and not a DVT. Compartment syndromes and wound complications accounted for over half the complications that were associated with patients developing a DVT. With regard to PEs, severe superficial injuries, cardiac arrest with cardiopulmonary resuscitation (CPR), and unplanned intubation were all strongly associated with an increased risk of developing a PE and not a DVT.

Of all the risk factors studied, there was one that had diverging odds ratios. Base deficit was associated with an increased risk of developing a DVT and a decreased risk for developing a PE. In addition, there was a risk factor that was shared between the two groups that had odds ratios with CIs that did not overlap. Decubitus ulcer was shown to have an odds ratio of 2.17 (CI = 1.9, 2.47) in patients with a DVT, and 1.41 (CI = 1.10, 1.82) in patients with a PE. This was considered a separate risk factor for patients who were diagnosed with a DVT versus a PE as the nonoverlapping CIs are statistically significantly different.

**DISCUSSION**

Although there are many overlapping risk factors for patients who are either diagnosed with a PE or DVT, there are independent risk factors that are not shared, and one risk factor with diverging odds ratios, which suggests that the two disease processes are different. Examining the groups separately shows trends that may explain how each group could be a different disease process.

Patients diagnosed with deep vein thrombosis appear to have the clinical picture of immobility. These patients have risk factors of presenting with a severe head injury (AIS score of ≥3), a GCS of <4, compartment syndromes (e.g., abdominal and extremity), and decubitus ulcers, all of which can be associated with immobility. In contrast, patients diagnosed with PE alone are more likely to have cardiac arrest with CPR, and severe superficial injury (AIS score of ≥3) that includes inhalational injuries. These two risk factors suggest a process more closely related to a chest injury.

This is supported by previous work with severe chest injuries being 40% more associated with a PE than a DVT in a previously published NTDB study. In addition, a recently published study by Van Gent et al described patients with chest wall injuries and a higher number of rib fractures and pulmonary contusions being more associated with PE than DVT.

In addition to severe chest injury, the previous NTDB study also noted a low overlap in the number of patients diagnosed with a PE and DVT. The lack of overlap between the two groups has been noted in multiple retrospective and prospective clinical studies, with a reported rate of 0 to 22% of patients who were diagnosed with a PE and a concomitant DVT. In one recent study, the authors noted that of the 46 patients identified on retrospective review who underwent a CT angiogram of the chest to diagnose a PE, followed by a CT venogram of the pelvis and lower extremities, only seven patients had a DVT.

After ruling out the lack of source of clot from the extremities in a previous study, the authors suspected that PE could be a presentation of a primary hypercoagulable event rather than an embolized thrombus. They hypothesized that a primary pulmonary thrombosis was due to autonomic dysfunction leading to adrenergia and inflammation followed by local vasospasm and eventual thrombosis of the vessel. Although the source of the local autonomic dysfunction was not noted in the article, the data from this study and previous studies suggest that severe chest trauma could be a contributing factor.

Another possible cause of pulmonary artery thrombosis not related to remote DVTs relates to thrombus being introduced from an external source. Historically, blood products were known to be a source of microthrombi. Thrombi would form as a result of the clotting of residual platelets, and white blood cells in the plasma of the blood product (e.g., liquid red blood cells, platelets). Before the use of a filter to remove these thrombi, patients would often have respiratory compromise after the transfusion of blood and PEs were often suspected. Considering that blood transfusions are a known risk factor for VTE, it is unclear how many current filters decrease the rate or completely eliminate the transfusion of microthrombi.

Blood transfusions or other means of resuscitation may explain the sole diverging odds ratio of a base deficit among patients with a VTE. Having a base deficit was associated with an odds ratio of 1.3 in patients with a DVT and 0.63 in patients with a PE. Although this could be considered in conflict with the proposed hypothesis of blood transfusions being a culprit for primary pulmonary thrombosis, this could be due to the way the patients were organized in the model and how base deficit was defined in this dataset. The exact reasoning for the association is not known.

Despite the lack of complete overlap of independent risk factors between the two groups, there were many shared statistically significant risk factors that were present in this study (Tables III and IV), most of which have been previously described in current literature. Patients who are older, male, smokers, obese, or have pneumonia have a higher association with developing a VTE. In trauma patients, traumatic brain injury, severe spine, chest, pelvic, and extremity fractures have also been previously described as risk factors for VTE. As stated in the methods, this study was not limited to known risk factors for VTE as it would have prevented discovery of new risk factors for PE and/or DVTs. The fact that the statistically significant overlapping risk factors for PE and DVTs are consistent with previously published data implies validation of the model developed for this study.

The clinical significance of trying to decipher if PEs could be a primary event is contingent on whether or not a...
patient would subsequently receive therapeutic anticoagulation or an IVC filter after diagnosis. VTEs are a risk factor in severely injured trauma patients, especially in combat soldiers, who likely have another injury with an increased risk of bleeding (e.g., traumatic brain injury, solid organ injury). IVC filters are often placed in these patients to thwart the bleeding risk of anticoagulation, and possible propagation of a DVT into a PE despite the thrombogenic associated risk.\(^3^3\)

If thrombus within the pulmonary arteries can occur primarily, an IVC filter and therapeutic anticoagulation along with their associated risks may not be indicated.

In addition to the clinical significance of primary pulmonary artery thrombosis versus a true PE, there is also a public policy component to the management of the diagnosis. Many governmental agencies, including the Agency for Healthcare Research and Quality and Medicare/Medicaid services consider VTEs as a potential measure of quality, and PE as the most common hospital cause of death.\(^7\) In certain patient populations, PEs and DVTs are considered a never event and are not reimbursed by Centers for Medicare and Medicaid Services.\(^8\) Although, there is no question that DVTs can lead to PEs or at the very least they are a risk factor for developing them, it is unclear how much they contribute to the overall incidence of pulmonary artery thrombosis. Further investigation into the pathophysiology is warranted to not only better understand and treat VTE disease, but to also help guide public policy regarding its management.

The greatest limitation of this study is the retrospective blinded design of the NTDB. Review of patient records for further information to help elucidate more specific disease risk factors for PEs and DVTs is not available. Nor is there access to individual hospital policies regarding VTE screening, prophylaxis and treatment algorithms to examine the more unique risk factors discovered in this study. Detection bias could also play a role in this study. For example, patients requiring CPR who have an unplanned intubation are more likely to receive imaging and are, therefore, more likely to be diagnosed with a PE and less likely to be diagnosed with a DVT. Future studies would benefit from prospectively collected data examining the incidence, and examination of clot location and burden to better understand the disease process and address these concerns.

Another limitation of the study is that 50% of the original number of patients in the cohort was eliminated from the study due to exclusion criteria and incomplete data. Despite the NTDB being the largest trauma registry in the United States and the world, it is limited by the quality of data submitted by each institution. An attempt to control for data quality resulted in the smaller study population. Nonetheless, this resulted in one of the strengths of this study in that all comorbidities and complications recorded by the NTDB were examined as risk factors and were controlled for confounding variables.

**CONCLUSION**

Despite the limitations of the study, mainly the retrospective nature and inability to survey diagnostic modality for VTEs, it shows trends that in conjunction with previously published work that PEs and DVTs may not always be related. In this study of the NTDB cohort between 2007 and 2010, despite patients with a DVT and PE having overlapping risk factors they have many different independent ones. These data further support that PE and DVTs may not always be the same disease process, and draws into question using PEs as a quality reimbursement metric, and whether to consistently treat them with either therapeutic anticoagulation or IVC filters.

**ACKNOWLEDGMENT**

We thank Dr. Brian Diggs for his assistance with statistical model development and management of the National Trauma Data Bank.

**REFERENCES**

Pulmonary Emboli and Deep Vein Thromboses

Dose Responses of Ibuprofen In Vitro on Platelet Aggregation and Coagulation in Human and Pig Blood Samples

Wenjun Z. Martini, PhD*; Cassandra M. Rodriguez, BS*; Rodolfo Deguzman, MT*; Jessica B. Guerra, BS*; Angela K. Martin, BS†; Anthony E. Pusateri, PhD‡; LTC Andrew P. Cap, MC USA*; Michael A. Dubick, PhD*

ABSTRACT

Introduction: Ibuprofen is commonly used by warfighters in the deployed environment. This study investigated its dose effects on in vitro coagulation in human and pig blood. Methods: Blood samples were collected from 6 normal volunteers and 6 healthy pigs and processed to make platelet-adjusted samples (100 × 10^7/μL, common transfusion trigger in trauma). Ibuprofen was added to the samples at concentrations of 0 μg/mL (control), the concentration from the highest recommended oral dose (163 μg/mL, 1×), and 2×, 4×, 8×, 10×, 12×, 16×, and 20×. Platelet aggregation by Chrono-Log aggregometer and coagulation by rotational thrombelastogram (Rotem) were assessed at 15 minutes after the addition of ibuprofen. Results: A robust inhibition of ibuprofen on arachidonic acid-induced platelet aggregation was observed at all doses tested in human or pig blood. Collagen-stimulated platelet aggregation was inhibited starting at 1× in human blood and 4× in pig blood. Rotem measurements were similarly compromised in pig and human blood starting at 16×, except clot formation time was prolonged at 1× in human blood (all p < 0.05). Conclusion: Ibuprofen inhibited platelet aggregation at recommended doses, and compromised coagulation at higher doses. Human blood was more sensitive to ibuprofen inhibition. Further effort is needed to investigate ibuprofen dose responses on coagulation in vivo.

INTRODUCTION

As a nonsteroidal anti-inflammatory drug (NSAID), ibuprofen is a commonly used analgesic and antipyretic. Although ibuprofen is well tolerated and safe at its recommended doses (400 mg–800 mg and 6–11 mg/kg), its widespread availability over the counter has increased the potential for accidental ingestion and misuse. Indeed, overdose (>800 mg) as high as 54 g (770 mg/kg, 70×) and 100 g (1428 mg/kg, 125×) have been reported in civilian hospitals. Ibuprofen overdose is associated with a wide range of toxic and adverse effects, including coma,2–5 seizures,6 metabolic acidosis,7 liver injury, gastrointestinal disturbances,3,5 acute renal failure,1 thrombocytopenia,3 and death.4,7 In the military, recent reports from military systems have revealed previously unrecognized overuse of NSAIDs and acetaminophen due to self-treatment of injuries sustained in combat environments,8,9 where NSAID overuse increased 40% annually from 2004 to 2008 and junior enlisted Soldiers are 6 times more likely to overdose than officers.9 Thus, there is concern that overuse of NSAIDs will predispose deployed warfighters on the battlefield to increased bleeding risk.

As the concept of damage control resuscitation has been introduced for treatment of injured warfighters, more attention has been directed at prevention of risk factors to hemostasis and coagulation. Previous reports have shown that ibuprofen can reduce platelet aggregation,10–12 but its dose responses on platelet aggregation and global measures of hemostasis such as rotational thrombelastogram (Rotem [TEM, Munich, Germany]) remain undefined. In addition, the swine model has been considered the animal model of choice to both investigate the pathophysiology of trauma-related hemorrhagic shock and search for new treatments for such injuries.13–19 However, it is unclear whether there are different responses to ibuprofen inhibition in platelets and coagulation between human blood and swine blood. As methodology development, we used blood samples from 4 pigs to test the dose responses of ibuprofen in vitro.20 As a follow-up, the present study compared dose-response effects on coagulation function and platelet aggregation in blood samples taken from 6 pigs and 6 human subjects.

MATERIALS AND METHODS

Human Blood Collection

This study was conducted under a protocol reviewed and approved by the U.S. Army Medical Research and Materiel Command Institutional Review Board and in accordance with the approved protocol. Native whole blood (NWB) samples were withdrawn via venous puncture from 6 normal volunteers with signed informed consent. Each participant was sampled once for a total of 100 mL blood. Exclusion criteria...
including pregnancy, ongoing therapeutic anticoagulation, and use of over-the-counter drugs such as aspirin, ibuprofen, herbal products, or NSAIDs within the previous 7 days.

**Preparation of Human Platelet Adjusted Whole Blood Samples**

Blood samples from each subject were collected into 22 citrate tubes (4.5 mL blue top containing 3.2% Na Citrate, Becton-Dickenson, Franklin Lakes, New Jersey), pooled and allowed to incubate for 15 minutes at room temperature. An aliquot of the NWB samples was used for measurements of complete blood cell counts (ABX Pentra 120 Hematology Analyzer, ABX Diagnostics, Inc., Irvine, California), including platelet counts and hematocrit (Hct). The remaining NWB samples were divided into two portions: one was reserved as NWB samples, and the other was processed to make platelet adjusted whole blood samples (PAWBs), following procedures developed at our Institute. Briefly, NWB samples were centrifuged at 2000× g for 15 minutes to separate platelet poor plasma (top layer), the buffy coat (middle layer, containing platelets), and red cells (bottom layer). Platelet poor plasma was first collected via aspiration from the top clear layer and red blood cells (RBCs) were collected on removal of the buffy coat layer. By mixing platelet poor plasma with collected RBCs (volume/volume) to achieve the same hematocrit level measured in the NWB samples, a stock of platelet poor whole blood was made. Afterward, NWB samples and platelet poor whole blood samples were mixed in appropriate amounts to make a blood sample that contained a platelet count of 100 × 10^3/μL, referred to as PAWBs. Since a platelet level of 100 × 10^3/μL was considered the critical level for platelet transfusion in trauma patients, this level was selected to assess the dose-response effects of ibuprofen in this study. The PAWBs were then divided into 8 tubes to test the dose responses of ibuprofen. The entire study was repeated in 6 separate experiments.

**Dosing of Ibuprofen**

The Food and Drug Administration-approved standard dose for Ibuprofen is up to 800 mg every 6 hours. Pharmacokinetic data indicate that an oral dose can be nearly 100% absorbed. Therefore, assuming an average person of 70 kg has blood volume of 70 mL/kg, then the estimated plasma concentration from the highest recommended dose would be 163 μg/mL. In the dose set of 8 tubes containing PAWBs, ibuprofen (Caldolor, [IV solution, 100 mg ibuprofen/mL]; Nashville, Tennessee) was added to reach the concentrations of 0 μg/mL (control), 163 μg/mL (the concentration from the highest recommended oral dose, referred to as the dose of 1×), and 2×, 4×, 8×, 10×, 12×, 16×, and 20×, respectively. Buffered blood bank saline (Isotonic solution 0.85% w/v, Thermo Scientific, Waltham, Massachusetts) was used for volume matching among the 8 tubes.

On the completion of dosing and volume matching, the Caldolor (ibuprofen)-treated pig or human blood sample in each tube was aliquoted for 3 measurements: (1) platelet aggregation using the Chrono-Log 700 aggregometer (Chrono-log, Havertown, Pennsylvania); (2) prothrombin time (PT), activated partial thromboplastin time (aPTT) and fibrinogen concentration by STA-R (Diagnostica Stago, Rue des Freres Chausson, France); and (3) thrombelastogram by Rotem (TEM). All measurements were made at 37°C.

**Platelet Aggregation**

Platelet impedance aggregometry was assessed in platelet-adjusted whole blood samples using a Chrono-Log 700 aggregometer. At 15 minutes after the addition of ibuprofen, the aggregation was stimulated with either collagen (2 μg/mL) or arachidonic acid (0.5 mM) as agonists. The area under the curve was used to compare changes of platelet aggregation.

**Thrombelastogram Rotem**

Coagulation profile from PAWBs was measured using Rotem. Three-hundred microliters of dosed blood samples were added to the measurement cup followed by addition of Extrem reagents, containing recombinant tissue factor and phospholipids. From the clotting curve tracing, the following parameters were generated to represent coagulation profiles: coagulation time (CT, the time from test start to an amplitude of 2 mm); clot formation time (CFT, the time between 2 and 20 mm amplitude); α-angle (angle between the baseline and a tangent to the clotting curve through the 2 mm point to represent clot rapidity); and maximum clot firmness (MCF, the maximum amplitude reached during the test to represent clot strength).
Statistical Analysis

Data were expressed as means ± standard error of the mean from 6 separate experiments and analyzed using SAS statistical software (Cary, North Carolina). After conduct of a one-way analysis of variance, Dunnett’s test was performed on the human and pig data to compare responses to the control (0 dose). A two-way analysis of variance with repeated measures in coordination with the post hoc Tukey’s test were conducted to compare the human and pig data at corresponding doses and to compare responses among the various dose levels. The statistically significant level was set at \( p < 0.05 \).

RESULTS

Blood Characteristics

As designed, the platelet count in PAWBs was reduced to about 100 × 10^3/μL, with no differences in RBC, Hct, hemoglobin, or fibrinogen concentration between NWB and PAWB samples (Table I).

Platelet Aggregation

Platelet aggregation was assessed in PAWBs using collagen or arachidonic acid as the agonist. In pig PAWBs, ibuprofen significantly inhibited collagen-induced platelet aggregation starting at the 4× dose (Fig. 1A). At 4×, 10×, and 20× of ibuprofen doses, collagen-induced platelet aggregation was reduced to 71% ± 5%, 45% ± 5%, and 10% ± 2% of the control (0 dose) value, respectively (all \( p < 0.05 \), Fig. 1A). In human PAWBs, ibuprofen inhibited collagen-induced platelet aggregation starting at the recommended dose (1×). At 4×, 10×, and 20× of ibuprofen doses, collagen-induced platelet aggregation was reduced to 45% ± 6%, 26% ± 4%, and 3% ± 2% of the control (0 dose) value, respectively (all \( p < 0.05 \), Fig. 1A). Compared to that occurring in pig PAWBs, a more sensitive inhibition of collagen-induced aggregation was observed in human vs. porcine PAWBs as demonstrated by the greater percent decreases in platelet aggregation at multiple doses of ibuprofen (Fig. 1A).

A robust inhibition of ibuprofen on arachidonic acid-induced platelet aggregation was observed at all doses tested in both pig and human PAWBs. At 1× and 2× of ibuprofen doses, arachidonic acid-induced platelet aggregation decreased to 8% ± 6% and 4% ± 3% of the control value in pig PAWBs, respectively (both \( p < 0.05 \), Fig. 1B), and to 20% ± 3% and 2% ± 2% of the control value in human PAWBs, respectively (both \( p < 0.05 \), Fig. 1B).

PT, aPTT, and Thromboelastogram Measurements

The effects of ibuprofen on PT and aPTT were assessed in whole blood samples. At all doses tested, PT did not change in pig PAWBs. In human PAWBs, a prolonged PT was observed at 20× (Table II). However, aPTT was prolonged at 16× and 20× in pig and human PAWBs.

Clotting initiation time (CT) was prolonged at ibuprofen doses of 16× and 20× in pig and human PAWBs (Fig. 2).

TABLE I. Blood Measurements From Samples Taken From 6 Normal Volunteers and 6 Healthy Pigs. Data Are Expressed as Means ±SE From 6 Separate Experiments

<table>
<thead>
<tr>
<th></th>
<th>RBC (10^3/μL)</th>
<th>Hct (%)</th>
<th>Hgb (g/dL)</th>
<th>Platelet (10^3/μL)</th>
<th>Fibrinogen (mg/dL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human NWB</td>
<td>4.3 ± 0.1</td>
<td>38.1 ± 1.2</td>
<td>12.9 ± 0.4</td>
<td>239 ± 24</td>
<td>370 ± 12</td>
</tr>
<tr>
<td>Human PAWB</td>
<td>4.2 ± 0.1</td>
<td>36.9 ± 1.3</td>
<td>12.8 ± 0.5</td>
<td>96 ± 2*</td>
<td>364 ± 12</td>
</tr>
<tr>
<td>Pig NWB</td>
<td>5.2 ± 0.2</td>
<td>26.3 ± 1.0</td>
<td>8.8 ± 0.3</td>
<td>335 ± 26</td>
<td>265 ± 13</td>
</tr>
<tr>
<td>Pig PAWB</td>
<td>5.2 ± 0.2</td>
<td>26.2 ± 1.3</td>
<td>8.6 ± 0.4</td>
<td>96 ± 5*</td>
<td>243 ± 9</td>
</tr>
</tbody>
</table>

Hct, hematocrit; Hgb, hemoglobin; NWB, native whole blood samples; PAWB, platelet-adjusted whole blood samples; RBC, red blood cells. *\( p < 0.05 \) PAWB vs. NWB.
with no statistically significant species difference seen. In contrast and similar to observations in platelet aggregation, a more sensitive inhibition from ibuprofen was observed in human blood on CFT. CFT was prolonged \( (p < 0.05) \) by ibuprofen starting at the recommended dose in human PAWBs (Fig. 2), but did not change in pig PAWBs until ibuprofen doses reached 16× and 20× (Fig. 2). Clotting rapidity (alpha) was decreased and clot strength (MCF) reduced at ibuprofen doses of 16× and 20× in both pig and human PAWBs (Fig. 3). Alpha and MCF were generally lower in human than pig blood up to 12×.

### DISCUSSION

Primary hemostasis involves a series of complex interactions of platelet adhesion and aggregation, thrombin generation, and ultimately hemostatic plug formation. Optimal platelet activation during this process is dependent on the synthesis of thromboxane A2 (TxA2) from prostaglandin H2. Prostaglandin H2 generation from arachidonic acid is catalyzed by cyclooxygenase. The antiplatelet effects of various NSAIDs, including ibuprofen, result from the inhibition of cyclooxygenase -1 activity and reduction of TxA2 synthesis.

In this study, we observed robust inhibition of ibuprofen on
Dose Responses of Ibuprofen In Vitro on Platelet Aggregation and Coagulation

arachidonic acid-induced platelet aggregation. Consistently, the dose-dependent effects of ibuprofen on inhibition of TxA2 production have been reported in whole blood samples and in healthy humans. Starting at the standard dose of 800 mg in the present study, arachidonic acid-induced platelet aggregation was nearly abolished by ibuprofen in human and pig blood samples. These findings are consistent with ibuprofen’s mechanism of action and previous reports of ibuprofen inhibition on platelet aggregation in vitro, in experimental animals, and in healthy humans at therapeutic doses.

In addition, collagen-induced platelet aggregation decreased to about 50% at 4× of recommended ibuprofen dose in human blood samples. Considering that subendothelial collagen can be exposed at injury, our observations are relevant to possible bleeding diathesis in injured warfighters. This inhibition of ibuprofen on platelet aggregation appears to be clinically relevant as an ibuprofen dose of 1800 mg (2.25×) inhibited platelet aggregation and prolonged bleeding time from 2.6 ± 1.3 minutes to 4.4 ± 2.2 minutes. Moreover, lower doses (1200 and 1500 mg, 1.5× and 1.9×) caused low levels of gastrointestinal distress in some subjects and higher doses caused stool blood loss in some subjects. Altogether these data confirm that ibuprofen, especially when taken at multiples of recommended doses, compromises hemostatic capacity and will predispose soldiers to bleeding complications on the battlefield.

The significant inhibitory effects of ibuprofen on platelet aggregation, however, were not reflected in Rotem measurements in this study. Neither clotting initiation time, clotting speed, nor clot strength changed until doses reached 16× and above. This disparity between aggregation and Rotem data is due to differences in aggregation and Rotem methodologies. Rotem measurements are based on the addition of reagents to generate thrombin for clot formation. The generated thrombin initiates and amplifies clot formation. Thrombin can also offset the inhibition of ibuprofen on platelet aggregation. Packham et al. demonstrated that thrombin induces platelet aggregation when the TxA2 and adenosine diphosphate pathways are inhibited, confirming that thrombin causes platelet aggregation via different mechanisms. Thus, the inhibitory effect of ibuprofen on platelet function was overcome by stimulatory effects of thrombin generated in the Rotem device. It should be noted that maximal thrombin generation in the closed Rotem or thrombelastograph system, which maximally stimulates platelets, represents an important limitation of these tools in determining the contribution of platelet dysfunction to coagulation disorders.

Swine have been widely used as an animal model for hemorrhage and resuscitation and as we have mentioned, more attention to coagulation and platelet aggregation responses are being made in these studies to justify including both swine and human blood in the present evaluation. The pigs often have much higher platelet counts than humans. To normalize the responses to response to ibuprofen and to be able to compare between species, we standardized platelet level to 100 × 10^3/μL in both human and pig blood samples in this study. This level has become a standard practice at our institute and was used by us in other in vitro studies investigating effects of dilution on thrombelastograph parameters, as might occur during resuscitation from hemorrhage. This level is relevant in trauma as a possible transfusion trigger. For example, if a service member had recently taken ibuprofen, gets injured and receives fluid resuscitation, it is quite possible that their platelet level could drop to 100 × 10^3/μL. Therefore, we studied coagulation and platelet aggregation at that level as it could be relevant in an injured warfighter. In addition, we observed that there are different sensitivities to ibuprofen inhibition in human and pig blood samples: collagen-induced platelet aggregation was inhibited at the recommended dose (1×) in human blood, but not until 4× recommended doses in pig blood; CFT was prolonged at the recommended dose in human blood, but did not occur until 16× recommended doses in pig blood; and PT was prolonged at 20× in human blood, but no changes occurred at any dose in pig blood. These differences suggest that human blood may be more sensitive to ibuprofen inhibition on platelet aggregation and coagulation, possibly due to the hypercoagulability observed in pigs, which makes pigs more resistant to ibuprofen inhibition. Thus, although the swine model is considered to be appropriate for hemorrhagic shock research, results from this study indicate that ibuprofen inhibitory effects on platelet aggregation and coagulation are likely underestimated when using the swine model or swine blood samples for evaluations.

A wide dose range of ibuprofen was selected in this study to assess effects of ibuprofen on coagulation. At the dose of 20×, platelet aggregation is abolished and all coagulation measurements are adversely affected, suggesting the severity of ibuprofen at high dose. Consistently, development of thrombocytopenia and bleeding complications has been reported in patients after ibuprofen overdose ranging from 20 g (25×) to 100 g (125×). It is worthwhile mentioning that the systemic effects of ibuprofen were not included in the current in vitro study. Since the toxicities of coma, seizures, metabolic acidosis, acute renal failure, and acute liver injury and death have been observed in ibuprofen overdoses, the adverse effects at high doses of ibuprofen are likely more detrimental than shown in the present study.

In conclusion, we investigated dose responses of ibuprofen on hemostasis in pig and human blood samples in vitro. Ibuprofen inhibited platelet aggregation and prolonged CFT at standard recommended doses, and prolonged aPTT and compromised coagulation profile starting at 16 times the standard dose. Compared to pig blood, human blood appears to be more sensitive to ibuprofen inhibition on coagulation, suggesting the likely underestimation of ibuprofen effects from use of a swine model. Although it appears that proper use of ibuprofen should not put deployed troops at greater risk of bleeding should they be injured, the translation of...
these current observations to the in vivo situation remains to be confirmed.

ACKNOWLEDGMENTS

We appreciate the support received from the Laboratory Support Section at the U.S. Army Institute of Surgical Research in coagulation measurements. We thank Mr. John Jones for his assistance with statistical analysis and Dr Harold Klemcke for reviewing the manuscript. Funded by U.S. Army Medical Research and Materiel Command.

REFERENCES

In Trauma, Conventional ROTEM and TEG Results Are Not Interchangeable But Are Similar in Clinical Applicability

Sandro Rizoli, MD, PhD†; Arimie Min†; Adic Perez Sanchez, MD‡; Pang Shek, PhD†; Richard Grodecki, MD‡; Precilla Veigas, MSc†; Henry T. Peng, PhD†

ABSTRACT Background: There is growing interest in viscoelastic hemostatic assays rotational thromboelastometry (ROTEM) and thromboelastography (TEG) for trauma. Despite shared features, it is unknown whether their results are interchangeable and whether one is clinically superior in predicting mortality, blood transfusion, and diagnosing early trauma coagulopathy. Methods: We conducted a prospective observational study comparing equivalent ROTEM and TEG parameters. Severely injured patients expected to receive massive transfusion were included. Assays were performed simultaneously on admission and repeated over subsequent 12 hours. International normalized ratio ≥1.2 or fibrinogen <1 g/L defined coagulopathy. TEG used kaolin as coagulation initiator and ROTEM used tissue factor (conventional). Spearman nonparametric analysis and Bland–Altman difference mean plot revealed parameter association. Logistic regression and receiver operating characteristic curves measured predictive values. Results: 33 patients (74 ROTEM, 74 TEG) were included; 79% were male, mean Injury Severity Score was 23.5 ± 14, admission international normalized ratio was 1.33 ± 0.4, and 63.4% received blood transfusions. Overall, parameter agreement fell outside acceptable limits, with weak or no association. Clinically, ROTEM maximum clot firmness and TEG maximum amplitude showed reasonable predictive accuracy for mortality, strong accuracy for any or massive blood transfusion, reasonable for plasma transfusion and similar poor predictive accuracy for diagnosing coagulopathy. Conclusions: ROTEM and TEG results are not interchangeable, arguably due to different coagulation triggers. Assays had similar clinical performance.

INTRODUCTION

Trauma is the leading cause of death among civilian and military populations. Over 5.8 million people of all ages and economic groups die every year from unintentional injuries and violence. Hemorrhage, particularly when complicated by coagulopathy, is the most preventable cause of death. The recent advances in trauma resuscitation were born from the growing understanding of the early trauma coagulopathy (ETC). Tissue damage and shock initiate the ETC via activation of protein C, leading to systemic anticoagulation and fibrinolysis, which are worsened by continuing blood loss, hypothermia, acidosis, and dilution. ETC is complex and involves multiple and variable combinations of failures at different stages of the coagulation process. Including fibrinogen depletion, platelet dysfunction, lack of clotting factors, endothelial dysregulation, and neurohormonal derangements among others. In this complex scenario, the role of conventional laboratory tests like prothrombin time and partial thromboplastin time has been questioned. Their restricted evaluation of the hemostasis, long time to results and dissociation from bleeding and transfusion requirements are limitations often cited. Viscoelastic hemostatic assays (VHAs) such as rotational thromboelastometry (ROTEM) and thromboelastography (TEG) were recently proposed for the early diagnosis and management of traumatic bleeding and coagulopathy. A growing number of trauma studies have focus on these tests including a recent systematic review from our group. VHAs evaluate the viscoelastic properties of coagulation in whole blood under low shear conditions, better reflecting the novel concepts of cell-based hemostasis instead of the classical coagulation cascade partitioned in intrinsic and extrinsic pathways. They provide a global and functional assessment of coagulation, from clot initiation to amplification/propagation and lysis. VHA can be done as point of care and the results immediately available. Positive tests correlate well with bleeding and may guide the clinical decisions to transfuse.

A question that remains mostly unanswered is whether the results obtained by TEG and ROTEM are similar (interchangeable), in particular in the trauma setting. The devices share the same fundamental principles and many common features, and at a preliminary evaluation, appear to differ only in complexity and aspects of ease of use, in their purchase and running costs. It has been argued that the choice of device is mostly determined by the geographical location of the institution, with North American institutions acquiring the U.S.-produced TEG, whereas Europeans favour ROTEM for the same reason. A recent study on the magnitude of changes from baseline in hypercoagulable or hypocoagulable samples showed equivalence between TEG and ROTEM indicating comparable use of the instruments. In contrast, another study by Hagemo et al concluded that the TEG and
ROTEM results were not interchangeable, without indicating the possible reasons for the differences.21,22

Our own experience with the use of both devices in the trauma patients is that the results differ significantly. We then proposed a study on the interchangeability of the conventionally performed TEG and ROTEM. We also investigated whether one test would be superior to the other in predicting mortality, the need for blood transfusion, and diagnosing early trauma coagulopathy.

METHODS

Study Design

This was a prospective observational study conducted at a level I adult trauma center of the University of Toronto in Canada. It included adult (age >16 years) severely injured (injury severity score >15) patients admitted directly from the scene within 1 hour of the trauma with significant bleeding and probable coagulopathy. Significant bleeding was defined as a patient expected to receive massive transfusion based on the ABC score for massive transfusion (score ≥2).23,24 Massive transfusion was defined as the replacement of ≥10 units of red blood cells (RBCs) in the span of 24 hours. Coagulopathy was defined as an international normalized ratio (INR) ≥1.2 and/or fibrinogen <1 g/L. Patients with known acquired coagulopathy, not received directly from the injury scene, ≤15 years or ≤50 kg if age unknown, or pregnancy were excluded. ROTEM and TEG were performed simultaneously in the same patients within 30 minutes of admission and repeated when clinically indicated during the first 12 hours. The tests were done conventionally, according to the manufacturer’s instructions. For ROTEM the blood samples were collected in a BD Vacutainer sodium citrate 1.8 mL tube, and processed in the hospital core laboratory by trained technologists. For TEG, samples were collected in a MONOJECT sodium citrate 2.7 mL tube and processed by a research assistant in the trauma room. ROTEM delta system (TEM Systems, Durham, North Carolina) used tissue factor (conventional) for the EXTEN assay, added of cytochalasin D as a platelet inhibitor for its FIBTEM assay. TEG 5000 Analyzer (Haemoscope, Niles, Illinois) used kaolin activation. The results of the VHA were not available to the clinicians and none of the clinical decisions made were based on the ROTEM or TEG results. The study was approved by the hospital research ethics board and used exception from informed consent.

Statistical Analysis

Interchangeability was tested initially using the Spearman non-parametric analysis to evaluate the direction and strength of the correlation between equivalent ROTEM and TEG parameters (ROTEM clotting time [CT] vs. TEG reaction time [R]); Alpha (ROTEM) vs. Alpha (TEG); clot formation time (CFT) vs. kinetics time (K); ROTEM maximum clot firmness (MCF) vs. TEG maximum amplitude (MA); ROTEM lysis index (LI) at 30 minutes after CT (LI30) vs. TEG LI at 30 minutes after MA (CL30). The larger the Spearman coefficient, stronger is the correlation between the two values. A Spearman coefficient >0.8 is considered very strong, >0.6 is strong, >0.4 is moderate, >0.2 is weak, and <0.2 is very weak. Next, the Bland–Altman difference mean plot was used to evaluate association.25 The mean values (A) were plotted on the y-axis against the difference (D) on the x-axis and the closeness between the TEG and ROTEM variables was assessed at each specific parameter mean value. The limit of agreement (LoA) was defined as: D ± 1.96 × SD, where D represents the sample mean difference and SD represents the sample standard deviation of the differences. The expectation was that the relationship between the differences and means of the results attained from the two tests would be nonuniform. Initial tests indicated that a log transformation, which is recommended to address this issue,25 would not be sufficient. If a significant linear association between the differences and means was found, then bivariate linear regression, defined as: D = α + β × A, was used to calculate the estimated difference. This value was also used to calculate the corresponding LoA as proposed by Bland and Altman.26 Consequently, the LoA given in this situation was defined as: (α + β × A) ± 1.96 × SD, where SD represents the estimated standard deviation of the residuals. Predefined clinically acceptable LoA has been defined in the literature as ±10% of the average values between methods.17,27

The predictive measurements of clot strength (ROTEM EXTEM MCF and ROTEM FIBTEM MCF and TEG MA) were assessed using logistic regression and receiver operating characteristic (ROC) curves, which plot the sensitivity on the y-axis against 1-specificity on the x-axis. This depicts the true positive rate that corresponds to each false positive rate. C-statistic and 95% confidence intervals were then calculated, which represent the area under each ROC curve for ROTEM MCF and TEG MA for predicting mortality, coagulopathy, and blood transfusion. A C-statistic of >0.7 is considered reasonable and >0.8 is considered strong.28 In contrast, a C-statistic of 0.5 would indicate a predictive value of 1 in 2. P values were calculated to determine possible statistical significance between each of the predictive curves.

The predictive measurements of clot lysis (CL) (ROTEM LI30 and TEG CL30) were also compared in this study by using one cut point and logistic regression rather than the ROC curves used for MCF and MA. This was due to the skewed non-normal distribution of the LI30 and CL30 measurements.

All analyses were performed using SAS version 9.3 (SAS Institute, Cary, North Carolina) and verified by a qualified biostatistician. Two-tailed type 1 error rate of 0.05 was used as the threshold for statistical significance.

RESULTS

From September 2012 to June 2013, approximately 800 severely injured patients were screened and 33 patients were
enrolled in the study, having 74 TEG and 74 ROTEM tests simultaneously performed during the first 12 hours of hospital admission. Table I details the descriptive statistics of the patients included in this study. Mean age was 40 ± 20 years, 79% were male, mean Injury Severity Score was 23.5 ± 14, mean admission INR was 1.33 ± 0.4, and 63.4% received blood transfusions. In this cohort, approximately ½ of the patients suffered a blunt trauma, in contrast to our usual population of >¾ of the injuries being blunt.

To evaluate interchangeability of the equivalent parameters, scatter plots were used to compare the ROTEM results plotted on the y-axis against the TEG results on the x-axis as shown in Figure 1. The respective Spearman correlation coefficients are listed in Table II indicating the strength of the correlations. All parameters show a statistically significant correlation (p < 0.002) except CT/R (p = 0.17). The strongest correlation was found between MCF/MA (p = 0.65), whereas the weakest is seen between CT/R (p = 0.19).

Next we used the Bland–Altman difference mean plots to determine the agreement between the TEG and ROTEM parameters as shown in Figure 2. Standard Bland–Altman difference mean plots were used for LI30/CL30 due to their non-normal distribution. Table III shows the LoA calculated from the results as described in methods. Based on the predefined clinically acceptable LoA of 10% threshold of the mean values, for the present study it was calculated and defined as: CT/R 3.0, MCF/MA 5.5, Alpha (ROTEM)/Alpha (TEG) 6.3, CFT/K 7.4, and LI30/CL30 9.6. A significant linear association was found between the difference (D) and average (A) for CT/R, CFT/K, and alpha/angle, but no significant linear association was found between MA and MCF. However, most importantly, none of the LoA for any of the parameters fell within the predefined clinically acceptable LoA other than LI30/CL30. The LoA were the following: CT/R 6.97, MCF/MA 23.33, a-angle 25.50, CFT/K 8.35, and LI30/CL30 8.50.

Having determined that the results from the two devices are not interchangeable, we then evaluated their predictive accuracy of clinical outcomes such as mortality, diagnosis of coagulopathy, and need of blood transfusion. Figures 3–5 display the ROC curves comparing the predictive accuracy of ROTEM EXTEM MCF, ROTEM FIBTEM MCF, and TEG MA for the predetermined clinical outcomes. Table IV lists the C-statistic and the 95% confidence intervals for each and includes a p value comparing EXTEM MCF with MA, and FIBTEM MCF with MA.

All variables, EXTEM MCF (C-statistic: 0.743), FIBTEM MCF (C-statistic: 0.755), and MA (C-statistic: 0.709), have reasonable predictive accuracy for mortality with no statistically significant differences between EXTEM MCF or FIBTEM MCF and MA.

For the diagnosis of coagulopathy, defined in this study by conventional laboratory tests, all variables (C-statistic: <0.7) independently performed poorly for INR ≥1.2. For predicting fibrinogen <1 g/L, MA (C-statistic: 0.743) performed reasonably well, whereas EXTEM MCF (C-statistic: 0.549) and FIBTEM MCF (C-statistic: 0.558) performed poorly. This difference was significant at only the 10% level. There was no significant difference in the predictive value of EXTEM MCF or FIBTEM MCF and MA for either INR or fibrinogen.

For predicting the need for blood transfusions, all measurements (C-statistic: approximately 0.8) had a strong performance accuracy for predicting massive RBC transfusion and reasonable accuracy (C-statistic: almost 0.7) for predicting any RBC transfusion. For predicting frozen plasma transfusion, EXTEM MCF (C-statistic: 0.717) appeared to have better predictive accuracy, defined as reasonable, compared to FIBTEM MCF (C-statistic: 0.574) or MA (C-statistic: 0.620). However, there was no significant difference in the predictive accuracy between the variables.
Finally, neither LI30 nor CL30 had any apparent value at predicting mortality, or RBC transfusion, evidenced by the extremely wide confidence intervals around the odds ratios and C-statistic close to 0.5 shown in Table V. Consequently, although within their clinically acceptable LoA, they are unlikely to have any clinical significance since they both predict outcomes poorly.

**DISCUSSION**

The present study was conducted in a population of severely injured patients identified on arrival to hospital as being at risk of requiring massive blood transfusion and thus being coagulopathic (ETC). This is the trauma patient population most likely to benefit from VHA and possibly in whom the growing number of studies with ROTEM and TEG will focus on. Few studies to date have addressed the question whether their results are similar or interchangeable. Anecdotal data suggests the results of the two tests are widely interpreted as similar.

**TABLE II.** Spearman Correlation Coefficients Between TEG and ROTEM Parameters

<table>
<thead>
<tr>
<th>Variable</th>
<th>Spearman</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT vs. R time</td>
<td>0.19</td>
<td>0.167</td>
</tr>
<tr>
<td>ROTEM Alpha vs. TEG Angle</td>
<td>0.40</td>
<td>0.0005</td>
</tr>
<tr>
<td>CFT vs. K Time</td>
<td>0.47</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>EXTEM MCF vs. MA</td>
<td>0.65</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>LI30 vs. CL30</td>
<td>0.38</td>
<td>0.001</td>
</tr>
</tbody>
</table>

**FIGURE 1.** Scatter plots of ROTEM vs. TEG parameters (CT/R, Alpha (ROTEM)/Angle TEG), CFT/K, EXTEM MCF/MA, LI30/CL30.
The results of our study indicate that when ROTEM and TEG are conventionally done, that the equivalent measurements are not interchangeable (ROTEM CT vs. TEG R; Alpha (ROTEM) vs. Alpha (TEG); CFT vs. K time; MCF vs. MA; LI30 vs. CL30). Despite the strength of the correlation of some parameters and the significant linear association of others, except for lysis indicators, all other parameters fell markedly outside the predefined clinically acceptable LoA. These results are similar to those reported recently by Hagemo et al.17

One possible explanation for the lack of interchangeability may come from the use of different coagulation triggers. Conventionally ROTEM, at least EXTEM and FIBTEM—the portions most used for trauma, is done using tissue factor.

**FIGURE 2.** Bland–Altman difference mean plots of ROTEM vs. TEG parameters (CT/R, Alpha(ROTEM)/Angle (TEG), CFT/K, MCF (EXTEM)/MA, LI30/CL30.
as trigger for the coagulation process. Tissue factor, which interacts with factor VIIa to subsequently activate factor X and prothrombin, activates the extrinsic pathway. In contrast, TEG is conventionally done using kaolin as coagulation trigger that activates contact-dependent factor XII and thus the intrinsic pathway. The fact that the conventional tests use different coagulation triggers (tissue factor, kaolin) activating different pathways may explain the lack of interchangeability, which to our knowledge, has not been mentioned to date. Another possible explanation might be the known limitations of VHA, their wide coefficient of variance. Although most conventional coagulation laboratorial tests have narrow and acceptable coefficients of variance, for VHA the coefficients range from 7.1% to 39.9% for TEG and 7.0% to 83.6% for ROTEM. Commonly the coefficient of variability for both tests is quoted as 30%. This wide variability may affect the calculation of the tests mean values and differences and consequently the evaluation of the closeness between ROTEM and TEG variables.

In addition to analysing the interchangeability of the conventional ROTEM and TEG parameters, we also studied whether one was superior to the other for clinical use. The clinical outcomes studied were the prediction accuracy of mortality, need for blood transfusion, and the diagnosis of coagulopathy.

As reported in a recent systematic review, measurements of clot firmness (ROTEM EXTEM MCF and FIBTEM MCF and their equivalent TEG MA) were good predictors of mortality without any of the parameters proving superior to the others. Our results, however, differ from those of Da Luz et al and the CL measurements (LI30 and CL30) were not significantly associated with mortality. The link between CL and mortality has been described by recent publications, whereas a large randomized control trial CRASH-2 demonstrated that the use of an antifibrinolytic medication reduces mortality in trauma. It is important to note however, that the association between CL and mortality changed according to the timing of the lysis (earlier lysis diagnosed <60 minutes of arrival carried significantly higher mortality rates) and also the amount of lysis. We included in this analysis not only the early CL measurements but all measurements made in the first 12 hours of admission. Furthermore, an analysis of our own experience with over 600 trauma patients (unpublished) was that no patient with TEG maximum lysis up to 99% died, whereas 75% of those with TEG maximum lysis of 100% did. We observed a similar but less clear distinction on the extent of lysis when analysing the CL30 measurement, where only markedly abnormal values were associated to mortality. The wide confidence intervals around the odds ratio for the LI30 and CL30 measurements may also have

| TABLE III. Limits of Agreement for ROTEM and TEG Parameters |
|-----------------|-----------------|-----------------|-----------------|
| Difference      | Difference      | Difference      | Difference      |
| α               | β               | ±α              | ±β              |
| CT vs. R Time   | −7.25           | 1.92            | ±6.97           |
| ROTEM Alpha vs. TEG Angle | −41.79       | 0.56            | ±25.50          |
| CFT vs. K Time  | −3.44           | 1.97            | ±8.35           |
| EXTEM MCF vs. MA| 0               | 3.22            | ±23.33          |
| LI30 vs. CL30   | 0               | 1.52            | ±8.50           |

FIGURE 3. MA, EXTEM MCF, FIBTEM MCF receiver operating characteristic curves for mortality.
accounted for the apparent lack of association of CL measurements and mortality.

Concerning blood transfusions, measurements of clot firmness (ROTEM EXTEM MCF and FIBTEM MCF and TEG MA) and CL (ROTEM LI30 and TEG CL30) were strong indicators of the need for any or massive blood transfusion, particularly for the latter, as well as of plasma transfusion. None of the parameters was clearly superior to others in these determinations. The association between abnormal clot firmness and/or excessive CL with the need for blood
transfusion has been consistently reported by recent studies.4,8,29 These studies also imply that VHA are superior to conventional coagulation laboratory tests in establishing the need for blood transfusions and other hemostatic interventions. If these assumptions are correct, then VHA such as ROTEM and TEG can assist the clinician in determining whether the injured patient needs transfusion, which hemostatic product to use and even the amount. These observations form the basis for the argument on using VHA for trauma resuscitation and led to the development of incipient VHA-based trauma transfusion guidelines.33

All ROTEM and TEG parameters performed similarly poor in diagnosing coagulopathy when defined by INR \(\geq 1.2\). The clot firmness parameters (ROTEM EXTEM MCF and FIBTEM MCF and TEG MA) performed better in diagnosing coagulopathy when defined by a fibrinogen level <1 g/L. None of the measurements performed statistically better than the others. These observations are arguably more a reflection of the poor association between VHA and conventional coagulation tests as previously reported, rather than a deficiency of the assays in diagnosing coagulopathy. The use of conventional coagulation laboratory tests such as INR in trauma has been severely criticized recently due to the lack of association with bleeding and blood transfusion. It has been reported that INR overestimated coagulopathy and should not be used to guide blood transfusion in stable trauma and surgical patients.34 At least one major trauma center has replaced the routine use of conventional coagulation assays on admission of severely injured patients for VHA.35

Overall, we were unable to identify any statistically significant differences between the two VHA in predicting

### TABLE IV. Area Under the Receiver Operating Characteristic (ROC) Curve for ROTEM MCF and TEG MA as Predictors of Mortality, Coagulopathy, and Transfusion.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Variable</th>
<th>Area Under the ROC Curve (95% CI)</th>
<th>p Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>MA</td>
<td>0.709 (0.563–0.855)</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>EXTEM MCF</td>
<td>0.743 (0.607–0.880)</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>FIBTEM MCF</td>
<td>0.755 (0.563–0.947)</td>
<td></td>
</tr>
<tr>
<td>Diagnosis of Coagulopathy</td>
<td>MA</td>
<td>0.743 (0.530–0.956)</td>
<td></td>
</tr>
<tr>
<td>Fibrinogen &lt;1 g/L</td>
<td>EXTEM MCF</td>
<td>0.549 (0.285–0.812)</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>FIBTEM MCF</td>
<td>0.558 (0.348–0.769)</td>
<td>0.12</td>
</tr>
<tr>
<td>INR (\geq 1.2)</td>
<td>MA</td>
<td>0.595 (0.452–0.738)</td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td>EXTEM MCF</td>
<td>0.566 (0.422–0.709)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FIBTEM MCF</td>
<td>0.595 (0.452–0.738)</td>
<td>0.76</td>
</tr>
<tr>
<td>Blood Transfusion</td>
<td>MA</td>
<td>0.668 (0.502–0.834)</td>
<td>0.84</td>
</tr>
<tr>
<td>RBC Any</td>
<td>EXTEM MCF</td>
<td>0.686 (0.545–0.827)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FIBTEM MCF</td>
<td>0.635 (0.493–0.777)</td>
<td>0.72</td>
</tr>
<tr>
<td>RBC Massive ((\geq 10) Units)</td>
<td>MA</td>
<td>0.812 (0.706–0.918)</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>EXTEM MCF</td>
<td>0.830 (0.734–0.927)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FIBTEM MCF</td>
<td>0.783 (0.646–0.919)</td>
<td>0.0001</td>
</tr>
<tr>
<td>Fresh Frozen Plasma</td>
<td>MA</td>
<td>0.620 (0.475–0.765)</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>EXTEM MCF</td>
<td>0.717 (0.603–0.832)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FIBTEM MCF</td>
<td>0.574 (0.431–0.716)</td>
<td>0.84</td>
</tr>
</tbody>
</table>

CI, confidence interval; INR, international normalized ratio; RBC, red blood cell. *Compared to MA.

### TABLE V. Predictive Value of LI30 vs. CL30

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Variable</th>
<th>Odds Ratio (95% CI)</th>
<th>C-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mortality</td>
<td>LI30</td>
<td>1.425 (0.147–13.807)</td>
<td>0.522</td>
</tr>
<tr>
<td></td>
<td>CL30</td>
<td>0.6 (0.103–3.506)</td>
<td>0.561</td>
</tr>
<tr>
<td>Diagnosis of Coagulopathy</td>
<td>LI30</td>
<td>1.056 (0.092–12.137)</td>
<td>0.503</td>
</tr>
<tr>
<td>Fibrinogen (&lt;1 g/L)</td>
<td>CL30</td>
<td>0.25 (0.025–2.489)</td>
<td>0.629</td>
</tr>
<tr>
<td>INR (\geq 1.2)</td>
<td>LI30</td>
<td>0.719 (0.163–3.176)</td>
<td>0.519</td>
</tr>
<tr>
<td></td>
<td>CL30</td>
<td>0.848 (0.311–2.317)</td>
<td>0.52</td>
</tr>
<tr>
<td>Blood Transfusion</td>
<td>LI30</td>
<td>0.673 (0.122–3.723)</td>
<td>0.65</td>
</tr>
<tr>
<td>RBC Any</td>
<td>CL30</td>
<td>0.51 (0.145–1.794)</td>
<td>0.58</td>
</tr>
<tr>
<td>RBC Massive ((\geq 10) Units)</td>
<td>LI30</td>
<td>1.19 (0.219–6.472)</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>CL30</td>
<td>0.556 (0.169–1.831)</td>
<td>0.57</td>
</tr>
</tbody>
</table>

CI, confidence interval; INR, international normalized ratio; RBC, red blood cell.
mortality, blood transfusion, and diagnosing ETC. Thus, there is no indication that sites currently using either ROTEM or TEG should consider changing their device. We did not investigate whether one assay adds significant information to the other, but this possibility seems unlikely. It may be relevant to emphasize that while the two laboratory assays investigated may have an impact on patient’s outcome, sound clinical principles such as early hemorrhage control may have a larger impact on outcome and should continue to be considered higher priority.9

This present study has limitations. Some of them may be inherent to the assays including their wide coefficient of variance discussed above, and the known variations due to age, gender, and race that were not considered in this analysis.36,37 Another inherent limitation is that VHA results are displayed both numerically and as a graph. At times, the VHA curve shapes can provide invaluable and timely information that cannot be measured or analyzed in a study such as this. The cohort analyzed is comprised arguably by the patient most likely to benefit from the introduction of VHA to trauma resuscitation but the sample size is small. The strengths are on the cohort of patients enrolled and the fact that both tests were done at the same time, by trained personnel and according to the highest standards of quality. Another limitation of this study is relatively small sample size especially for mortality. Larger sample size may allow us to see more predictability of VHA and any differences in their clinical performance.

Despite their limitations, VHA could have a significant role in the early resuscitation of bleeding trauma patients, enabling more cost-effective treatment, guiding blood transfusion, and improving clinical outcome.58

CONCLUSIONS
The results from TEG and ROTEM, when conventionally performed, failed to reach acceptable limits of agreement and thus are not interchangeable. Any guidelines developed for one instrument should not be extrapolated for the other. The difference may result from the use of different activators of the coagulation, which trigger different pathways. Consequently, guidelines developed for one instrument should not be extrapolated for the other. Although the results are not interchangeable, both VHA appear to have a similar clinical performance in predicting mortality, the need for blood transfusion, and diagnosing early trauma coagulopathy.

ACKNOWLEDGMENT
This study was supported by Defence Research and Development Canada through a contract W7719-125093.

REFERENCES
Bayesian Scoring Systems for Military Pelvic and Perineal Blast Injuries: Is it Time to Take a New Approach?

Somayyeh Mossadegh, BM, MRCS*; Shan He, BEng, MSc, PhD†; Paul Parker, FIMC, FRCS (Orth)‡

ABSTRACT  Background: Various injury severity scores exist for trauma; it is known that they do not correlate accurately to military injuries. A promising anatomical scoring system for blast pelvic and perineal injury led to the development of an improved scoring system using machine-learning techniques. Methods: An unbiased genetic algorithm selected optimal anatomical and physiological parameters from 118 military cases. A Naïve Bayesian model was built using the proposed parameters to predict the probability of survival. Ten-fold cross validation was employed to evaluate its performance. Results: Our model significantly out-performed Injury Severity Score (ISS), Trauma ISS, New ISS, and the Revised Trauma Score in virtually all areas; positive predictive value 0.8941, specificity 0.9027, accuracy 0.9056, and area under curve 0.9059. A two-sample t test showed that the predictive performance of the proposed scoring system was significantly better than the other systems (p < 0.001). Conclusion: With limited resources and the simplest of Bayesian methodologies, we have demonstrated that the Naïve Bayesian model performed significantly better in virtually all areas assessed by current scoring systems used for trauma. This is encouraging and highlights that more can be done to improve trauma systems not only for our military injured, but also for civilian trauma victims.

INTRODUCTION

Trauma scoring systems are used by the United Kingdom (U.K.) Defence Medical Services, in order to provide accurate risk assessments for injured service personnel. This is seen as essential research and investment as part of the quality assurance process, in order to maintain and improve clinical practice.1 As these scoring systems are well known internationally, it allows comparison of clinical performances between institutions and countries.

An in-depth analysis of pelvic and perineal blast injuries in U.K. military service personnel over an 8-year period of the war in Afghanistan provided the evidence for our study injury pattern and created a didactic module for the Military Operational Surgical Training course at the Royal College of Surgeon of England.2 Review of actual injuries sustained by this cohort of patients identified that Injury Severity Score (ISS) did not correlate with the severity of pelvic and perineal injuries,3 a finding that is well known amongst the military medical profession and one that carries a high mortality rate risk assessments for injured service personnel. This is necessary to develop this scoring system into a valid model worthy of use in current trauma systems. This article outlines the preliminary steps in developing a mortality prediction algorithm using Naïve Bayesian (NB) analysis specifically in a population of combat trauma patients exposed to an improvised explosive devices (IED) blast injury. We hypothesize that using a nonlogistic regression based algorithm, we were able to produce a more accurate scoring system with highly significant implications for future development on a global scale.

METHODS

The cohort of casualties in this article was initially analyzed in two previous publications.3,8 The U.K. Joint Theatre Trauma Registry (JTTR) was used to identify all U.K. military service personnel who had sustained a perineal injury over an 8-year period (January 2003–December 2010). Coalition military were excluded due to nonavailability of follow-up. To ensure complete data capture, anal and genitourinary injuries were also specified in search terms.

Anatomical injuries affecting the pelvis and perineal region were firstly identified using abbreviated injury scale (AIS) codes and then corroborated using free-text descriptions to confirm that they matched the AIS Military (2005) descriptions.24 Although this was time consuming, it was necessary as frequent input errors were identified and corrected. Physiological data were also available relating to observations at various times, from point of wounding (Role 1
95 cases had TRISS and RTS values. The standard coefficients used for both the U.K. and U.S. JTTR are still based on the 1980’s blunt and penetrating injuries. The same standard weighting factors were also used for the RTS. Ultimately, for blast injuries, the most predominant injury would decide if a blunt or penetrating coefficient would be used. If intubated, a casualty would have been given a Glasgow Coma Scale of 3, respiratory rate was recorded as 0 if being ventilated and the rate achieved if breathing spontaneously through the tube.

All anatomical injuries in the pelvis and perineal region (Table II) as well as the physiological parameters were put through a process of optimization using an unbiased genetic algorithm to select the most important variables using the Waikato Environment for Knowledge Analysis. This is a relatively easy data-mining tool used by many trauma registries. A genetic algorithm is a search heuristic that mimics the process of natural selection. The algorithm first generates a population of random solutions (e.g., a number of physiological measures as features). Then the algorithm will execute selection, e.g., selecting those good solutions (sets of features) that can generate better classification results. Genetic operators, e.g., mutation (changing the features randomly) and crossover (combining two sets of features) will then be applied to create a new generation of solutions. The algorithm will then perform selection and the iterations continue until some stopping criteria are met. This process is represented diagrammatically in Figure 2. The algorithm selected 3 variables from the physiological measures: “BP systolic R3,” “BP diastolic R3,” and “pulse R3.” We also included all pelvic and perineal measurements: “pelvis,” “penis,” “testes,” “scrotum,” and “anorectum.” The pelvic measurements referred to severity of pelvic fracture sustained. This scale was derived from the Young and Burgess classification.

These variables were then used to build a Naïve Bayes Classifier model. All the algorithms were programmed using MATLAB (Mathworks, Natick, Massachusetts). The average performance of our system was evaluated using 30 runs of 10-fold cross validation and this compared against the performance of ISS, NISS, TRISS, and RTS.

TABLE I. All Variables Considered, Variables Selected, and Their Coefficients

<table>
<thead>
<tr>
<th>All Variables</th>
<th>Variables Selected</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anatomical</td>
<td>Pelvis</td>
<td>-132.781</td>
</tr>
<tr>
<td></td>
<td>Penis</td>
<td>-93.22</td>
</tr>
<tr>
<td></td>
<td>Scrotum</td>
<td>-63.540</td>
</tr>
<tr>
<td></td>
<td>Anus</td>
<td>-54.486</td>
</tr>
<tr>
<td></td>
<td>Testes</td>
<td>53.492</td>
</tr>
<tr>
<td></td>
<td>Ructum</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Urethra</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Perineum</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hip/Buttocks</td>
<td></td>
</tr>
<tr>
<td>Physiological</td>
<td>BP Systolic,</td>
<td>14.872</td>
</tr>
<tr>
<td></td>
<td>R3, R2, R1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BP Diastolic,</td>
<td>-22.233</td>
</tr>
<tr>
<td></td>
<td>R3, R2, R1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pulse R3, R2, R1</td>
<td>7.581</td>
</tr>
<tr>
<td></td>
<td>Respiratory Rate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>R3, R2, R1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SpO2 R3, R2, R1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Temperature R3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>WCC</td>
<td></td>
</tr>
</tbody>
</table>

BP, blood pressure; R3, Role 3; R2, Role 2; R1, Role 1; SpO2, oxygen saturation; WCC, white cell count; ISS – Injury Severity Score; NISS, New ISS; TRISS, Trauma ISS; RTS – Revised Trauma Score.

RESULTS

The JTTR contained 4,808 coalition casualties, of which 2,204 were U.K. Military and 118 (5.4% of 2,204) were identified as having sustained a perineal injury. The mean age of the cohort was 25, case fatality rate was 47%, and there were 62 (53%) survivors. The demographics of this cohort of patients as well as the distribution of ISS are summarized in Table III. The overall case fatality rate for all U.K. military IED related casualties was 28% (all mechanisms was 25%). Using the definition of an unexpected survivor having a TRISS of less than 50% in a surviving casualty, there were 9 unexpected survivors (17%). Twenty three patients did not have TRISS assigned because of index parameters not being recorded, 13 from this group survived. Genitourinary injuries were identified in 85 (72%) of 118 patients. There was a significantly lower rate of mortality in patients with perineal injuries alone compared to
Severity score based on abbreviated injury scale (AIS); APC, anterior posterior compression; CMI, combined mechanism injury; LC, lateral compression; SJJ, sacroiliac joint; VS, vertical shear.

FIGURE 2. Diagrammatic representation of the whole data analysis process.

TABLE II. Anatomical Components of Cumulative Pelvic and Perineal Trauma Scoring System

<table>
<thead>
<tr>
<th>Structure Components</th>
<th>Pelvic Fracture Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penis</td>
<td>0</td>
</tr>
<tr>
<td>Scrotum</td>
<td>1</td>
</tr>
<tr>
<td>Testes</td>
<td>2</td>
</tr>
<tr>
<td>Urethra</td>
<td>3</td>
</tr>
<tr>
<td>Perineum</td>
<td>4</td>
</tr>
<tr>
<td>Anus</td>
<td>5</td>
</tr>
<tr>
<td>Rectum</td>
<td>6</td>
</tr>
</tbody>
</table>

Over time there have been major improvements and modifications were made for military trauma over the past 40 years. These are based on anatomical or physiological descriptors or some combination of the two.1,12-17 Over time there have been major improvements and modifications to these scoring systems. However, the simple fact remains that these scoring systems using anatomical descriptors are based on the AIS, originally founded in the early 1970s, by the U.S. automotive industry.18 This scale was developed to measure the severity of (frequently blunt) vehicular trauma at that time. It became useful in measuring other traumatic injuries from different mechanisms of injury, and since its inception there have been several revisions. The most recent revision is AIS 2005 update 2008.19-23 Some modifications were made for military trauma (AIS 2005-Military) using expert military surgeons to account for multiple injury etiology from high-energy weapons and explosives, but these did not envisage any triple amputee pelvi-perineally injured IED survivors.24,25

In light of this fact, a Military Injury Scoring Summit was convened in 2008 at the U.S. Army Institute of Surgical Research in San Antonio, Texas. It comprised of a panel of military and civilian experts and resulted in the Military Combat Injury Scale and the Military Functional Incapacity Scale, which was introduced into the literature in 2013.26 This may be reproducible and relatively simple to use, however, with further advances in technology and cross-disciplinary collaborations between computer sciences and medicine, it may already be outdated. The algorithm for our NB model was created using MATLAB (Mathworks)—this clearly is not a program used by clinicians on a daily basis; however, this is where cross-disciplinary collaboration is paramount. Yet et al27 have demonstrated that the development of a Bayesian Network (BN) utilizes not only the raw data available, but also integrates domain expertise to develop and refine a prediction model. Once a model has been created, the next step is to create an interface that is user friendly and readily available. Naturally, there are many steps before a usable interface is created; however, it is possible and as an example, a coagulopathy prediction tool is available online usable interface is created; however, it is possible and as an example, a coagulopathy prediction tool is available online with a fully working BN.27-29

A NB analysis is a form of machine learning, where NB classifiers represent a family of simple probabilistic classifiers based on applying Bayes’ theorem using strong (naïve) independence assumptions between the features.30,31 A Naïve
Bayes classifier assigns a new observation to the most probable class, assuming the features are conditionally independent given the class value. Bayesian updating is especially important in analyzing sequences of data inference, and has found application in a range of fields including science, engineering, philosophy, medicine, and law.30,31

Some might argue that maybe all that is needed is a new blast TRISS coefficient; however, we would argue that this is frequently based on retrospective data and requires the information to be in a linear distribution. Bayesian algorithms are nonlinear and can use prospective data to learn and develop with more information, including subject matter experts. This concept is currently used to create consensus statements; however, it has not been incorporated into trauma scoring models.

Regarding the seemingly contradictory positive effect of testicular loss, recent studies have shown that there is indeed a sex-related difference in outcome in trauma patients with males being more susceptible to multiple organ failure, sepsis, and mortality after trauma.32 Female sex was associated with improved organ function following traumatic injury and hemorrhagic shock, and in particular, in the reproductive age groups (16–44 years).33,34 Testicular injury was the only anatomical area with a positive coefficient in our study, thus corroborating the notion that there may be a protective mechanism as the degree of testicular injury worsened. Further evidence is required before the use of sex steroids or testosterone blockade as a therapeutic intervention in critically ill trauma patients can be advocated.

In statistics, logistic regression is a method of analyzing multiple independent variables in a dataset in order to determine the outcome, which is a binary measurement. It aims to identify the best fitting, yet biologically reasonable, model to describe the relationship between the two outcomes and the set of independent variables. Coefficients are generated to predict a logit (log odds) transformation of the probability of presence of the characteristic of interest.32 Unfortunately, it does not account for missing data. Bayesian methodology does, and this analysis was used for our study.

Finally, we believe that this data is encouraging and highlights an exciting prospect for the future of improving trauma systems, especially as the NB is the simplest of all the Bayesian models and is purely data driven. As more information is given to the model and developments change, the model will learn from the accumulating data and will allow a more meaningful comparison between different trauma systems, theatres, and time periods. Future work will be directed toward using nonlinear trauma models such as BN to aid clinical governance as well as develop mortality prediction tools looking at total injury burden, which may lead to creating a more accurate global trauma scoring system.

TABLE III. Demographics and Distribution of Injury Severity Scores

<table>
<thead>
<tr>
<th></th>
<th>ALL</th>
<th>Perineal Injury Alone</th>
<th>Perineal Injury and Pelvic Fracture</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Casualties, n</td>
<td>118</td>
<td>56 (47)</td>
<td>62 (53)</td>
<td>NS</td>
</tr>
<tr>
<td>Mortality, n (%)</td>
<td>56 (48)</td>
<td>11 (20)</td>
<td>45 (73)</td>
<td>0.0001</td>
</tr>
<tr>
<td>KIA, n (%)</td>
<td>35 (67)</td>
<td>8 (23)</td>
<td>27 (77)</td>
<td>0.0001</td>
</tr>
<tr>
<td>DOW, n (%)</td>
<td>21 (37)</td>
<td>3 (14)</td>
<td>18 (86)</td>
<td>0.0001</td>
</tr>
<tr>
<td>CFR, %</td>
<td>47</td>
<td>20</td>
<td>73</td>
<td>0.0001</td>
</tr>
<tr>
<td>Survivors, n</td>
<td>62</td>
<td>45</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Exclude KIA, n (%)</td>
<td>83</td>
<td>48 (58)</td>
<td>35 (42)</td>
<td>NS</td>
</tr>
</tbody>
</table>

**Mean** | **Median (IQR)**

| Age         | 25 (21–29) |
| ISS (n = 118)| 41.03 (29–57) |
| NISS (n = 118) | 53.14 (36–75) |
| TRISS (n = 95) | 43.51 (0.36–96.69) |
| RTS (n = 95)      | 7.95 (0.11–92.88) |

CFR, case fatality rate; DOW, died of wounds; KIA, killed in action; NS, not significant; ISS, Injury Severity Score; NISS, New ISS; TRISS, Trauma ISS; p, significance.

TABLE IV. Naïve Bayesian Analysis of Novel Score vs. Current Trauma Scoring Systems

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Precision</th>
<th>Recall</th>
<th>Specificity</th>
<th>Accuracy</th>
<th>AUC</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISS</td>
<td>0.858</td>
<td>0.800</td>
<td>0.879</td>
<td>0.842</td>
<td>0.844</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NISS</td>
<td>0.891</td>
<td>0.857</td>
<td>0.905</td>
<td>0.882</td>
<td>0.880</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>TRISS</td>
<td>0.784</td>
<td>0.954</td>
<td>0.761</td>
<td>0.853</td>
<td>0.859</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>RTS</td>
<td>0.788</td>
<td>0.929</td>
<td>0.774</td>
<td>0.848</td>
<td>0.851</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Novel Scoring System</td>
<td>0.894</td>
<td>0.909</td>
<td>0.903</td>
<td>0.906</td>
<td>0.906</td>
<td></td>
</tr>
</tbody>
</table>

ISS, Injury Severity Score; NISS, New ISS; , TRISS, Trauma ISS, RTS, Revised Trauma Score; AUC, area under curve (calculated by 30 runs of 10-fold cross validation); p, significance.
SUMMARY
This article describes a novel concept for developing a scoring system that correlates better to injuries sustained by U.K. military personnel from IEDs. It has not taken into account any other injuries at present and is no doubt a weak study, which we fully acknowledge, yet it still outperforms current scoring systems. Its progression from a simple cumulative scoring system to a NB mathematical model, independent of any formal funding or working groups, is testament to the authors’ determination that this is the future of trauma scoring systems. When used for the purpose that they were designed for—moderate speed automotive crashes, the AIS is an excellent tool. However, newer concepts such as en route airway care, prehospital hemostatic blood transfusion, damage control resuscitation and surgery, and surgeon and institutional trauma volume now means that an ISS of 75 does not always mean death. As treatments evolve—so must scoring systems.

REFERENCES
Preflight Variables Are Associated With Increased Ventilator Days and 30-Day Mortality in Trauma Casualties Evacuated by Critical Care Air Transport Teams: An Exploratory Retrospective Study

Surgeon Lieutenant Commander Ed Barnard, Royal Navy*†; Alejandra G. Mora, BS*; Lt Col Vikhyat S. Bebarta, USAF MC*

ABSTRACT  Background: There are no tools to predict outcomes in the U.S. Air Force Critical Care Air Transport Team (CCATT) trauma patients. The objective of this study was to identify associations between preflight variables and outcomes that could assist planning of ongoing critical care. Methods: This Institutional Review Board approved retrospective study included all patients evacuated from Afghanistan by CCATT between 2007 and 2011. Preflight variables were assessed for associations and examined in logistic regression models. Ventilator time over 72 hours, and 30-day mortality were the primary and secondary outcomes respectively. Results: 1,308 trauma patients (24 years, 98% male) were included; 72% blast. Injury severity score (odds ratio [OR] = 1.04 [1.03–1.06]), preflight packed red blood cell units transfused (OR = 1.05 [1.04–1.07]), and preflight intubated status (OR = 11.9 [8.53–16.89]) were independently associated with increased ventilator days; a composite produced an area under the curve of 0.85 with 86% sensitivity and 56% specificity. Injury severity score (OR = 1.06 [1.03–1.09]), prothrombin time (OR = 2.13 [1.18–4.47]), preflight intubated status (OR = 9.2 [1.88–166.11]), and whole blood (OR = 3.18 [1.38–7.04]) were associated with 30-day mortality; a composite produced an area under the curve of 0.84 with 71% sensitivity and 57% specificity. Conclusion: In our large CCATT study a number of preflight variables were associated with outcomes, which may assist in the future planning of critical care services.

INTRODUCTION

There have been considerable improvements in outcomes from combat trauma in the past 10 years.1,2 This improvement is multifactorial,3 but is in part due to the development of robust aeromedical evacuation (AE) systems.4–7 United States and United Kingdom combat trauma patients injured during Operation Enduring Freedom (Afghanistan, 2001 to 2014) have typically been retrieved to a medical treatment facility by an AE helicopter.8 There is limited published evidence of effectiveness in this military AE system during Operation Enduring Freedom. Two published studies have compared the level of military AE provider and outcomes: Both studies demonstrated that a higher level of provider (Emergency Medical Technician-BASIC [EMT-Bs] compared to Critical Care Flight Paramedics,9 and Pararescuemen and EMT-Bs compared to U.K. prehospital physicians)10 is associated with reduced mortality.

After initial resuscitation and surgery, patients who require critical care support are transferred to hospitals in Europe by aircraft equipped to provide this high level of care: The U.S. Critical Care Air Transport Team (CCATT), and the U.K. Critical Care Air Support Team.11–13 Physiological scoring systems that predict critical care outcomes in the civilian setting are abundant.14–16 However, there are no tools that specifically examine combat intensive care unit (ICU) patients.

For military logistic planning purposes, it would be useful to determine if preflight physiology and necessary treatments were associated with relevant patient outcomes postflight. Specifically, in times of increased demand (e.g., a major incident) or when there is a reduced capability (for example during expeditionary warfare) it would be useful to have a tool that could assist logistical planning of the expected level of medical care that these critically injured patients were likely to require, and their likely survival rate. The U.S. system generally assigns a level of ICU capability per hospital, rather than to individual patients’ needs.17 However, the U.K. system assigns a “level of care” to individual patients—allowing identification of patients who require the greatest amount of medical resource. The highest level of ICU care in the U.K. system is “Level 3”—defined as the requirement for invasive mechanical ventilator support (intubated, or with a tracheostomy), and/or the need for 2 or more organ support.18 Survival is a binary, objective, and relevant outcome, and in trauma research is typically reported as 30-day mortality.19 ICU and hospital length of stay outcomes were not used in our study owing to their questionable validity (subjectivity, and bias from nonclinical factors),20 and have not been reported in recent large trials of traumatic injury.21,22

We, therefore, designed an exploratory retrospective study, reanalyzing data captured in a preapproved prospective database of CCATT patients, to evaluate the association between preflight variables and increased ventilator time.
METHODS

This exploratory retrospective study is a reanalysis of a prospectively recorded database of U.S. CCATT patients who were evacuated from Afghanistan between 2007 and 2011 (CCATT database). The primary aim of the “CCATT database” was to enable the creation of a comprehensive record of CCATT activity, from which it would be possible to retrospectively address research questions. The study protocol was approved by the Wilford Hall Ambulatory Surgical Center Institutional Review Board.

All trauma patients transported by CCATT in this time period were included. Data were inputted into a spreadsheet (Microsoft, Excel 2010, Redmond, Washington) and validated for errors. All data, with the exception of injury severity scores (ISS), total ventilator time, and 30-day mortality were captured directly from the CCATT medical record. ISS, total ventilator time, and 30-day mortality were obtained by cross-referencing patient records with the Department of Defense Trauma Registry (DoDTR). DoDTR (formerly the

**TABLE I.** Overall Patient Characteristics: Injury Description, Preflight Physiology, and Administration of Blood Products (n = 1,308)

<table>
<thead>
<tr>
<th>Injury Description</th>
<th>Median [IQR] or % (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISS</td>
<td>22 [16–29]</td>
</tr>
<tr>
<td>Blast</td>
<td>72% (939)</td>
</tr>
<tr>
<td>Penetrating</td>
<td>18% (229)</td>
</tr>
<tr>
<td>Blunt</td>
<td>9% (118)</td>
</tr>
<tr>
<td>Burn</td>
<td>2% (22)</td>
</tr>
<tr>
<td><strong>Physiology</strong></td>
<td></td>
</tr>
<tr>
<td>Heart Rate (bpm)</td>
<td>100 [84–116]</td>
</tr>
<tr>
<td>SBP (mm Hg)</td>
<td>121 [109–143]</td>
</tr>
<tr>
<td>Base Deficit</td>
<td>0 [–2–2]</td>
</tr>
<tr>
<td><strong>Blood Products</strong></td>
<td></td>
</tr>
<tr>
<td>PRBC (Unit)</td>
<td>4 [0–13]</td>
</tr>
<tr>
<td>FFP (Unit)</td>
<td>3 [0–12]</td>
</tr>
<tr>
<td>Whole Blood</td>
<td>8% (100)</td>
</tr>
</tbody>
</table>

IQR, interquartile range; ISS, injury severity score; bpm, beats per minute; PRBC, packed red blood cells; FFP, fresh frozen plasma.

**TABLE II.** Evaluation of Demographics, Injury, and Preflight Physiological and Medical Interventions Against Ventilator Time ≤72 Hours and Over 72 Hours (n = 1,308). Data are presented as a percentage (%) or median [interquartile range].

<table>
<thead>
<tr>
<th>Variable</th>
<th>≤72 Hours Ventilation (n = 735)</th>
<th>&gt; 72 Hours Ventilation (n = 573)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender (% Male)</td>
<td>98%</td>
<td>99%</td>
<td>NS</td>
</tr>
<tr>
<td>Age (Years [IQR])</td>
<td>24 [21–29]</td>
<td>24 [21–30]</td>
<td>NS</td>
</tr>
<tr>
<td>Injury Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blast</td>
<td>69%</td>
<td>76%</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Penetrating</td>
<td>19%</td>
<td>16%</td>
<td>NS</td>
</tr>
<tr>
<td>Blunt</td>
<td>11%</td>
<td>7%</td>
<td>NS</td>
</tr>
<tr>
<td>Burn (Burn Only)</td>
<td>4% (2%)</td>
<td>7% (2%)</td>
<td>NS (NS)</td>
</tr>
<tr>
<td>Preflight Parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intubated</td>
<td>38%</td>
<td>90%</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Systolic Blood Pressure (mm Hg)</td>
<td>123 [112–136]</td>
<td>118 [106–131]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Diastolic Blood Pressure (mm Hg)</td>
<td>63 [56–71]</td>
<td>60 [54–67]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Heart Rate (bpm)</td>
<td>96 [81–111]</td>
<td>105 [89–121]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Shock Index</td>
<td>0.8 [0.6–0.9]</td>
<td>0.9 [0.7–1.1]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>INR</td>
<td>1.1 [0.9–1.3]</td>
<td>1.2 [1–1.3]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Platelet Count (×10⁹/L)</td>
<td>126 [99–166]</td>
<td>107 [88–135]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Base Deficit</td>
<td>0 [–2–2]</td>
<td>0 [–2–2]</td>
<td>NS</td>
</tr>
<tr>
<td>Preflight Resuscitation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crystalloid (% Yes)</td>
<td>82%</td>
<td>94%</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Colloid (% Yes)</td>
<td>22%</td>
<td>40%</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Blood Products (% Yes)</td>
<td>56%</td>
<td>84%</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Plasma Units, (% Yes)</td>
<td>7 [3–13], 46%</td>
<td>12 [4–24], 79%</td>
<td>&lt;0.0001, &lt;0.0001</td>
</tr>
<tr>
<td>RBC Units, (% Yes)</td>
<td>7 [3–13], 53%</td>
<td>13 [6–25], 77%</td>
<td>&lt;0.0001, &lt;0.0001</td>
</tr>
<tr>
<td>Platelets, (% Yes)</td>
<td>2 [1–4], 29%</td>
<td>4 [2–6], 60%</td>
<td>&lt;0.0001, &lt;0.0001</td>
</tr>
<tr>
<td>Cryoprecipitate Units, (% Yes)</td>
<td>3 [1–10], 17%</td>
<td>5 [2–11], 43%</td>
<td>&lt;0.01, &lt;0.0001</td>
</tr>
<tr>
<td>Whole Blood Units, (% Yes)</td>
<td>3 [2–8], 3%</td>
<td>5 [4–10], 14%</td>
<td>&lt;0.05, &lt;0.0001</td>
</tr>
<tr>
<td>Factor VIIa (% Yes)</td>
<td>3%</td>
<td>18%</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

GCS, Glasgow Coma Scale score; IQR, interquartile range; ISS, injury severity score; bpm, beats per minute; NS, nonsignificant; INR, international normalized ratio; RBC, red blood cells.
Joint Theater Trauma Registry) is a performance improvement military medical database of casualties treated in Iraq, Afghanistan, and other deployment areas.23

**Primary Analysis**

Twenty-seven preflight variables were examined in this study: including patient demographics (gender, age, mechanism of injury, and ISS), patient physiology (shock index, blood pressure, heart rate, Glasgow coma scale score, pH, base deficit, platelet count, hematocrit, and prothrombin time/international normalized ratio [INR]), and medical interventions (intubated status, and administration of at least one unit of: any blood product, fresh whole blood, fresh frozen plasma, packed red blood cells (PRBCs), platelets, cryoprecipitate, and recombinant activated factor VIIa).

**Primary Outcome**

The primary outcome was increased ventilator time. The median duration of ventilation required for this cohort of CCATT patients was 72 hours; therefore, ventilator time over 72 hours was used as the definition of “increased ventilator time.”

**Secondary Outcome**

The secondary outcome was 30-day mortality.

**Secondary Analysis**

The primary aim of this study was to examine which preflight variables affected two relevant outcomes in CCATT patients. However, to give an idea of the magnitude of the effect of different variables a secondary analysis was performed to quantify the effect of individual variables on the primary and secondary outcomes.

**Data Analysis**

Population descriptive data are presented as medians with interquartile ranges (IQR). Statistical analyses were conducted with JMP version 10 (SAS Institute Inc., Cary, North Carolina). The 27 preflight variables were independently examined for associations with the primary and secondary outcomes. After univariate analyses, progressive stepwise logistic regression analyses were performed to establish the significant parameters for each model. We began with saturated models to include all variables and progressively eliminated variables when significance was equal to or greater than 0.2. Akaike’s information criterion as a measure of best fit was then used to identify the optimal model with most appropriate parameters to predict the primary and secondary outcomes. Finally, these variables were evaluated as a composite in receiver operating characteristic curves to ascertain their combined probability of ruling in or ruling out the primary and secondary outcomes. The cut point of the curve was the point closest to the top-left corner, and the area under the curve (AUC) calculated to demonstrate the accuracy of the composite variable model in predicting the predefined outcomes.

**RESULTS**

In our study, 1,308 combat trauma patients (median age of 24 years, 98% male, ISS 22 [16–29]) were included (Table I). Sixty-one (4.7%) patients did not have 30-day mortality data available, and therefore were not included in the secondary outcome analysis.

**Primary Analysis**

**Primary Outcome**

When modeling for increased ventilator time, the significance of 27 preflight variables was independently assessed for their association with the primary outcome (Table II). Twenty variables independently demonstrated significance between groups. While modeling for association with prolonged ventilator time, variables with the least significance were sequentially removed to identify those with the greatest effect. ISS (OR = 1.04 [1.03–1.06]), PRBC units transfused (OR = 1.05 [1.04–1.07]), and intubated status (OR = 11.9 [8.53–16.89]) were independently associated with requiring ventilation for greater than 72 hours. A composite of the variables produced an AUC of 0.85 with 86% sensitivity and 56% specificity (Fig. 1).

**Secondary Outcome**

When modeling for increased ventilator time, the significance of 27 preflight variables was independently assessed for their association with the primary outcome (Table II). Twenty variables independently demonstrated significance between groups. While modeling for association with prolonged ventilator time, variables with the least significance were sequentially removed to identify those with the greatest effect. ISS (OR = 1.04 [1.03–1.06]), PRBC units transfused (OR = 1.05 [1.04–1.07]), and intubated status (OR = 11.9 [8.53–16.89]) were independently associated with requiring ventilation for greater than 72 hours. A composite of the variables produced an AUC of 0.85 with 86% sensitivity and 56% specificity (Fig. 1).

**Secondary Outcome**

When modeling for 30-day mortality, the significance of 27 preflight variables was independently assessed for their association with the secondary outcome (Table III).

![FIGURE 1. Receiver operating characteristic curve for the primary outcome (ventilator time over 72 hours). The composite of injury severity score, packed red cells transfusion, and intubated status produced an area under the curve of 0.85, and 86% sensitivity and 56% specificity.](image-url)
Seventeen variables independently demonstrated significance between groups. While modeling for association with 30-day mortality, variables with the least significance were sequentially removed to identify those with the greatest effect. Using 30-day mortality as the outcome, ISS (OR = 1.06 [1.03–1.09]), INR (OR = 2.13 [1.18–4.47]), intubated (OR = 9.2 [1.88–166.11]), and whole blood (OR = 3.18 [1.38–7.04]) were associated with death. The combination of variables produced an AUC of 0.84 with 71% sensitivity and 57% specificity (Fig. 2).

### Secondary Analysis

#### Primary Outcome

For every one unit increase in ISS or PRBCs there was a corresponding 4% and 5% increase in the odds of requiring >72 hours of invasive ventilation respectively. Patients who were intubated preflight had a 12-fold increase in odds of requiring >72 hours of invasive ventilation.

#### Secondary Outcome

For every one unit increase in ISS there was a corresponding 6% increase in the odds of 30-day mortality. Patients who were intubated preflight had a nine-fold increase in the odds of 30-day mortality, and for every 2 second increase in prothrombin time there was a 13% increase in the odds of 30-day mortality. Patients who received a whole blood transfusion had a three-fold increase in the odds of 30-day mortality.

### DISCUSSION

The primary aim of this study was to determine if preflight variables could be mathematically combined to predict military logistic planning relevant outcomes in CCATT patients. The benefit of that could be realized at times of increased demand (e.g., a major incident), and or at times of reduced capability (e.g., during future expeditionary warfare).

Our study demonstrates the potential to rule out the requirement for ventilation over 72 hours, and therefore the need for the highest level of ICU care, by using the ISS, the preflight transfusion of PRBCs, and if the patient was intubated preflight or not, with a sensitivity of 86%. This exploratory study has also shown that 30-day mortality can be ruled out (and therefore short-term survival ruled in), with a 71% chance by examining the ISS, the administration

---

**TABLE III.** Evaluation of Demographics, Injury, and Preflight Physiological and Medical Interventions Against 30-Day Mortality (n = 1,247). Data are presented as a percentage (%) or median [inter-quartile range].

<table>
<thead>
<tr>
<th>Variable</th>
<th>30-Day Survival (n = 1,197)</th>
<th>30-Day Mortality (n = 50)</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender (% Male)</td>
<td>98%</td>
<td>98%</td>
<td>NS</td>
</tr>
<tr>
<td>Age (Years [IQR])</td>
<td>24 [21–29]</td>
<td>23 [21–29]</td>
<td>NS</td>
</tr>
<tr>
<td>Injury Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blast</td>
<td>73%</td>
<td>62%</td>
<td>NS</td>
</tr>
<tr>
<td>Penetrating</td>
<td>17%</td>
<td>32%</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Blunt</td>
<td>9%</td>
<td>2%</td>
<td>NS</td>
</tr>
<tr>
<td>Burn (Burn Only)</td>
<td>5% (2%)</td>
<td>18% (4%)</td>
<td>&lt;0.01 (NS)</td>
</tr>
<tr>
<td>ISS</td>
<td>22 [16–29]</td>
<td>33 [27–49]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Preflight Parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCS</td>
<td>7 [3–15]</td>
<td>3 [3–3]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Intubated</td>
<td>60%</td>
<td>96%</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Systolic Blood Pressure (mm Hg)</td>
<td>121 [109–134]</td>
<td>120 [102–143]</td>
<td>NS</td>
</tr>
<tr>
<td>Diastolic Blood Pressure (mm Hg)</td>
<td>62 [55–69]</td>
<td>64 [58–84]</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Heart Rate (bpm)</td>
<td>100 [84–115]</td>
<td>110 [88–124]</td>
<td>0.03</td>
</tr>
<tr>
<td>Shock Index</td>
<td>0.8 [0.7–1.0]</td>
<td>0.9 [0.7–1.1]</td>
<td>NS</td>
</tr>
<tr>
<td>INR</td>
<td>1.1 [1.0–1.3]</td>
<td>1.4 [1–1.7]</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Hematocrit</td>
<td>30 [27–35]</td>
<td>31 [27–38]</td>
<td>NS</td>
</tr>
<tr>
<td>Platelet Count (x10⁹/L)</td>
<td>116 [93–153]</td>
<td>100 [72–123]</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Base Deficit</td>
<td>0 [−2–2]</td>
<td>−2.5 [−6.3 to −0.5]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Preflight Resuscitation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crystalloid (%Yes)</td>
<td>88%</td>
<td>94%</td>
<td>NS</td>
</tr>
<tr>
<td>Colloid (%Yes)</td>
<td>31%</td>
<td>30%</td>
<td>NS</td>
</tr>
<tr>
<td>Blood Products (%Yes)</td>
<td>70%</td>
<td>88%</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Plasma Units, (%Yes)</td>
<td>9 [4–18], 62%</td>
<td>10 [5–32], 82%</td>
<td>NS, &lt;0.01</td>
</tr>
<tr>
<td>RBC Units, (%Yes)</td>
<td>9 [4–18], 66%</td>
<td>14 [6–29], 66%</td>
<td>0.03, NS</td>
</tr>
<tr>
<td>Platelets Units, (%Yes)</td>
<td>3 [1–6], 43%</td>
<td>2 [1–6], 62%</td>
<td>NS, &lt;0.01</td>
</tr>
<tr>
<td>Cryoprecipitate Units, (%Yes)</td>
<td>4 [2–10], 29%</td>
<td>10 [2–30], 44%</td>
<td>NS, 0.03</td>
</tr>
<tr>
<td>Whole Blood Units, (%Yes)</td>
<td>4 [3–9], 7%</td>
<td>19 [7–29], 26%</td>
<td>&lt;0.01, &lt;0.0001</td>
</tr>
<tr>
<td>Factor VIIa (%Yes)</td>
<td>9%</td>
<td>34%</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

GCS, Glasgow Coma Scale score; IQR, interquartile range; ISS, injury severity score; bpm, beats per minute; NS, nonsignificant; INR, international normalized ratio; RBC, red blood cells.
of fresh whole blood, prothrombin time, and whether the patient was intubated or not.

A 2014 publication of CCATT data by Ingalls et al \(^7\) compared survival with physiological and injury characteristics mirrors the findings of this study; ISS, volume of whole blood transfused, and prothrombin time (reported as INR) were significantly associated with increased 30-day mortality. This study also reported that the worst recorded base deficit was associated with an increased 30-day mortality. \(^7\)

There is however no examination of other preflight characteristics or an analysis of the data that could be used to rule in or rule out outcomes—additionally 60% of the CCATT population were excluded from data analysis, predominantly because the time of arrival at Role 4 was unavailable. The remainder of published studies of CCATT patients examine the incidence of secondary brain injury, \(^24\) the epidemiology of patients on CCATT, \(^11,25\) review the literature and provide expert opinion on AE of patients with traumatic brain injury, \(^26\) or outcomes limited to the CCATT flight itself. \(^5\)

It is likely that all of the variables that are significantly associated with ICU outcomes in the Ingalls et al study \(^7\) and ours are in fact surrogates for injury severity, and which perform better than ISS (or other individual variables) alone. Further research that better describes the use of physiological and anatomical injury data to predict ICU outcomes could also be of use in planning the management of combat trauma patients.

Our study has limitations. The association of preflight variables and outcomes are only valid in this examined cohort, and application in another population should be undertaken to check their general applicability (e.g., using the U.K. Critical Care Air Support Team database). Our study included predominantly young, male, physically fit military patients with a high proportion of blast injury, it may therefore not be applicable to civilian settings, and even civilians injured in combat. The data in this study, although collected prospectively, were reanalyzed retrospectively and thus is limited to the accuracy and availability of the medical records. The variable “intubated preflight” is included in both outcome models. This variable, like others already mentioned, is probably a surrogate for true injury severity. It may be postulated that the decision to keep a patient intubated for an intercontinental flight is not based entirely on the patient’s physiology or injury pattern, but instead a logistical decision to increase patient safety during AE transfer. However, even if this is the case, this complex medical decision is unlikely to be significantly different in future operations, and we believe is, therefore, still a valid preflight variable to include in the model. The DoDTR calculated ISS, used in this study, would not be available preCCATT flight, however an initial ISS is calculated in theatre by trauma nurse coordinators. We were unable to access this initial ISS, and therefore there is an assumption that any difference between these two ISS would not significantly affect its utility in this model.

CCATT and military ICU practitioners should be aware that the variables identified in this study may have a predictive effect in the requirement for prolonged ventilation, and therefore the highest level of ICU care, and 30-day mortality. These data may, therefore, have utility in planning ongoing ICU provision, and with urgent evacuation of combat patients, particularly with a large number of casualties, and or with limited ICU resources.

In our large CCATT study a number of preflight variables were independently associated with a need for prolonged ventilator time (and therefore the highest level of critical care), and composites of these produced reasonable rule in and rule out percentages that could be used by military medical planners at times of increased demand or limited capability.

**ACKNOWLEDGMENT**

Department of Defense Joint Program Committee (JPC-6) grant awarded through the Geneva Foundation (W81XWH-13-2-0033).

**REFERENCES**

Dysphagia Management and Research in an Acute-Care Military Treatment Facility: The Role of Applied Informatics

Nancy Pearl Solomon, PhD, CCC-SLP*; Angela M. Dietsch, PhD, CCC-SLP†*; Katie E. Dietrich-Burns, MS, CCC-SLP‡*; Edda L. Styrmisdottir, MSc‡; Christopher S. Armao, PhD‡

ABSTRACT  Purpose: This report describes the development and preliminary analysis of a database for traumatically injured military service members with dysphagia. Methods: A multidimensional database was developed to capture clinical variables related to swallowing. Data were derived from clinical records and instrumental swallow studies, and ranged from demographics, injury characteristics, swallowing biomechanics, medications, and standardized tools (e.g., Glasgow Coma Scale, Penetration-Aspiration Scale). Bayesian Belief Network modeling was used to analyze the data at intermediate points, guide data collection, and predict outcomes. Predictive models were validated with independent data via receiver operating characteristic curves. Results: The first iteration of the model (n = 48) revealed variables that could be collapsed for the second model (n = 96). The ability to predict recovery from dysphagia improved from the second to third models (area under the curve = 0.68 to 0.86). The third model, based on 161 cases, revealed “initial diet restrictions” as first-degree, and “Glasgow Coma Scale, intubation history, and diet change” as second-degree associates for diet restrictions at discharge. Conclusion: This project demonstrates the potential for bioinformatics to advance understanding of dysphagia. This database in concert with Bayesian Belief Network modeling makes it possible to explore predictive relationships between injuries and swallowing function, individual variability in recovery, and appropriate treatment options.

INTRODUCTION  Clinical providers in military treatment facilities (MTFs) face multiple challenges in evidence-based clinical decision-making. First, given the tremendous range of care-related decisions that occur within a single patient’s hospital admission, there is relatively limited literature to inform many of those choices. Second, whatever literature is available is likely to be based on civilian sample populations and thus may have limited application to the military population being served. Third, an abundance of clinical data regarding outcomes of various diagnoses and interventions exists, but this information is not readily available in a form that can be analyzed statistically.

Dysphagia, or swallowing dysfunction, is one example of a clinical sequela that has limited research evidence to guide caregivers in MTFs. Internal estimates suggest that up to 20% of military active duty service members (ADSMs) injured during deployment who were admitted to Walter Reed Army Medical Center (WRAMC) and Walter Reed National Military Medical Center (WRNMMC) exhibited dysphagia. Difficulty eating and swallowing can have significant health care implications including aspiration pneumonia, dehydration, malnutrition, and in severe and prolonged cases, death. However, the published literature lacks systematic examination of dysphagia in the ADSM population.

A preliminary analysis of 50 blast-injured ADSMs referred for dysphagia evaluation at WRAMC during the early stages of the conflicts in Iraq and Afghanistan revealed that all demonstrated some aspect of pharyngeal dysphagia.1 Deficits reportedly included delayed onset of the pharyngeal swallow response, pharyngeal residue, and tracheal aspiration. Reports from civilian populations with traumatic brain injury or spinal cord injury tend to support these results. For example, videofluoroscopic swallowing studies (VFSSs) from 53 patients with dysphagia subsequent to closed-head injuries revealed 81% with delayed or absent swallow response, 50% with abnormal tongue control, and aspiration in greater than one-third of the patients.2 Records reviewed from 131 patients with spinal cord injury identified the co-occurrence of brain injury, history of spinal surgery, and the presence of a tracheostomy tube to be predictive of dysphagia and tracheal aspiration.3 These results must be applied with caution by clinicians in MTFs, however, because traumatically injured ADSMs (TI-ADSM) are generally younger and more physically fit than the civilian patients included in these studies, and the mechanisms of injury can differ substantially from those incurred in the civilian population.

To better understand the nature, course, and management of dysphagia in a TI-ADSM population, this project was designed to systematically review and catalog patients referred...
for dysphagia evaluation after sustaining combat-related injuries over the past decade. By organizing this information into a normalized, relational database, clinicians and researchers can answer questions that previously were available only through expert opinion and clinical experience. The purpose of this report is to describe the development and initial examination of the database to illustrate how existing clinical data can be harnessed by clinician-investigators to inform best practice patterns for patient referral, prognosis, and management.

**METHODS**

**Patient Selection**

Patients eligible for inclusion were (1) TI-ADSM admitted to a National Capital Regional MTF (WRAMC, National Naval Medical Center [NNMC], and WRNMMC) since the beginning of 2004; (2) referred to the Speech Pathology Clinic for a dysphagia evaluation that was completed during the inpatient hospitalization; and (3) between 18 and 50 years of age. Hospital coders provided lists of potential candidates by querying International Classification of Diseases, 9th Revision, Clinical Modification and “Healthcare Common Procedure Coding System” codes from inpatient records. Candidates’ records were screened by research team members to identify and enroll qualifying cases. Once enrolled, patients whose recorded VFSS were available for reanalysis were prioritized for detailed medical record extraction into the database.

Records from patients admitted before August 2012 were exempt from providing informed consent; inpatients recruited thereafter consented and were prospectively enrolled into the study in accordance with the rules and regulations of the WRNMMC Department of Research Programs Institutional Review Board (IRBNet no. 357205).

**Database Development**

**Structure**

The WRNMMC Dysphagia Database was constructed in Microsoft Access. Its overall design utilized a normalized relational format to accommodate the longitudinal nature of prolonged hospitalizations. Figure 1 illustrates the basic design of the database with selected variables. Essentially, it is a layered tree-like structure wherein each subject has a unique identifier or primary key as the main trunk. Each patient’s hospital admission was assigned a secondary key that linked back to the primary key, and each date with relevant medical information was tagged with a tertiary “encounter” key that linked to the admission identification. Each layer of data included variables that were anticipated to be stable for the duration of time represented at that level, and one or more tables within Microsoft Access were utilized to capture data from each layer. For example, the “Accrual” table corresponded to the first layer described above and included the patient’s study identification code, date of birth, sex, race, ethnicity, and height. The second “Admission” layer included

---

**FIGURE 1.** General structure and contents of the Dysphagia Database. Each row represents a layer in the tree structure, and each box represents a table; the identification keys linking information between layers are italicized. Note that there may be multiple encounters (1 = admission, n or last = discharge) other encounters (2 through n-1) were triggered by changes in status or a speech-language pathology (SLP) consult. SLP consults generated an additional layer of data. SLP Consult 1 typically involved a clinical (or “bedside”) swallow examination and SLP Consult 2 often included an instrumental (videofluoroscopic/VFSS or fiberoptic-endoscopic swallow study. Retrospective analyses (Penetration-Aspiration Scale/PAS, Modified Barium Swallow Impairment Profile/MBSImP, and biomechanical measures) for available archived VFSS recordings were completed, and medications taken on the day of the VFSS were entered; these data were not included in the preliminary models illustrated in this article. Only selected and simplified variables are listed in the figure for clarity; see text or contact authors for more detail.
one table with the nature and severity of injuries sustained as well as dates of injury, admission, key consultations, and discharge. The third “Encounter” layer held information that was likely to change on a daily basis, such as artificial airway status, feeding status or diet orders, cognitive status, speech and voice status, and weight. An encounter was generated for every admission, discharge, and any change in status that was likely to affect dysphagia management. The fourth “Speech-Language Pathology (SLP) Consult” layer comprised several tables specific to each swallow-related encounter. Figure 1 illustrates a typical scenario in which the first SLP consult involves a clinical swallow examination, the second involves an instrumental swallow study, and subsequent consults involve additional evaluations, interventions, or changes in recommendations. Each was linked back to the key for the relevant encounter date. Date (or 1 day) was determined a priori to be the narrowest temporal window of relevance for this study.

The layered tree structure enabled the database to accommodate multiple entries of similar data (e.g., weight, diet status) efficiently while maintaining the ability to collapse data as necessary for specific analyses. Within each table, each column contained one variable and each row indicated a separate patient (in the “Accrual” table), admission (in the “Admission” table), or date of service (in the “Encounter” and “SLP Consult” tables). Thus, a given patient would have a single entry in the Accrual table, typically a single entry in the Admission table (rarely, a patient was redeployed, reinjured, and admitted for a second entry), and perhaps a dozen in the Encounter and SLP Consult tables.

Once the tables were finalized, forms were created within Microsoft Access to facilitate data entry. Forms were organized by topic, and each entry was linked to a single variable within the tables. Customized drop-down response menus provided fixed or binary choices for each question, usually with an option for “other” or “unknown.” This necessitated more questions/variables at the data entry stage, but allowed for better management of queries and Bayesian Belief Network (BBN) analysis later. Because the existing electronic health record (EHR) was mostly in narrative form rather than templates, investigators reviewed pages of narrative text to extract relevant variables and enter them into forms.

**Variables**

The process of determining which variables to include in the database took into account the scope of the database project’s goals, current research literature, clinical experience, and patient population demographics. Two research SLPs (AMD, KDB) with combined 25 years of previous medically based clinical experience generated lists of possible variables, organized them into relevant layers and categories with input from statistical collaborators, and beta tested them on an initial set of 10 patient retrospective records. Based on this testing, categories and variables were added, deleted, and modified as appropriate.

Standardized measures, whether available in the EHR or applied retrospectively to data within the medical record, were included in the database whenever possible. For example, nursing notes regularly provided Glasgow Coma Scale (GCS) scores, which offer a relatively universal measurement of cognitive status. Though not a part of the original EHR, an Injury Severity Score was determined for each admission based on injury characteristics described in history and physical notes.

To address questions about predictors for dysphagia symptoms and outcomes, a broad range of medical variables was included in the database. Complete accounting of the mechanism and site(s) of injury can enable investigation of the primary causes of dysphagia as well as the impact of comorbidities on functional outcomes. Details about the timing and results of consult from a variety of potentially relevant disciplines (i.e., neurology, neurosurgery, otolaryngology, audiology, maxillofacial/dental, pulmonary, and gastroenterology) allow for exploration of the course of medical intervention juxtaposed with identification of and recovery from dysphagia. By tracking changes in airway management, nutritional support, and cognitive status, the database offers tremendous capacity to explore dysphagia-specific questions such as the effects of prolonged versus repeated intubations on swallow function as well as more general questions such as average days of intubation or parenteral nutritional support in this population.

Specific to dysphagia, we reviewed archived medical records and SLP notes to assign a score on the swallowing component of the American Speech-Language-Hearing Association’s National Outcome Measures System (ASHA NOMS). In addition, the database accommodated detailed information from each swallow-related encounter. Virtually every type of diet modification, direct treatment modality, and compensatory strategy recommended in the course of dysphagia management was captured in variables within the SLP Consult tables.

Additional dysphagia metrics, not included in the preliminary BBN models presented in this article, included Penetration-Aspiration Scale (PAS) scores, the Modified Barium Swallow Impairment Profile (MBSImP), and biomechanical/morphometric analysis. These analyses are being determined retrospectively for individual swallow trials from archived VFSS recordings. Finally, because medications have potential sensory, motor, visceral, and cognitive effects that can impact swallowing, each VFSS Analysis table includes medication information (names, dosages), generated by EHR IT specialists. Before database entry, we coded medications by class (e.g., opioid, antipsychotic, anticonvulsant, and prokinetic) based on pharmacological standards and potentially relevant side effects.

**Statistical Analysis**

BBN classification was the primary technique used to analyze these data. The analysis is based on Bayes’ Theorem that
relates prior probabilities to future probabilities. It effectively links correlated variables based on how they are distributed in a dataset. Recent approaches to BBN modeling utilize machine learning, which permits predictions of outcomes based entirely on data (evidence). This study used a machine learning tool, FasterAnalytics, that learns the structure and joint probability distributions of source study data. It then represents the resulting associations in a hierarchical graphical format. FasterAnalytics allows users to examine cascades of codependent factors that may contribute to a given outcome and to make predictions regarding the potential impact of an intervention on an outcome of interest.10,11

BBNs are inherently robust and tolerant of heterogeneous and incomplete data sets, making them powerful tools for analyzing complex clinical data. Several iterations of BBN models were compiled at different stages of database entry. Once the database contained data from nearly 100 patients, a BBN model was trained using 80% of randomly selected cases. The model was tested with the remaining 20%, generating a receiver operating characteristic curve and calculating area under the curve (AUC). Preliminary analyses guided refinement of the dataset and model characteristics for subsequent analyses.

Sample Size Estimation
The BBN analysis will accept and model small sample sizes with large numbers of variables, and the model can continuously update as data are added. A power analysis is, therefore, not required for BBN, though ideally the model will eventually stabilize such that it becomes robust even with the addition of future cases.

Sample size estimation was conducted to accommodate traditional statistical procedures, taking into account the availability of records that met eligibility requirements and feasibility for a 2-year period of data retrieval and entry. Results indicated that the desired statistical power could be achieved with a sample of 200 records. For estimates of proportions, a sample of 100 subjects provides a 95% confidence interval (CI) of ±10% and a sample of 200 subjects provides a 95% CI of ±7%. Controlling the probability of a Type I error at α = 0.05, a sample of 200 subjects has 80% power to detect a correlation as low as r = 0.20. For exploratory regression analyses, approximately 15 to 20 records for each of independent variable are necessary for a multivariate model. Therefore, with 200 records, up to 10 predictor variables can be considered.

RESULTS
Potential candidates between the ages of 18 and 50 admitted to WRMC between January 2004 and April 2010 were screened, leading to the identification of 455 eligible retrospective cases. In addition, 59 eligible patients admitted to NNMC or WRNMMC between April 2010 and August 2012 were identified. Finally, 43 eligible participants admitted to WRNMMC...
since September 2012 consented to be included. In sum, the pool of potential patients eligible for database entry numbers 557 to date. Separate preliminary BBN analyses were conducted when 48, 96, and 161 cases had been extracted and compiled into the database. The reader is cautioned that these models are not comparable because the structure of the tree and the amount of branch pruning differed markedly as the sample size grew.

The initial model was intended only as an examination of relational associations between variables at the time of the first swallow-related encounter. Figure 2 illustrates this first iteration of 48 cases, which yielded a large and complex tangle of associations that was virtually undecipherable. This was because of the complexity of the patient population, the large number of variables included, and the relatively small sample size.

Although meaningful predictive relationships could not be extrapolated from this model, it was helpful in delineating variables that were inherently related (e.g., “maxillofacial fracture” and “maxillofacial consult”). This guided flattening for the second iteration of the model (Fig. 3), which yielded an AUC of 0.68. The primary outcome variable, selected for both its functionality and objectivity, was a discharge score of 5 to 7 on the ASHA NOMS, indicating that the patient was allowed an unrestricted diet with minimal or no restrictions or cues. The model was trained on 80% of 96 cases and validated with the remaining 20% of the data. Certain factors started to emerge as relationally relevant. Specifically, pre-admission intubation, tracheostomy, broken or damaged teeth, and ethnicity were first-degree associates for the primary outcome variable.

At 161 cases, the BBN model became somewhat more refined and interpretable (Fig. 4). It revealed that diet restrictions at the time of the first dysphagia consult was a first degree associate for the primary outcome variable of an unrestricted diet at discharge. Second-degree associates included GCS score, intubation status, and diet change and restrictions based on the first dysphagia evaluation. Model validation at this stage yielded an AUC of 0.86.

**DISCUSSION**

Clinicians treating unique populations such as TI-ADSMs are often left to determine best practices without appropriate supporting literature, and the prospect of addressing these gaps can be daunting to clinicians and researchers alike. In the case of dysphagia, the lack of existing research regarding this complex consequence of combat-related injuries inspired the development of a comprehensive database rooted in existing clinical records and based on modern informatics principles.

Resources for designing and executing the project included a readily available software platform, experienced clinicians to identify potentially germane variables, guidance from existing literature, and supplemental input from research statisticians regarding optimal database organization. Investigators relied on clinical experience and current literature to interpret preliminary BBN models for refinement of the database. Subsequent analyses were then able to elucidate associations between injury characteristics, medical and surgical

**FIGURE 3.** Preliminary Bayesian Belief Network model based on 96 cases using a discharge score of 5 or greater on the American Speech-Language-Hearing Association’s National Outcome Measures System/ASHA NOMS for swallowing (indicating unrestricted diet) as the outcome variable. Area under the curve/AUC = 0.68 for model validation.
interventions, recovery benchmarks, and the outcome variable of interest. These relationships will motivate additional lines of inquiry to which traditional statistical methodologies can be applied to test hypotheses for associations revealed within the model.

Dysphagia in the TI-ADSM population typically accompanies serious injuries that require prolonged hospitalization with multidisciplinary care. Analysis of the partially developed BBN model suggests that history of intubation, GCS scores, and dysphagia management recommendations at early stages of rehabilitation are predictive of swallow function at the time of discharge. These results are consistent with those previously reported in other populations.12–16

The process of extracting and analyzing data revealed a number of challenges and subsequent opportunities for improvement in clinical workflow and research design. For example, the mostly narrative format of the existing EHR required investigators to read through volumes of records to extract relevant variables. More standardized note templates with predefined fields and responses in EHR would have obvious clinical benefits and could also enable relevant data to be populated directly into the database. In addition, analysis of archived VFSS recorded over the past decade revealed differences in protocols, framing, and temporal resolution that affected study interpretability. This observation led the Speech Pathology and Radiology services to collaborate and recommend procedures and equipment settings that balance the needs of information gathering and radiation exposure for all clinical VFSS procedures.6,17,18 The comprehensive nature of a large database such as the WRNMMC Dysphagia Database is likely to identify many other trends and opportunities for improvement and investigation across the continuum of multidisciplinary health care.

The next iteration of the BBN analysis will be initiated after we complete entries for 200 patients into the WRNMMC Dysphagia Database. In addition, the completed database will include medications and results from VFSS analyses. The inclusion of standardized dysphagia rating scales (PAS, MBSImP) and quantitative measures of swallowing biomechanics will enrich the database with information that is meaningful across settings. Secondary analyses using multivariate regression and hypothesis testing are planned to further characterize the study sample and to explore relationships between specific variables of interest. The volume and scope of data available supports a wide range of inquiries. BBN analysis will identify superfluous variables that increase bias and overfitting, thereby facilitating selection of key variables to include in regression models. Ultimately, the goal is to

FIGURE 4. Bayesian Belief Network model based on 161 cases using a discharge score of 5 or greater on the American Speech-Language-Hearing Association’s National Outcome Measures System/ASHA NOMS for swallowing (indicating unrestricted diet) as the outcome variable. Area under the curve/AUC = 0.858 for model validation.
make a simplified, user-friendly database available to other Department of Defense and civilian clinics so that individual patient data can be entered, and a BBN model can be used to predict clinical outcomes based on each patient’s specific injury and dysphagia characteristics.

SUMMARY
This article utilizes a multiyear Department of Defense funded project on dysphagia in TI-ADSMs to illustrate how existing medical records can be harnessed to reveal associations between various aspects of clinical care. These relationships could help clinicians predict the effects of care decisions, and guide the exploration of clinically relevant research questions. BBN modeling was the primary analysis tool, and clinically fluent investigators plus a highly accessible relational database structure formed the basis for project development. Early iterations of the BBN model on 48 cases, followed by 96 cases, and at the time of this report 161 cases, were useful for flattening data and elucidating predictive relationships that warrant further statistical analysis. The validity of the BBN model is already high (0.86) and is expected to improve further as more cases are added to the database. BBN methodology has clinical relevance, as it can guide treatment decisions for dysphagia and other complex medical issues. The richness of the Dysphagia Database invites queries about many aspects of wounded warrior injuries and care including hospitalization, medications, comorbid impairments, and treatment outcomes. Results from these analyses have the potential to enhance dysphagia care and maximize functional outcomes for patients with polytraumatic injuries.

ACKNOWLEDGMENTS
The authors gratefully acknowledge the contributions of Todd Radano and Elizabeth Silvius towards database development and Bayesian modeling; Carolyn Hamm and Surbhi Paul for hospital-record coding; and Laura Cord, Sid Fels, Captain Gerald Grant, Andrew Ho, WJ Bryan Lewis, William Pearson, and Ian Stavness for their collaboration and support. This work is supported by a grant from the U.S. Defense Medical Research Development Program (W81XWK-12-2-0021) administered through the Henry M. Jackson Foundation for the Advancement of Military Medicine.

REFERENCES
En Route Use of Analgesics in Nonintubated, Critically Ill Patients Transported by U.S. Air Force Critical Care Air Transport Teams

Alejandra G. Mora, BS*; Victoria J. Ganem, RN, BSN*; Alicia T. Ervin, RN, BSN*; Joseph K. Maddry, MD†; Vikhyat S. Bebarta, MD†

ABSTRACT
Introduction: U.S. Critical Care Air Transport Teams (CCATTs) evacuate critically ill patients with acute pain in the combat setting. Limited data have been reported on analgesic administration en route, and no study has reported analgesic use by CCATTs. Our objective was to describe analgesics used by CCATTs for nonintubated, critically ill patients during evacuation from a combat setting. Methods: We conducted an institutional review board–approved, retrospective review of CCATT records. We included nonintubated, critically ill patients who were administered analgesics in flight and were evacuated out of theater (2007–2012). Demographics, injury description, analgesics and anesthetics, and predefined clinical adverse events were recorded. Data were presented as mean ± standard deviation or percentage (%). Results: Of 1,128 records, we analyzed 381 subjects with the following characteristics: age 26 ± 7.0 years; 98% male; and 97% trauma (70% blast, 17% penetrating, 11% blunt, and 3% burn). The injury severity score was 19 ± 9. Fifty-one percent received morphine, 39% hydromorphone, 15% fentanyl, and 5% ketamine. Routes of delivery were 63% patient-controlled analgesia (PCA), 32% bolus intravenous (IV) administration, 24% epidural delivery, 21% continuous IV infusions, and 9% oral opioids. Patients that were administered local anesthetics (nerve block or epidural delivery) with IV opioids received a lower total dose of opioids than those who received opioids alone. No differences were associated between analgesics and frequency of complications in flight or postflight. Conclusion: About half of nonintubated, critically ill subjects evacuated out of combat by CCATT received morphine and more than half had a PCA. In our study, ketamine was not frequently used and pain scores were rarely recorded. However, we detected an opioid-sparing effect associated with local anesthetics (regional nerve blocks and epidural delivery).

INTRODUCTION
U.S. Air Force Critical Care Air Transport Teams (CCATTs) evacuate critically ill and injured patients within and out of the combat setting. To provide expeditious transport of critically ill patients to higher levels of definitive care, CCATTs were developed to consist of a physician, a critical care nurse, and a respiratory therapist per team.1 CCATT is capable of providing care similar to that of a hospital-based intensive care unit (ICU); and thus, it is at times referred to as “a flying ICU” where en route care is a continuation of patient care provided in the preceding medical facility.1–3 However, the CCATT environment is different in that patients experience additional stressors inherent to flight transport.4 In the recent Iraq and Afghanistan conflicts, patients required prompt movement after injury and CCATTs had to manage patients with severe, acute pain during long flights of up to 8 or more hours. Providers are challenged with addressing the pain needs of the patient while simultaneously maintaining hemodynamic stability. In addition, the isolated aircraft environment poses unique difficulties not seen in the hospital setting such as vibration, g-force, temperature extremes, gas expansion, dehydration, hypoxia, and movement.4–6 Some of these conditions worsen pain and may prompt an increase in analgesic use. Buckenmaier et al7 surveyed injured soldiers at Landstuhl Regional Medical Center (LRMC) in Germany following recent evacuation from Iraq and Afghanistan. In this study, only 65% attained 50% or less reduction of their pain during transport supporting the dif-
En Route Analgesic Use in Nonintubated, Critically Ill Patients by CCATTs

pain management after injury is a known priority for service members, the Air Mobility Commander, and the U.S. Air Force Medical Service. However, an evidence-based understanding is needed to provide appropriate clinical standardization and adequate clinical practice guidelines for CCATTs and other en route care platforms. Few studies exist that describe the analgesic methods used by CCATT providers, and there are no published reports on patient outcomes. The objectives of our study were to (1) describe analgesics used by CCATTs; (2) document clinical complications in flight; and (3) report postflight outcomes in nonintubated, critically ill patients transported out of the combat setting by CCATTs. The nonintubated, critically ill patients in our study could be considered to be most vulnerable to flight stressors (i.e., hypoxia), analgesic-induced changes in vitals, and consciousness of pain levels experienced. We intended to characterize current practices in relation to relevant clinical complications and outcomes. With this, we expect to facilitate understanding of CCATT analgesic modalities and associated risks in the combat setting.

METHODS
Following approval from the Wilford Hall Ambulatory Surgical Center Institutional Review Board, we performed a retrospective record review.

We requested CCATT medical records for this study under a data sharing agreement between the CCATT Pilot Unit and the En route Care Research Center. The available CCATT medical records (n = 1,128) were reviewed and subjects were enrolled on the basis of the following inclusion criteria: (1) subjects transported via CCATTs from Afghanistan or Iraq to LRMC between January 2007 and December 2011, (2) subjects who were not intubated, and (3) subjects who received analgesics during flight. CCATT medical records of intratheater transports and transports to the United States were excluded along with ventilated patients. We developed an electronic study database (Microsoft Excel 2010; Microsoft Corporation, Redmond, Washington) with predefined fields where we entered data abstracted from CCATT medical records. Research staff was trained to interpret CCATT medical records, abstract data, determine events using a standardize tool, and enter data consistently. Missing data were not imputed or included in analysis; and thus, data reported reflect provider documentation. The secure electronic study database was maintained and accessible only to institutional review board–approved study personnel.

The data we abstracted included demographics, injury description, departure and arrival locations, inflight vital signs and hemodynamics, laboratory values, analgesics, complications, and procedures. Preflight vital signs, hemodynamic measures, and laboratory values were recorded when available to establish a baseline for each subject. When available, pain assessments were collected as pain assessed (yes or no) and pain scale score (score: 0–10). Complications were predefined and were identified by calculating a percentage change from baseline using standard clinical normal values in addition to provider descriptions of events from the CCATT progress notes (Table I). Analgesics and local anesthetics data were collected to include dose and route. Analgesics comprised of opioids (morphine, fentanyl, and hydromorphone) and ketamine. A morphine equivalent (ME) dose was calculated for fentanyl and hydromorphone for statistical analysis. Mode of administration was recorded as continuous IV infusion, bolus IV infusion, patient-controlled analgesia (PCA), or oral administration. Epidurals and local regional blocks were considered to be local anesthetics in this study. Local anesthetics included lidocaine, ropivacaine, bupivacaine, or opioids. For analysis, subjects were categorized on the basis of medication and method of delivery. Subjects that received IV opioids (morphine, fentanyl, or hydromorphone) were evaluated (OPIOID). The following groups were developed to characterize subjects who received concomitant medications for en route pain management: IV opioids and a local anesthetic (OPIOID-LOCAL ANES); IV opioids and ketamine (OPIOID-KET); and IV opioids, ketamine, and local anesthetic (OPIOID-KET-LOCAL ANES). The local anesthetic (LOCAL ANESTHETIC) group received local anesthetics and the ketamine (KETAMINE) group received ketamine alone. Subjects whose pain management consisted of bolus IV administrations only or oral administration only were categorized as NO MAINTENANCE. Propofol, midazolam, and lorazepam were defined as sedative adjuncts.

<table>
<thead>
<tr>
<th>TABLE I. Definitions for Inflight Complications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperthermia</td>
</tr>
<tr>
<td>Respiratory</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Hemodynamic</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Heart Rate</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Decreased Urine Output</td>
</tr>
<tr>
<td>Bleeding</td>
</tr>
<tr>
<td>Neurologic</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Medication Reaction</td>
</tr>
<tr>
<td>Pain</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

bpm, beats per minute; CCATT, Critical Care Air Transport Team; CVP, central venous pressure; FiO2, fraction of inspired oxygen; L/min, liters per minute; MAP, mean arterial pressure; pCO2, blood carbon dioxide level; SBP, systolic blood pressure; SpO2, blood oxygen saturation.
Outcome data for enrolled subjects were obtained from the Department of Defense Trauma Registry (DoDTR). The DoDTR data included postflight vital signs, hemodynamic measures, and laboratory values. The first set of laboratory results at LRMC was considered to be postflight. In addition, we collected postflight outcomes such as the incidence of complications experienced at LRMC, number of days on mechanical ventilation, number of days in an ICU, total length of hospital stay, and mortality through 30 days. If the same complication was present preflight and postflight, the complication was not considered as new.

STATISTICS
We performed a descriptive analysis of the data. Data were reported as frequencies and proportions. We statistically compared OPIOID vs. OPIOID-LOCAL ANES groups and excluded other groups (OPIOID-KET, LOCAL ANESTHETIC, KETAMINE, and NO MAINTENANCE).

Statistical analysis was conducted using SAS JMP version 10 (Cary, North Carolina) and Microsoft Excel 2010 (Microsoft Corporation). Categorical variables were compared using $\chi^2$ or Fisher’s exact test when appropriate and were reported as percentages. After conducting the Shapiro–Wilk test for normality, Student’s $t$ tests were performed for continuous variables and summarized as mean (± standard deviation [SD]) or median (interquartile range). All statistical testing was two-sided, with a significance level set at $p < 0.05$.

RESULTS
Overall Subject Demographics
Out of 1,128 CCATT medical records reviewed, 381 subjects were not ventilated and received analgesics en route by CCATTs while being evacuated out of the zone of operations to LRMC during the study period. Subjects were a mean age of 26 (SD ± 7.0) years and predominantly male (98%) with traumatic injuries (97%). The traumatic injuries were 70% blast, 11% blunt, 17% penetrating, and 3% burn. One percent had an inhalational injury. Subjects with traumatic injuries had a mean injury severity score (ISS) of 19 (SD ± 9.0) (Table II).

Overall Pain Management
Most subjects (98%, $n = 372$) were administered opioids en route. Five percent of our subjects ($n = 18$) received ketamine en route. Almost a quarter (24%) received a local anesthetic (regional block or epidural). Fifteen percent received a sedative during transport. Pain score assessments were available for 5% and of those the median score was 7 (6–9).

Of the 381 patients, 63% were transported with a PCA, 32% received bolus IV administration, 24% epidural delivery, 21% continuous IV infusions, and 9% oral opioids. The majority (86%) of subjects receiving hydromorphone were transported with a PCA (Table III).
TABLE III. Summary of Opioid and Ketamine Mode of Delivery With Median Doses Administered by CCATT

<table>
<thead>
<tr>
<th>Group</th>
<th>Morphine</th>
<th>Hydromorphone</th>
<th>Fentanyl</th>
<th>Ketamine</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCA, % (n)</td>
<td>54% (105)</td>
<td>86% (128)</td>
<td>&lt;2% (1)</td>
<td>—</td>
</tr>
<tr>
<td>IV gtt, % (n)</td>
<td>8% (16)</td>
<td>5% (8)</td>
<td>63% (37)</td>
<td>100% (18)</td>
</tr>
<tr>
<td>Total Dose per 7 Hours, Median(IQR), (n)</td>
<td>1.1 (0.7–4.4), (13)</td>
<td>0.4 (0.1–0.7), (4)</td>
<td>3.6 (0.9–9.3), (19)</td>
<td>1.7 (0.6–5.0), (11)</td>
</tr>
<tr>
<td>Total Dose, Median(IQR), (n)</td>
<td>8.0 (5.0–31.0), (13)</td>
<td>2.5 (1.0–4.75), (4)</td>
<td>25.0 (6.5–65.0), (19)</td>
<td>12.0 (4.0–35.0), (11)</td>
</tr>
<tr>
<td>IVP, % (n)</td>
<td>43% (83)</td>
<td>16% (24)</td>
<td>42% (25)</td>
<td>0% (0)</td>
</tr>
<tr>
<td>Total Dose per 7 Hours, Median(IQR), (n)</td>
<td>0.7 (0.6–1.1), (79)</td>
<td>0.3 (0.1–0.3), (25)</td>
<td>0.7 (0.3–9.3), (22)</td>
<td>—</td>
</tr>
<tr>
<td>Total Dose, Median(IQR), (n)</td>
<td>5.0 (4.0–8.0), (79)</td>
<td>2.0 (1.0–2.0), (25)</td>
<td>5 (2–13), (22)</td>
<td>—</td>
</tr>
<tr>
<td>Total ME Dose, Median(IQR), (n)</td>
<td>6.0 (0.9–8.0), (88)</td>
<td>10.0 (96.7–13.3), (38)</td>
<td>6.8 (2.1–16.9), (22)</td>
<td>—</td>
</tr>
</tbody>
</table>

IV gtt, continuous IV infusion; IVP, IV opioid push; ME, morphine equivalent; PCA, patient-controlled analgesia.

Opioids

Of the OPIOIDS, more than half (51%) of subjects were administered morphine, whereas 39% received hydromorphone and 15% fentanyl. A similar total ME dose was administered to subjects when comparing doses on the basis of opioid type (Table III).

The NO MAINTENANCE group had a total of nine subjects who received opioids either IV bolus only (n = 4) or oral administration only (n = 5) during transport. Of the IV only subjects (n = 4), two were administered fentanyl, one morphine, and one hydromorphone. Three subjects in the NO MAINTENANCE group received a sedative and two were administered a fluid bolus during transport. No subjects received vasopressors and none had a documented inflight complication. Two subjects had an adverse respiratory event postflight. One incident in each of the following events was noted in this study groups: coagulopathy, unexpected bleeding, or abnormal laboratory. All survived and had a median of 0 (0–3) ventilator, 4 (2–5) ICU, and 11 (4–23) hospital days.

Local Anesthetics (Regional Nerve Block or Epidural Delivery)

Seven subjects had LOCAL ANESTHETIC only as treatment for pain management en route. All had trauma involving their extremities and skin. One subject received a vasopressor and three subjects were administered at least one fluid bolus in flight. The following incidences were recorded during transport: pain-related event (n = 1), hyperthermia (n = 4), and hypercapnia (n = 1). Five subjects had cardiac event postflight. Also, in the LOCAL ANESTHETIC group, we detected one incidence of the following events after transport: coagulopathy, bleeding, hemodynamic, respiratory, and renal. All survived and spent a similar number of days on the ventilator, in the ICU, and in the hospital in comparison to OPIOID and OPIOID-LOCAL ANES (2 [0–3], 4 [2–6], and 20 [4–45] days, respectively).

Ketamine

Although we evaluated all study groups, interpretation and the implication of findings were limited by sample size. In our study, 18 subjects received ketamine in flight. Of these, 14 received ketamine in conjunction with IV opioids (OPIOID-KET) and two received a combination of ketamine with IV opioids and a local anesthetic (OPIOID-KET-LOCAL ANES). Two subjects were administered ketamine alone (KETAMINE). Of the concomitant opioids administered, morphine (50%) was the most common followed by fentanyl (28%, n = 5) and hydromorphone (22%, n = 4). All subjects that received ketamine in flight survived.

The OPIOID-KET group received a median ME total dose of 18 mg (4–35) during flight. Only one subject in the OPIOID-KET group received a vasopressor, three were administered a fluid bolus, and one received blood products during transport. The following inflight incidences were recorded: pain-related event (n = 6), hypothermia (n = 4), tachycardia (n = 2), and hypercarbia (n = 5). During their hospital course following transport the following events were recorded: coagulopathy (n = 3), adverse cardiac event (n = 4), hemodynamic instability (n = 1), abnormal laboratory values (n = 2), adverse respiratory event (n = 6), and renal event (n = 1). The OPIOID-KET group had a median of 2 (0–4) days on ventilator, 6 (3–13) days in the ICU, and 17 (5–32) days in the hospital.

Both subjects in the OPIOID-KET-LOCAL ANES group received an ME total dose of 5 mg. The OPIOID-KET-LOCAL ANES group had a median of 1 (0–1) days on ventilator, 26 (2–5) days in the ICU, and 50 (35–65) days in the hospital. The KETAMINE group had a median of 8 (3–13) days on ventilator, 14 (6–22) days in the ICU, and 29 (23–34) days in the hospital.

Opioids vs. Opioids With Adjunct Local Anesthetic

Further analysis was conducted by comparing the OPIOID group to OPIOID-LOCAL ANES group. Also, the number of dose increases (4% vs. 1%; OPIOID vs. OPIOID-LOCAL...
ANES, respectively) and the frequency of starting a new adjunct analgesic (1% vs. 0%; OPIOID vs. OPIOID-LOCAL ANES, respectively) to manage subjects’ pain in flight were similar between groups. However, when comparing the total ME equivalent dose administered, the OPIOID group received more opioids than OPIOID-LOCAL ANES (12 mg [5–90] vs. 1 mg [2–18], p = 0.01). There was no difference in frequency of complications inflight (Table IV) or postflight (Table V). In addition, subjects had a similar number of days on ventilator, in the ICU, and in the hospital following flight with CCATTs (Table V).

**DISCUSSION**

Although studies have reported analgesics with associated short-term findings\(^1\) and others have described CCATT capabilities,\(^1,4–17\) this is the first study to report both the en route use of analgesics in critically ill patients transported by CCATTs in association with 30-day outcomes. Of the opioids evaluated in our study, morphine was administered to approximately half of the patients and was the most common. This was similar to that of a report of pain management practices of wounded service members.\(^18\) More than half of the subjects were transported with a PCA and for those patients the opioid administered tended to be hydromorphone. When comparing the ME doses among the various opioids there was no difference. About a quarter of our study population received either an epidural analgesic or a regional block for pain management en route. Ketamine was an infrequently administered analgesic used to manage the pain of nonintubated, critically ill patients.

We did not detect increased adverse events associated with opioid administration or, specifically, IV morphine. Although studies have reported adverse effects associated with morphine, such as hypotension and respiratory depression, in our study, the incidence of these adverse events was low. Such low frequencies may simply reflect infrequent or absent measures. For example, on CCATT flights laboratory measurements of respiratory status are typically only assessed with mechanically ventilated patients. Since patients were not mechanically ventilated, availability of blood gas measures (e.g., \(\text{pO}_2\), \(\text{PCO}_2\)) was low. If present, such data reflected undisclosed clinical factors that influenced provider decision-making. However, we used the combination of increased liters of oxygen delivered and oxygen saturation levels of blood along with laboratory parameters to more accurately evaluate respiratory events.

In addition to the overall assessment of analgesic practices, we sought to determine if there was an associated opioid-sparing effect related to use of epidural, regional block, or ketamine. Total ME doses were lower when local anesthetics were also present in comparison to patients that received IV opioids alone. Injury severity was similar between groups but, differences in injury types were noted. In a future study, pain management could be studied in relation to injury to facilitate understanding analgesic modality preferences and support evidence-based recommendations. However, in this study, we were able to conclude that epidural use during transport did not increase risk on the basis of the similar ME doses among the various opioids there was no difference. About a quarter of our study population received either an epidural analgesic or a regional block for pain management en route. Ketamine was an infrequently administered analgesic used to manage the pain of nonintubated, critically ill patients.

**TABLE IV.** Incidence of Event Experienced Inflight; Opioids (OPIOID) vs. Opioids With Adjunct Local Anesthetic (OPIOID-LOCAL ANES)

<table>
<thead>
<tr>
<th>Event</th>
<th>OPIOID</th>
<th>OPIOID-LOCAL ANES</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pain Event</td>
<td>32% (84/265)</td>
<td>35% (29/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Respiratory Event</td>
<td>9% (23/265)</td>
<td>12% (10/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Renal Event</td>
<td>3% (7/265)</td>
<td>4% (3/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Bleeding Event</td>
<td>0% (0/265)</td>
<td>0% (0/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Medication Event</td>
<td>&lt;1% (1/265)</td>
<td>0% (0/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Hypothermia</td>
<td>10% (21/220)</td>
<td>7% (5/75)</td>
<td>NS</td>
</tr>
<tr>
<td>Hyperthermia</td>
<td>30% (65/220)</td>
<td>44% (33/75)</td>
<td>NS</td>
</tr>
<tr>
<td>Bradycardia</td>
<td>7% (18/263)</td>
<td>4% (3/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Tachycardia</td>
<td>7% (18/263)</td>
<td>4% (3/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Hypotension</td>
<td>2% (6/261)</td>
<td>2% (2/81)</td>
<td>NS</td>
</tr>
<tr>
<td>Hypoxia</td>
<td>3% (7/250)</td>
<td>4% (3/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Hypocarbia</td>
<td>4% (2/50)</td>
<td>15% (2/13)</td>
<td>NS</td>
</tr>
<tr>
<td>Hypercarbia</td>
<td>48% (24/50)</td>
<td>46% (6/13)</td>
<td>NS</td>
</tr>
</tbody>
</table>

**TABLE V.** Outcomes Following Transport; Opioids (OPIOID) vs. Opioids With Adjunct Local Anesthetic (OPIOID-LOCAL ANES)

<table>
<thead>
<tr>
<th>Event</th>
<th>OPIOID</th>
<th>OPIOID-LOCAL ANES</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postflight Events</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coagulopathic</td>
<td>18% (48/265)</td>
<td>10% (8/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Bleeding</td>
<td>8% (20/265)</td>
<td>7% (6/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Cardiac</td>
<td>11% (28/265)</td>
<td>15% (12/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Hemodynamic</td>
<td>2% (6/265)</td>
<td>2% (2/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Infection</td>
<td>&lt;1% (1/265)</td>
<td>0% (0/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Abnormal Lab</td>
<td>15% (41/265)</td>
<td>17% (14/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Neurologic</td>
<td>0% (0/265)</td>
<td>0% (0/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Respiratory</td>
<td>28% (74/265)</td>
<td>33% (27/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Renal</td>
<td>2% (6/262)</td>
<td>2% (2/82)</td>
<td>NS</td>
</tr>
<tr>
<td>Ventilator Days</td>
<td>2 ± 4.2 1 (0–2)</td>
<td>2 ± 1.4 1 (0–2)</td>
<td>NS</td>
</tr>
<tr>
<td>ICU Days</td>
<td>5 ± 5.6 4 (3–6)</td>
<td>5 ± 3.7 5 (3–6)</td>
<td>NS</td>
</tr>
<tr>
<td>Hospital Days</td>
<td>21 ± 22.0 13 (5–30)</td>
<td>25 ± 24.2 17 (6–41)</td>
<td>NS</td>
</tr>
<tr>
<td>Mortality</td>
<td>&lt;1% (1/265)</td>
<td>0% (0/82)</td>
<td>NS</td>
</tr>
</tbody>
</table>

NS, not significant (p > 0.05).
as a result of the aircraft environmental stressors known to exacerbate pain. Of note, patients receiving ketamine in flight had higher ISS values. Ketamine administration may have been utilized as an adjunct analgesic to alleviate patients’ pain during the transport. In the current analysis, ketamine administration was not associated with increased risk. However, the low number of subjects receiving ketamine during transport limited statistical comparisons and implications.

Buckenmaier and O’Connor determined at least half or more of casualties failed to have a minimum of 50% reduction in pain with the administration of opioids. Although IV opioid administration has been regarded as an aggressive approach for pain management, it may be ineffective as evidenced by reports of continued severe pain even 90 minutes after administration. Pain assessments were rarely documented in our study population. As a result, we were not able to assess efficacy of various analgesic modalities, changes in pain levels (increase or decrease), or the persistence of pain experienced in flight. Our study population was not intubated and potentially capable of verbalizing pain experienced. The lack of documentation may be because pain assessments were infrequent or the provider flow sheet did not prompt an assessment. A better developed medical record with focused fields may facilitate improved documentation of pain scores. Prompting providers to perform assessments at predefined intervals may also be advised. Improved documentation would enable a better understanding of analgesic use and its association with the hemodynamics, clinical changes, adverse events, and outcomes of patients transported with CCATTs.

Although Bridges et al have published a study on less severely ill patients, a prospective evaluation may be needed to better evaluate the administration of analgesics for pain management during en route care of the critically ill patients. Although provider preference may influence pain management regimen, acuity and injury type may also prompt type of analgesics used. An analysis of analgesics and injury descriptions may be warranted. Understanding the relationship between opioid and other analgesics used with long-term outcomes is also needed. The use of opioids has been associated with chronic pain and dependence. As a result, adjunct nonopioid analgesics for pain management are recommended to decrease the amount of opioids administered. In addition, given the incidence (69%) of polytrauma observed in combat-related injuries, dual mechanism therapeutics are recommended as an opioid-sparing approach to reduce the likelihood of tolerance, and mitigate future dependence. Furthermore, negative correlations have been documented between post-traumatic stress disorder and early morphine or ketamine use. Future studies to characterize the influence of analgesic use on long-term cognitive and mental health effects are needed.

**Limitations**

There were inherent limitations in our retrospective study. As a record review, the data were limited to the accuracy and completion of the medical records performed by clinical providers. Our study personnel were trained and adhered to a standardized method for abstracting data. Parameters of interest were predefined to maintain consistency and adherence to a standardized data collection tool ensured data quality. Eighteen percent of IV dose data were unavailable; however, we did not substitute or interpolate missing data. We were not able to evaluate clinical interventions before CCATT flight or account for provider variations. In particular, provider preferences may be a confounding variable and may explain the differences in practice. Moreover, our results are reflective of young, relatively healthy military personnel; and so, our findings may not be generalizable to other military or civilian populations. The sample size of some of the groups was small; and thus, interpretation of findings is limited. Specifically, findings with regard to the use of ketamine and assessment of pain scores are limited. In addition, outcome data for medical patients were limited to availability in the DoDTR which primarily captures trauma data. We limited our outcome data to 30 days and outcomes beyond 30 days may yield different results.

**CONCLUSION**

In this first study of the en route use of analgesics for critically ill, nonintubated patients transported via CCATTs, we found that about half received morphine and over 60% had a PCA. We detected an opioid-sparing effect associated with local anesthetics (regional nerve blocks and epidural delivery), ketamine was infrequently used, and pain scores were rarely recorded. We did not identify a difference in the incidence of inflight complications or 30-day postflight outcomes associated with analgesic approaches during transport.

**ACKNOWLEDGMENT**

Funded by the Department of Defense Joint Program Committee (JPC-6) and awarded to the Geneva Foundation (W81XWH-13-2-0034).

**REFERENCES**

Is Hydroxyethyl Starch Safe in Penetrating Trauma Patients?

Casey J. Allen, MD; Xiomara D. Ruiz, MD; Jonathan P. Meizoso, MD; Juliet J. Ray, MD; Alan S. Livingstone, MD; Carl I. Schulman, MD, PhD; Nicholas Namias, MD; Kenneth G. Proctor, PhD

ABSTRACT  Objectives: For logistic reasons, a bolus of 6% hydroxyethyl starch (HES 450/0.7 in lactated electrolyte injection) is recommended for battlefield resuscitation even though it has risks of mortality and acute kidney injury (AKI) in certain patient populations. The purpose of this study was to test the hypothesis that victims of penetrating trauma have no increased risks of AKI and/or death when receiving a single bolus of HES during initial fluid resuscitation. Methods: 816 consecutive admissions with penetrating trauma were reviewed. Patients who died within 24 hours were excluded. Propensity scores and a 1:1 fixed ratio nearest neighbor matching were used to compare those who received HES to those who did not. Data were expressed as mean ± SD and significance was assessed at \( p < 0.05 \). Results: The cohort was 88% male, age 35 ± 14 years, injury severity score of 10 ± 10, with a 3.8% rate of AKI, and 3.2% rate of mortality. HES was administered to 121 (14.8%) patients. In HES and no HES propensity matched groups, the rate of AKI was 3.8% vs. 4.8% \( (p = 0.749) \) and the 90-day mortality rate was 3.8% vs. 4.8% \( (p = 0.749) \). Conclusion: An increased risk of mortality or AKI was not observed in penetrating trauma patients who were resuscitated with low volume HES.

INTRODUCTION

The Department of Defense Committee on Tactical Combat Casualty Care (TCCC) has issued guidelines for battlefield fluid resuscitation.1 For logistic reasons, in hypotensive or obtunded patients, a half-liter bolus of hydroxyethyl starch (HES) solution is recommended for initial resuscitation if blood products are not immediately available.2–4 With the recent TCCC guidelines update in June 2014, HES is retained as the preferred option over crystalloids because smaller volumes are required during long evacuations in the military setting.5

In June 2013, the U.S. Food and Drug Administration (FDA) issued a Black Box Warning on the use of HES based on reports of increased rates of mortality and acute kidney injury (AKI) after large volume resuscitation in septic, critically ill medical patients.6–8 For this reason, we evaluated the long-term outcomes associated with HES use in trauma patients. The data showed that HES independently increases risks of AKI and death after blunt, but not penetrating, trauma.9 Major limitations of that study included the relatively small population of penetrating trauma and the lack of propensity matching. The purpose of this study was to fill those gaps. We hypothesize that victims of penetrating trauma have no increased risks of AKI and/or death when receiving a single bolus of HES during initial fluid resuscitation.

METHODS

The University of Miami and Jackson Memorial Hospital institutional review board approved this retrospective study with waiver of informed consent. The data registry of the Ryder Trauma Center was queried for all patients sustaining penetrating trauma between December 1, 2009 and December 31, 2012. Patients admitted to the floor, intensive care unit (ICU) or operating room (OR) were included. Patients admitted to orthopedics or neurosurgery, under the age of 18, with known kidney disease, pregnant, incarcerated, with isolated traumatic brain injury (TBI), and/or died within 24 hours of arrival were excluded.

Demographics, injury severity score (ISS), admission vital signs and laboratory values, initial diagnostic imaging, and outcome variables including survival were queried. Patients who received 6% HES 450/0.7 in a lactated electrolyte injection (Hextend, http://www.hospira.com/) and/or blood products within 24 hours of arrival to the hospital were identified. The variables included patient age, initial systolic blood pressure (SBP), heart rate (HR), hematocrit, ISS, Glasgow Coma Scale (GCS), requirement of blood transfusion during initial resuscitation, need for operative intervention within 24 hours of arrival or direct ICU admission, the use of HES, and 90-day mortality.

AKI was defined as a rise in creatinine >1.5 times above baseline for at least 48 consecutive hours, according to RIFLE (Risk, Injury, Failure, Loss, and End-stage) and AKIN (Acute Kidney Injury Network) criteria.10,11 Propensity scores were assigned for each patient based on a logistic regression model for predicting the use of HES using patient characteristics (i.e., age and sex) and injury characteristics (i.e., ISS, initial SBP, HR, GCS, base excess, hematocrit, need for blood transfusion, ICU admission, and operation). A 1:1 fixed ratio nearest neighbor matching was performed to compare those who received HES to those who did not.

Statistical analyses were performed using SPSS version 21 (IBM Corporation, Armonk, New York). Data were reported
as mean ± SD if parametric (normally distributed) or median (interquartile range) if nonparametric (abnormally distributed). Continuous variables were compared with Student t test for parametric data and Mann–Whitney U test for nonparametric data. Categorical data were compared with Fisher’s exact test. Significance was assessed at \( p < 0.05 \).

**RESULTS**

Table I describes a population of 816 consecutive penetrating trauma admissions characterized as 88% male, age 35 ± 14 years, ISS of 10 ± 10, with a 3.8% rate of AKI, and 3.2% rate of mortality. HES was administered to 121 (14.8%) patients. Of those patients who received HES, the dose of HES was 500 (500-500) mL.

Table II compares those who received HES to those who did not. HES was generally administered to sicker patients, as reflected by lower initial SBP, lower hematocrit, higher ISS, lower GCS, and more blood transfusions (all \( p < 0.001 \)). Between the HES and no HES groups, the rate of AKI was 5.0% vs. 3.6% (\( p = 0.443 \)) and the mortality rate was 5.0% vs. 2.9% (\( p = 0.257 \)). Although there is a trend of increased mortality and AKI in the HES population, it is not statistically significant; therefore, we cannot reject the null hypothesis. Thus, HES did not increase the risk of mortality or AKI despite having a worse injury and hemodynamic compromise. Therefore, because of these differences between the HES and non-HES populations, a propensity score match was performed to more adequately assess the effects of HES alone.

Table III compares the propensity score matched groups. The matching achieved similar age, ISS, hemodynamics, rates of transfusion, ICU admission, and operative intervention between the groups. When comparing the HES and no HES matched groups, the rate of AKI was 3.8% vs. 4.8% (\( p = 0.749 \)) and the mortality rate was 3.8% vs. 4.8% (\( p = 0.749 \)). The major findings from this study are that there is no increased risk of either mortality or AKI in penetrating trauma patients who received a bolus of HES during initial fluid resuscitation. In fact, there was a trend toward a benefit; AKI and mortality were both absolutely reduced about 1% in the HES group. Using a power analysis, between 2,100 and 2,300 patients matched to a similarly sized cohort would be required to obtain statistical significance for the difference in AKI and mortality risk observed in Table III.

We previously reviewed 1,714 consecutive trauma admissions and observed that overall mortality in patients who received a HES bolus was significantly lower compared to those who received only standard of care crystalloids. In a follow-up study, in those requiring an operation after disposition from the ICU, we observed a similar trend of decreased mortality in those that received HES (4.8% vs. 3.6% \( p = 0.749 \)).

**DISCUSSION**

This is the fourth in a series of studies from our group evaluating the safety and/or efficacy of HES in trauma patients.

| **TABLE I.** Study Population (n = 816) |
|-----------------|-----------------|-----------------|
| Age (Years)     | 35 ± 14         | 36 ± 14         |
| Sex             | Male            | Female          |
| Race            | Latin           | White           |
| ISS             | 10 ± 10         | 14 ± 11         |
| Disposition     | Floor           | ICU             |
| Transfusion     | 16%             | 14.8%           |
| HES             | 3.2%            | AKI             |
| Mortality       | 3.2%            | 3.2%            |

ICU, Intensive Care Unit; OR, Operating Room; HES, hydroxyethyl starch; AKI, acute kidney injury.

| **TABLE II.** Initial Physiologic and Laboratory Values of HES and No HES Groups (n = 816) |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Value           | No HES          | HES             | p Value         |
| Age (Years)     | 35 ± 14         | 36 ± 14         | 0.499           |
| SBP (mm Hg)     | 131 ± 31        | 104 ± 27        | <0.001          |
| HR (Beats/min)  | 81 ± 37         | 85 ± 49         | 0.301           |
| Hematocrit (%)  | 39 ± 6          | 36 ± 5          | <0.001          |
| GCS             | 14 ± 2          | 13 ± 3          | 0.004           |
| ISS             | 9 ± 9           | 14 ± 11         | <0.001          |
| Disposition     | Floor           | ICU             |
| Transfusion     | 12%             | 39%             | <0.001          |
| AKI             | 3.6%            | 5.0%            | 0.443           |
| Hemodialysis    | 1.7%            | 3.6%            | 0.286           |
| Mortality       | 2.9%            | 5.0%            | 0.257           |

HES, hydroxyethyl starch; SBP, systolic blood pressure; HR, heart rate; GCS, Glasgow Coma Scale; ISS, Injury Severity Score; ICU, Intensive Care Unit; OR, Operating Room; AKI, acute kidney injury.

| **TABLE III.** Initial Physiologic and Laboratory Values of Propensity Score Matched HES and No HES Groups (n = 214) |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Value           | No HES          | HES             | p Value         |
| Age (Years)     | 35 ± 14         | 37 ± 15         | 0.425           |
| SBP (mm Hg)     | 105 ± 32        | 104 ± 28        | 0.920           |
| HR (Beats/min)  | 87 ± 37         | 83 ± 49         | 0.518           |
| Hematocrit (%)  | 36 ± 7          | 36 ± 5          | 0.780           |
| GCS             | 14 ± 2          | 14 ± 2          | 0.616           |
| ISS             | 14 ± 11         | 13 ± 10         | 0.953           |
| Disposition     | Floor           | ICU             |
| Transfusion     | 41%             | 37%             | 0.675           |
| AKI             | 4.8%            | 3.8%            | 0.749           |
| Hemodialysis    | 3.4%            | 2.1%            | 0.685           |
| Mortality       | 4.8%            | 3.8%            | 0.749           |

SBP, systolic blood pressure; HR, heart rate; GCS, Glasgow Coma Scale; ISS, Injury Severity Score; ICU, Intensive Care Unit; OR, Operating Room; AKI, acute kidney injury.
penetrating injury, HES significantly reduced transfusion requirements, improved 24-hour urine output and fluid balance, yet had no detectable effect after blunt trauma. In both of those studies, deaths within 24 hours were included.

Recently, the U.S. FDA convened a Public Workshop to discuss the risks and benefits of HES solutions. Panelists presented data from numerous studies that showed increased mortality and/or renal injury when HES was used in certain critically ill adult patients. In June 2013, a Black Box Warning was issued to include the risk of mortality and severe renal injury.

Despite the warnings, the current TCCC guidelines recommend HES when blood products are not available. In response to the FDA warnings, we observed that following blunt trauma, HES independently increases risks of AKI and death, whereas after penetrating trauma, there was no evidence of risk. Those findings revealed a fundamental difference between blunt and penetrating injury, but further study was necessary. One major limitation of that study was that the population of penetrating patients was less than half of that of the blunt population (451 penetrating vs. 956 blunt) with only 64 penetrating patients receiving HES. Another limitation was that propensity matching was not used.

In this present study, we extended our analysis on penetrating trauma to 816 patients who survived at least 24 hours; 121 received HES and this sample size was similar to that of the blunt population of our previous study. Propensity score matching was not necessary to show there was no risk associated with HES following penetrating injury. Still, this population size allowed us to match over 100 patients to HES and no HES groups and we again observed no difference in long-term outcomes between the groups. In fact, trends favored improved rates of AKI and mortality in those who received HES when compared to a matched cohort.

To note, there were equal numbers of patients who developed AKI or died in the propensity score matched analysis; 4.8% or 3.8% for the no HES or HES groups, respectively. To clarify, these were not always the same patients. In the no HES group, 4 patients developed AKI, 4 patients died, and 1 patient both developed AKI and died. In the HES group, 2 patients developed AKI, 2 patients died, and 2 patients both developed AKI and died. Thus, the identical rates of AKI and mortality seem to be a coincidence.

We are not the first to demonstrate a differential effect of HES in blunt vs. penetrating trauma. Although there is a global risk of using HES numerous studies directly addressing the trauma population have shown short-term benefit within the penetrating population alone. We have no direct explanation for the benefit of HES use in only the penetrating population.

In a recent analysis of battlefield fatalities, explosive injury caused the majority (72%) of deaths, yet many were deemed “nonsurvivable.” In contrast, the mortality in the “potentially survivable” group was predominantly due to hemorrhage from firearm-related injuries (80%). The proportion of preventable deaths from penetrating trauma in the military is profoundly higher (~80%) than that of the civilian population (~15%). For this reason, the differential effect of HES following penetrating or blunt injury necessitates clarification for its justified use in the military setting. As most treatable injuries in military are from penetrating mechanisms, and since there is evidence of early resuscitative advantages of HES in this population without the adverse long-term effects as described in our previous report in conjunction with this present analysis, a single low-dose bolus of HES during initial trauma resuscitation appears a safe alternative in the military setting as detailed by the recent TCCC guidelines.

There are study limitations. This is a retrospective review, and therefore major baseline differences between groups could not be directly controlled. All deaths within 24 hours were excluded. Resuscitative strategies were based upon physician discretion instead of a standard resuscitation protocol. Given the size of the study sample and that many physicians likely follow similar Advanced Trauma Life Support guidelines, these differences are likely minimal. Unlike the military, the time between injury and definitive care was usually <1 hour. Long-term follow-up is another potential weakness. Some patients were lost to follow-up. We may have missed some who developed organ dysfunction or died outside of the hospital. Still, patients discharged are hemodynamically stable, and thus unlikely to experience ongoing organ dysfunction or death.

CONCLUSIONS

In conclusion, this is the largest analysis of HES use during initial resuscitation following penetrating trauma. The major findings are that there was no increased risk of mortality or AKI in those who received HES, and a trend toward a benefit. As most treatable military fatalities are a result of penetrating injury, and since there is no evidence of adverse long-term effects, a single low-dose bolus of HES during initial trauma resuscitation appears a safe alternative in the military setting.

ACKNOWLEDGMENTS

We would like to recognize the nursing and administrative staff at Ryder Trauma Center for their Co-operation and assistance with the patients and their families. This study supported in part by Grants no. N140610670 from the Office of Naval Research and no. 09078015 from the U.S. Army Medical Research and Materiel Command. CJA is directly responsible for all aspects of this study. He participated in the collection, analysis, and interpretation of data; drafting and revision of the manuscript, figures, and tables. JPM, JJR, and XRB participated in the experimental design, collection of data, revision of the manuscript, figures, and tables. CJS, NN, and ASL were medically responsible for the patients and participated in the review and revision of the manuscript, figures, and tables. KGP had overall responsibility for the study; including conception and experimental design; analysis and interpretation of data; drafting and revision of the manuscript, figures, and tables; statistical expertise and evaluation; obtaining funding for this project; and supervision.
Is Hydroxyethyl Starch Safe in Penetrating Trauma Patients?

REFERENCES


ABSTRACT  In the current theater of operation, medical devices are often shipped and stored at ambient conditions. The effect of storage at hot and cold temperature extremes on ventilator performance is unknown. We evaluated three portable ventilators currently in use or being evaluated for use by the Department of Defense (731, Impact Instrumentation; T1, Hamilton Medical; and Revel, CareFusion) at temperature extremes in a laboratory setting. The ventilators were stored at temperatures of 60°C and −35°C for 24 hours and were allowed to acclimate to room temperature for 30 minutes before evaluation. The T1 required an extra 15 to 30 minutes of acclimation to room temperature before the ventilator would deliver breaths. All delivered tidal volumes at room temperature and after storage at temperature extremes were less than the ±10% American Society for Testing and Materials standard with the Revel. Delivered tidal volumes at the pediatric settings were less than the ±10% threshold after storage at both temperatures and at room temperature with the 731. Storage at extreme temperature affected the performance of the portable ventilators tested. This study showed that portable ventilators may need an hour or more of acclimation time at room temperature after storage at temperature extremes to operate as intended.

INTRODUCTION

In the current theater of military operations and in disaster situations, medical devices are often shipped and stored at room temperatures. These devices may be operated with little time to acclimate to room temperature. The effect of storage at temperature extremes on ventilator performance is unknown and to our knowledge no evaluations of these effects in modern devices have been published. Safe ventilatory support of patients require ventilators deliver desired settings accurately. Consistent tidal volume (VT) is of critical importance especially in patients with acute respiratory distress syndrome (ARDS) as shown by the results of the ARDS Network study showing that using low VT (6 mL/kg of predicted body weight) improved mortality,1 Other ventilator settings such as respiratory rate and positive end expiratory pressure (PEEP), if altered may lead to acid–base imbalance and hypoxemia and/or hemodynamic instability, respectively. Following storage at temperature extremes, we evaluated the performance of three portable ventilators currently used or are being considered for use during military operations.

METHODS

We evaluated three commercially available portable ventilators: 731 (Impact Instrumentation, West Caldwell, New Jersey), T-1 (Hamilton Medical, Reno, Nevada), and Revel (Carefusion, San Diego, California) in a laboratory setting. Studies were conducted in an altitude/environmental chamber at the University of Cincinnati. The devices were stored at temperatures of 60°C and −35°C in the chamber for 24 hours and operated after placement outside the chamber at room temperature for 30 minutes. Room temperature was 21°C. We used the Department of Defense Test Method Standard (MIL-STD-810G)2 and Joint Enroute Care Equipment Test Standard3 as guidance in selecting the testing temperatures.

After storage at each temperature for 24 hours and 30-minute acclimation to room temperature, ventilators were connected to alternating current power and attached to a two-chamber test lung (Michigan Instruments, Grand Rapids, Michigan) via the manufacturer-supplied circuit and evaluated using the combinations of ventilator settings shown in Table I, using pediatric and adult lung models shown in Table II. A Fleisch pneumotachograph (Series 4700; Hans Rudolph, Shawnee, Kansas) was connected between the ventilator circuit and the test lung and the signals for airway pressure, flow, and volume were collected on a breath-to-breath basis by a research data collection system (RSS 100, Hans Rudolph) and recorded to a personal computer for later analysis. The pneumotachograph was calibrated before each set of measurements using a 3-L super syringe. We recorded the time from powering on the ventilator until satisfactory operation. After a 1 minute stabilization period, a minimum of 1 minute of data were collected at each combination of lung model and ventilator settings. All tests were performed at each temperature a minimum of two times. Delivered and set
VTs were compared using the American Society for Testing and Materials (ASTM)\(^4\) standard of ±10% of set VT.

A paired \( t \) test was performed with each at each combination of settings and lung models to determine if there were statistical differences in delivered VTs when comparing the room temperature measurements to the same measurements after storage at \(-35°C\) and \(60°C\). Statistical significance was \( p < 0.05\). All the ventilators tested have the option to operate solely on battery, but we did not assess the performance using the battery as the power source.

**RESULTS**

Operation and/or performance was affected by storage at temperature extremes with all the ventilators tested. The T1 required an additional 15 to 30 minutes (45–60 minutes total) of time to acclimate to room temperature before the ventilator would begin delivering breaths. Additionally, at both the high and low temperatures, the T1 pressure limited at pressures less than 50 cm H\(_2\)O with the 750 mL settings despite the high pressure limit setting of 70 cm H\(_2\)O, therefore failing to deliver the full VT. At all other conditions, the T1 delivered VTs within the 10% threshold throughout the range of settings. Differences in delivered VT from room temperature measurements were statistically significant at all settings except the 100 mL and 250 mL VT settings after storage at \(-35°C\).

After storage at temperature extremes, there were no alterations in set respiratory rate or inspiratory time. The Revel began operating immediately after 30 minutes of acclimation to room temperature following extreme temperature storage. There were no alarms or alerts displayed concerning battery temperature or ability to deliver the ventilator settings chosen. The delivered VTs in both the pediatric and adult settings were all less than the ASTM ±10% range at room temperature and both temperature extremes.

Figures 1 to 5 show delivered VT as compared to the ± ASTM standard at room temperature and both temperature extremes using pediatric and adult settings. Nearly half of the delivered VTs after storage at \(-35°C\) were statistically different from room temperature measurements. Delivered VTs at all pediatric settings were not statistically different after \(-35°C\) storage but were after storage at \(60°C\). All VTs after storage at \(60°C\) were statistically different from room temperature except at settings of 250 mL, 0 cm H\(_2\)O PEEP, and lung compliances of 20 and 100 mL/cm H\(_2\)O.

After storage at extreme temperatures, there were no alterations in set respiratory rate or inspiratory time. The

<table>
<thead>
<tr>
<th>TABLE I.</th>
<th>Pediatric and Adult Ventilator Settings Used in the Evaluation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ventilator Settings</td>
<td>Lung Model</td>
</tr>
<tr>
<td>---------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Pediatric</td>
<td>30</td>
</tr>
<tr>
<td>Adult</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE II.</th>
<th>Pediatric and Adult Lung Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung Model</td>
<td>Lung Compliance</td>
</tr>
<tr>
<td>Pediatric</td>
<td>0.01 L/cm H(_2)O</td>
</tr>
<tr>
<td>Adult Normal</td>
<td>0.1 L/cm H(_2)O</td>
</tr>
<tr>
<td>Adult Restrictive</td>
<td>0.02 L/cm H(_2)O</td>
</tr>
</tbody>
</table>

FIGURE 1. Measured VT at room temperature and temperature extremes using set VT of 50 mL, 0 PEEP, lung compliance of 10 mL/cm H\(_2\)O, and airway resistance of 20 cm H\(_2\)O/L/s.

FIGURE 2. Measured VT at room temperature and temperature extremes using set VT of 100 mL, 0 PEEP, lung compliance of 10 mL/cm H\(_2\)O, and airway resistance of 20 cm H\(_2\)O/L/s.
addition of PEEP, using fraction of inspired oxygen of 1.0, or varying lung compliance did not markedly alter delivered VTs in any of the settings, with any of the ventilators tested. The digital displays with each of the devices continued to operate appropriately during testing at all conditions. There was a tendency with all the devices for VTs to decrease as compared to room temperature VTs, after storage at 60°C, although this did not result in the VTs being outside the ±10% range at room temperature. Because of the number of measurements taken at each setting and the small variance between each measurement, many of the differences were statistically significant but not necessarily clinically significant. Clinical significance was determined by applying the ASTM ±10% standard as compared to the set VT.

DISCUSSION
The main findings of the study demonstrate that following acclimation to room temperature all three ventilators accurately delivered set VTs. We also demonstrated that one of the ventilators required 45 minutes of stabilization before it would operate. While many of the differences in set and delivered VTs were statistically significant, the clinical importance of these differences is negligible.

Devices that are designed for use under optimal conditions are often used in austere environments. During military conflicts, medical devices may be subject to storage temperatures that are outside the manufacturers’ specifications. Therefore, the ability of the devices to perform under these conditions is unknown. High temperature on the Air Force flight-line can reach 140°F (60°C) during the summer and as low as −51°F (−46°C) during winter. Laboratories utilized by the branches of the military such as the Air Force Research Laboratory and U.S. Army Aeromedical Research Laboratory provide environmental testing covering a broad spectrum of harsh conditions in which a device may be deployed. This study provided clinical validation by evaluating and measuring the delivered parameters as compared to set parameters over a range of clinically relevant settings and lung conditions. Life-sustaining medical devices, such as the ventilators included in this study, must accurately provide the parameters set by the caregiver to ensure patient safety.

We conducted a comprehensive literature search and found four documents pertaining to testing of ventilation devices. Bruckart et al evaluated medical devices, including two portable ventilators at temperature extremes, based on MIL-STD-810D. The authors found that neither of the ventilators tested failed after storage at temperature extremes (63°C and −46°C). The details of the testing procedure, ventilator settings, or ventilator type or model were not detailed in the publication. Barnes and Stockwell evaluated the performance of ten manual resuscitators at −18°C and 50°C. The results of the study showed that only two of the resuscitators met ASTM and International Organization for Standardization performance standards at both temperature extremes. The remaining two publications evaluated a total of five portable ventilators (Motivus, Pneupac 2, Logic 07, Oxylog, and Maxaman) at temperature extremes and

FIGURE 3. Measured VT at room temperature and temperature extremes using set VT of 250 mL, 0 PEEP, lung compliance of 20 mL/cm H₂O, and airway resistance of 5 cm H₂O/L/s.

FIGURE 4. Measured VT at room temperature and temperature extremes using set VT of 500 mL, 0 PEEP, lung compliance of 20 mL/cm H₂O, and airway resistance of 5 cm H₂O/L/s.

FIGURE 5. Measured VT at room temperature and temperature extremes using set VT of 750 mL, 0 PEEP, lung compliance of 20 mL/cm H₂O, and airway resistance of 5 cm H₂O/L/s.
found that although all the devices functioned well at high temperatures, only two would operate at temperatures below 0°C. The devices evaluated in these two studies are no longer commercially available.

The results of our study show that ventilator storage at extremely cold or hot temperatures can be problematic for the newest generation of portable ventilators. Recommended storage temperatures reported in the operator’s manual of the devices were −15°C to 70°C with the T1 (Hamilton T1 Operator’s Manual: Hamilton Medical Version 624369/02); −25°C to 49°C with the 731 (Operation manual, 731 Series Electrical Mini Ventilator (EMV+), Impact Instrumentation Revision 6.01, http://www.impactinstrumentation.com/731.html); and −20°C to 60°C with the Revel. Additionally, the Revel operator’s manual states the device will operate at least 20 minutes when returned to room temperature for 10 minutes, after storage at temperatures of −30°C to 70°C. Storage at extremely cold temperatures prevented the T1 from immediately starting up after the 30-minute warm-up period and altered performance as demonstrated by false high pressure readings and alarms with the 750 mL V_T settings. Even though we did not operate the ventilators on direct current power, the high and low battery temperature alerts on the 731 warned that the temperature conditions were outside of the specifications. There were no such alerts with the T1 and Revel, but the T1 “Technical Fault” alarms on start-up after storage at −35°C were followed by codes but the cause was not specified on the device or in the operator’s manual. Despite the start-up and battery problems associated with storage at extreme temperatures outside the ventilators’ published temperature storage ranges, the delivered V_Ts were not markedly different from those recorded at room temperature. Although the majority of the V_T differences between room temperature and temperature extremes were statistically significant, with the exception of the Revel and those detailed in the Results section with the T1 and 731, delivered V_Ts were within 10% of the set V_Ts across a range of settings.

CONCLUSIONS

Storage of medical devices at extremes in temperature is often necessary in the theater of operations in austere environments. After storage, devices must be able to acclimate to room temperature within a reasonable time frame and operate accurately as intended. Caregivers must have confidence that the device settings are reliably delivered to the patient. This study showed that portable ventilators may need an hour or more of acclimation time at room temperature after storage at temperature extremes to operate as intended. Caregivers must be aware of the performance limitations when portable ventilators cannot be stored at optimal temperatures. To define optimal storage temperatures and duration of storage, future studies should investigate a range of temperatures and storage durations and evaluate the effect on portable ventilator performance.

ACKNOWLEDGMENT

This study was funded by the U.S. Air Force.

REFERENCES

Evaluation of Oxygen Concentrators and Chemical Oxygen Generators at Altitude and Temperature Extremes

Thomas C. Blakeman, MSc*; Dario Rodríguez Jr., MSc†; TSgt Tyler J. Britton, USAF‡; Col Jay A. Johannigman, USAF MC*; Lt Col Michael C. Petro, USAF MC†; Richard D. Branson, MSc*

ABSTRACT Oxygen cylinders are heavy and present a number of hazards, and liquid oxygen is too heavy and cumbersome to be used in far forward environments. Portable oxygen concentrators (POCs) and chemical oxygen generators (COGs) have been proposed as a solution. We evaluated 3 commercially available POCs and 3 COGs in a laboratory setting. Altitude testing was done at sea level and 8,000, 16,000, and 22,000 ft. Temperature extreme testing was performed after storing devices at 60°C and −35°C for 24 hours. Mean FIO2 decreased after storage at −35°C with Eclipse and iGo POCs and also at the higher volumes after storage at 60°C with the Eclipse. The iGo ceased to operate at 16,000 ft, but the Eclipse and Saros were unaffected by altitude. Oxygen flow, duration of operation, and total oxygen volume varied between COGs and within the same device type. Output decreased after storage at −35°C, but increased at each altitude as compared to sea level. This study showed significant differences in the performance of POCs and COGs after storage at temperature extremes and with the COGs at altitude. Clinicians must understand the performance characteristics of devices in all potential environments.

INTRODUCTION
Supplemental oxygen can be lifesaving in emergency situations, although the burden of providing oxygen during transport and in remote areas is substantial in cost, transport, and materials. Oxygen cylinders are heavy and present a number of potential hazards including fire and projectile risks. Liquid oxygen systems provide a large amount of gas with a smaller footprint but are heavy, exhaust gas over time, and present a burn risk if handled improperly. In addition, the output of both of these oxygen systems is finite and requires refilling, which presents logistical issues in far forward military operations. Simpler, lighter, and longer lasting oxygen delivery systems are needed for military and mass casualty operations. As possible materiel solutions, we evaluated portable oxygen concentrators (POCs) and chemical oxygen generators (COGs) at altitude and temperature extremes. Understanding performance of these devices under deployed conditions is crucial to safe and effective use.

METHODS
We evaluated 3 commercially available POCs (Eclipse 3 and Saros, Chart Sequal Technologies, Ball Ground, Georgia, and iGo, DeVillbis Healthcare, Somerset, Pennsylvania) and 3 COGs (O2PAK, Pacific Precision Products, Irvine, California, Traumaid-26, HABCO Industries, Glastonbury, Connecticut, and BUDI Oxygen Bag [BOB], Green Dot Systems, Miami, Florida) in a laboratory setting. The devices were evaluated at sea level and at altitudes of 8,000, 16,000, and 22,000 ft corresponding to respective barometric pressures of 760, 565, 412, and 321 mm Hg in a man-rated altitude chamber at Wright–Patterson Air Force Base, Dayton, Ohio. An altitude of 8,000 ft was chosen to represent a simulated cabin altitude during a Critical Care Air Transport Team flight. An altitude of 22,000 ft was chosen to represent the upper limit of crew functionality in the case of aircraft decompression and conditions for Special Operations Forces mission requirements. The devices were also evaluated after storage for 24 hours at temperature extremes of −35°C (−31°F) and 60°C (140°F) in an altitude/environmental chamber at the University of Cincinnati. The devices were allowed to acclimate to room temperature for 30 minutes after placement outside the chamber before measurements were made. Room temperature was 21°C (70°F).

The COG flow output was obtained by attaching the oxygen tubing to the device and to a Fleisch pneumoachograph (Series 4700, Hans Rudolph, Shawnee, Kansas). Measurements of liter flow, total oxygen volume, and duration of operation were measured continuously after activation of the devices until flow ceased. Oxygen concentration was continuously measured with a fast laser diode oxygen analyzer (O2CAP, Oxigraf, Inc., Mountain View, California) throughout the duration of operation. Surface temperature of the COGs was measured intermittently throughout the duration of operation with a noncontact infrared thermometer (62 Max, Fluke Corporation, Everett, Washington).

Measurements of flow, volume, and fraction of inspired oxygen (FIO2) were accomplished by attaching oxygen tubing to the outlet of the POCs and to the inlet of an oxygen concentrator tester (Hans Rudolph) and running the device.
in either continuous flow or bolus mode. The concentrator tester has the ability to provide negative pressure to simulate inspiratory effort, which triggers the concentrator to deliver a predetermined bolus of oxygen. Concentrators were tested at 1, 2, and 3 L/min continuous flow and throughout the range of bolus volumes with each device at respiratory rates of 20 and 30 breaths/min, with each bolus setting. Data were recorded every 100 milliseconds with continuous flow mode and breath to breath in bolus mode. Concentrators were allowed 1 minute of stabilization and a minimum of 1 minute of data were collected at each setting.

Flow and volume accuracy was determined by comparing the measured values to the device specifications detailed in the operator’s manual of the Eclipse and Saros. The iGo operator’s manual did not report an accuracy range for flow and bolus volume, so we used the ranges documented for the other two devices. Reported flow accuracy was ±10% or 200 mL/min, whichever was larger. Bolus volume accuracy was reported as ±15% of the set volume. $\text{FIO}_2$ accuracy range was determined by the documented range in the operator’s manual for each device. The reported $\text{FIO}_2$ range was 90% ± 3% for the Eclipse, 91% ± 3% with the iGo, and 93% ± 3% with the Saros. In addition, battery life of the POCs was evaluated at room temperature after charging for 24 hours, using continuous flow at 3 L/min and the highest pulse dose setting at a respiratory rate of 30 bpm. Two devices of each model were evaluated and all tests with each devices was accomplished a minimum of 2 times. Data were continuously recorded to a personal computer for later analysis.

**Device Description**

**Portable Oxygen Concentrators**

Figure 1 shows the POCs evaluated in this study. These devices were chosen because each produced the highest commercially available continuous flow output and bolus size. All the devices weighed less than 20 lbs. The Eclipse 3 and iGo can either be carried via handle or placed in a wheeled cart for transport. A harness which attaches to the Saros that includes a shoulder strap provides a hands-free method in which to transport the device. Table I shows the specifications for the concentrators evaluated in this study.

**Chemical Oxygen Generators**

Current COGs typically contain 1 or more of the following solid compounds: sodium chlorate, sodium perchlorate, potassium superoxide, or peroxide species’ sodium percarbonate, or percarbamide peroxide.¹ When combined with a catalyst, the resulting chemical reaction releases oxygen and produces heat. Figure 2 shows the COGs evaluated in this study.

**O₂PAK**

The main ingredient in the O₂PAK is sodium chlorate in addition to small quantities of disodium peroxide, disodium oxide, mica, magnesium, sodium perchlorate, glass powder, and zinc peroxide. The device is cylindrical, 9.8 inches in height, and 4.0 inches in diameter, weighing 3.0 lbs.² The device is self-contained, sealed, and internally insulated, and is supplied with a nylon cover for further insulation. A pin attached to a wire is pulled to activate the device. Oxygen

---

**FIGURE 1.** Oxygen concentrators evaluated in this study.
begins to flow within seconds of activation. The O2PAK has a small flow indicator at the end of the outlet tubing where oxygen tubing is attached and also connects to a nasal cannula or simple mask for patient use. The cost of the device is $675 each.

**Traumaid**

The main ingredient in the Traumaid is sodium perchlorate with smaller quantities of iron powder, manganese dioxide, and mica. Similar to the O2PAK, the device is cylindrical, 8.2 inches in height, and 3.5 inches in diameter, weighing 2.3 lbs. The device is self-contained, sealed and internally insulated, and may be fitted with an optional nylon cover for additional insulation. The device has 2 pins that must be pulled to initiate the reaction process. Oxygen flow begins seconds after activation. As with the O2PAK, oxygen tubing is attached from the device outlet to a nasal cannula or simple mask for patient use. The cost of the device is $895 each.

**BUDI Oxygen Bag**

The BOB system requires the user to add ingredients to a plastic bag to initiate oxygen production. The reusable bag and chemicals are supplied as a kit. The user places premeasured sodium percarbonate and manganese into the bag, adds 450 mL tap water, swirls the bag to mix, place the top on the bag, and attach oxygen tubing from the outlet in the top to a nasal cannula or simple mask for use. Oxygen flow begins several minutes after mixing the chemicals. The top of the device contains desiccant consisting of silica beads to absorb excess moisture as the gas exists the bag. The cost for a single device kit is $163, and consists of 1 reusable bag and chemicals for 2 separate runs. A refill kit consisting of enough dry chemicals to run 4 reactions costs $50.

**RESULTS**

**Portable Oxygen Concentrators**

The mean FIO₂ with the Eclipse was within the manufacturer stated range during all altitudes and temperatures in continuous flow mode and at all altitudes in bolus mode, but fell to <87% at bolus volumes of 128, 164, and 192 mL at both temperature extremes. Figure 3 shows the ranges and mean FIO₂ with the 192 mL bolus setting at a respiratory rate of 30 at all test conditions. The FIO₂ difference from room temperature was statistically significant (p < 0.0001). Delivered FIO₂ was higher at altitude than at sea level especially with the bolus volumes of 64 mL and greater. Using the 2 L/min continuous flow setting, the mean flow was 1.7 L/min at all 3 altitudes, which was slightly below the accuracy range of 1.8 to 2.2 L/min. All continuous flow settings were within the accuracy range at temperature extremes. In bolus mode at the 128 and 160 mL settings at a respiratory

<table>
<thead>
<tr>
<th>TABLE I.</th>
<th>Concentrator Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eclipse 3</td>
</tr>
<tr>
<td>Size (H × W × D) (in)</td>
<td>19.3 × 12.3 × 7.1</td>
</tr>
<tr>
<td>Weight With Battery (lbs)</td>
<td>17.4</td>
</tr>
<tr>
<td>Continuous Flow Settings</td>
<td>0.5–3.0 lpm (0.5 lpm Increments)</td>
</tr>
<tr>
<td>Pulse Dose Settings</td>
<td>16–96 mL, 128 mL, 16 mL, 196 mL</td>
</tr>
<tr>
<td>O₂ Concentration</td>
<td>90% ± 3%</td>
</tr>
<tr>
<td>AC/DC Operation</td>
<td>Yes</td>
</tr>
<tr>
<td>Battery Life</td>
<td>1.3–5.4 Hours</td>
</tr>
<tr>
<td>Storage Temperature (°C)</td>
<td>–20–60</td>
</tr>
<tr>
<td>Operating Temperature (°C)</td>
<td>10–40</td>
</tr>
<tr>
<td>Altitude Range (ft)</td>
<td>0–13,123</td>
</tr>
</tbody>
</table>

AC, alternate current; DC, direct current; lpm, liters per minute.

**FIGURE 2.** Chemical oxygen generators evaluated in this study.
rate of 30, and 192 mL at respiratory rates of 20 and 30, at all conditions the measured bolus volumes were less than reported accuracy range by 5% to 45%.

The mean FIO\textsubscript{2} with the Saros was within the specified range at all altitudes and temperature extremes in both continuous flow and bolus modes with the exception of 3 L/min continuous flow after storage at −35°C (88% ± 3%) and 96 mL bolus modes after storage at both temperature extremes (88% ± 4%). Figure 4 shows the ranges and mean FIO\textsubscript{2} at the 96 mL bolus setting at a respiratory rate of 30 at all test conditions. The differences from room temperature values were not statistically significant (\textit{p} > 0.05). In all continuous flow settings, the Saros liter flows were less than reported accuracy range by 0.1 to 0.2 L/min after storage at −35°C and at all altitudes using the 2 L/min setting and at 16,000 and 22,000 ft using the 1 L/min setting. Flow accuracy was within specifications at sea level and after storage at 60°C. Pulsed dose volumes were 0.1 to 0.4 mL less than stated accuracy at 16,000 and 22,000 ft using the 16 mL setting with respiratory rates of 20 and 30 bpm and after storage at −35°C using the 16 mL setting at a respiratory rate of 20 breaths/minute. All other bolus volumes were within specifications.

The iGo produced a mean FIO\textsubscript{2} that was within the specified range in both continuous flow and bolus modes at sea level and 8,000 ft and after storage at 60°C. At 16,000 ft, the FIO\textsubscript{2} fell to 81% in continuous flow mode and failed to operate in bolus mode. After storage at −35°C, the FIO\textsubscript{2} range was 73% ± 0.3% to 78% ± 9% in continuous flow mode. The difference from room temperature measurements was statistically significant (\textit{p} < 0.01). Figure 5 shows the range and mean FIO\textsubscript{2} at the 3 L/min continuous flow setting at all test conditions. In continuous flow mode at the 2 L/min setting after storage at 60°C and at the 3 L/min setting after storage at −35°C, measured flow rate was less than specifications by 0.1 and 0.5 L/min, respectively. All bolus modes at sea level, 8,000 ft, and after storage at both temperature extremes were within specified accuracy range.

Battery life varied widely between the concentrators. Table II shows the battery life of each device at the highest pulse dose setting using a respiratory rate of 30 breaths/minute and at the highest continuous flow setting.

One of each of the POCs was rendered inoperable after storage at −35°C. The Eclipse and iGo would start, but the membrane pads to make mode and flow adjustments would not respond and the Saros would not start. These devices were reevaluated after having been at room temperature for 24 hours but the problems remained and were permanent.

**Chemical Oxygen Generators**

As compared to sea level at room temperature, flow rate, duration of operation, and total volume of oxygen produced varied widely between devices and within the same devices when exposed to temperature extremes and increased altitude. The inter-device variability was greatest with the BOB at all conditions, but this device was least affected by temperature extremes. As compared to room temperature
measurements at sea level, mean liter flow and total oxygen volume increased with each increase in altitude with all COGs (Table III). Duration of operation did not markedly change with the O2PAK and Traumaid with changes in altitude and was inconsistent with the BOB. Mean oxygen concentration was 99.9%, 95% confidence interval (CI) (99.87%, 99.94%) with the O2PAK, 99.9%, 95% CI (99.89%, 99.96%) with the Traumaid, and 80.9%, 95% CI...

FIGURE 4. Range and mean FIO2 on the 96 mL bolus setting with the Sacros at a respiratory rate of 30 during all test conditions.

FIGURE 5. Range and mean FIO2 at the 3 L/min continuous flow setting with the iGo at a respiratory rate of 30 during all test conditions.
The oxygen concentration measurements started when flow began and continued until flow ceased.

As compared to room temperature measurements, after storage at −35°C mean flow rate was lower with the O2PAK. The duration of operation was longer, but the total oxygen was not markedly different. The mean flow rate and total oxygen volume was lower with the Traumaid, but the duration of operation was unchanged. The mean flow rate and total oxygen volume was slightly higher with the BOB, but the duration of operation was less.

After storage at 60°C, mean flow rate was higher with the O2PAK but the total oxygen volume and duration of operation were lower. Mean flow rate was higher with the BOB, total oxygen volume was unchanged, but duration of operation was less. Table IV shows mean flow rate, mean total oxygen volume, and mean duration of the devices at all test conditions.

The highest surface temperature with each device was 172°F with the O2PAK, 167°F with the Traumaid, and 173°F with the BOB. These temperatures occurred at or near the time oxygen generation ceased.

**DISCUSSION**

**Portable Oxygen Concentrators**

Oxygen concentrators were first employed as an alternative for compressed oxygen for use in long-term oxygen therapy (80.16%, 81.39%) with the BOB. The oxygen concentration measurements started when flow began and continued until flow ceased.

TABLE II. Mean Battery Life of Each Concentrator at the Highest Bolus and Continuous Flow Settings

<table>
<thead>
<tr>
<th>Concentrator</th>
<th>Bolus Setting (mL)</th>
<th>Battery Life (minute ± SD)</th>
<th>Continuous Flow Setting (lpm)</th>
<th>Battery Life (minute ± SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eclipse</td>
<td>192</td>
<td>75.5 ± 0.6</td>
<td>3</td>
<td>76.0 ± 0.8</td>
</tr>
<tr>
<td>Saros</td>
<td>96</td>
<td>57.0 ± 4.1</td>
<td>3</td>
<td>34.0 ± 2.2</td>
</tr>
<tr>
<td>iGo</td>
<td>84</td>
<td>184.5 ± 0.7</td>
<td>3</td>
<td>108.5 ± 2.1</td>
</tr>
</tbody>
</table>

TABLE III. Mean Liter Flow and Mean Total Oxygen With (% Increase) From Sea Level Measurements With the COGs at Each Altitude

<table>
<thead>
<tr>
<th>COG</th>
<th>Sea Level</th>
<th>8,000 ft</th>
<th>16,000 ft</th>
<th>22,000 ft</th>
</tr>
</thead>
<tbody>
<tr>
<td>O2PAK</td>
<td>Liter Flow</td>
<td>6.5</td>
<td>7.6 (17)</td>
<td>10.5 (62)</td>
</tr>
<tr>
<td>Volume</td>
<td>181.3</td>
<td>220.5 (22)</td>
<td>301.0 (66)</td>
<td>421.3 (132)</td>
</tr>
<tr>
<td>Traumaid</td>
<td>Liter Flow</td>
<td>5.6</td>
<td>5.6 (0)</td>
<td>8.8 (57)</td>
</tr>
<tr>
<td>Volume</td>
<td>139.2</td>
<td>152.9 (10)</td>
<td>222.2 (60)</td>
<td>279.6 (101)</td>
</tr>
<tr>
<td>BUDI</td>
<td>Liter Flow</td>
<td>0.8</td>
<td>1.8 (125)</td>
<td>3.1 (287)</td>
</tr>
<tr>
<td>Volume</td>
<td>23.4</td>
<td>29.9 (28)</td>
<td>55.2 (136)</td>
<td>95.0 (306)</td>
</tr>
</tbody>
</table>

**TABLE IV.** Mean (±SD) Flow Rate, Total Oxygen Volume, and Duration of Operation at Sea Level at Room Temperature, at Altitude, and After Storage at Temperature Extremes

<table>
<thead>
<tr>
<th>COG</th>
<th>Mean Flow (lpm)</th>
<th>Mean O2 Volume (L)</th>
<th>Mean Duration (Minute)</th>
</tr>
</thead>
<tbody>
<tr>
<td>O2PAK</td>
<td>Room Temperature/Sea Level</td>
<td>6.5 ± 2.6</td>
<td>181.3 ± 37.8</td>
</tr>
<tr>
<td>8,000 ft</td>
<td>7.6 ± 2.7</td>
<td>220.5 ± 6.3</td>
<td>28.9 ± 1.6</td>
</tr>
<tr>
<td>16,000 ft</td>
<td>10.5 ± 3.8</td>
<td>301.0 ± 1.4</td>
<td>28.5 ± 0.9</td>
</tr>
<tr>
<td>22,000 ft</td>
<td>14.4 ± 5.2</td>
<td>421.3 ± 19.1</td>
<td>29.1 ± 1.8</td>
</tr>
<tr>
<td>−35°C</td>
<td>5.5 ± 1.6</td>
<td>188.5 ± 10.6</td>
<td>33.0 ± 2.0</td>
</tr>
<tr>
<td>60°C</td>
<td>7.9 ± 2.9</td>
<td>179.2 ± 10.6</td>
<td>22.6 ± 0.6</td>
</tr>
<tr>
<td>Traumaid</td>
<td>Room Temperature/Sea Level</td>
<td>5.6 ± 2.8</td>
<td>139.2 ± 36.2</td>
</tr>
<tr>
<td>8,000 ft</td>
<td>5.6 ± 3.4</td>
<td>152.9 ± 24.2</td>
<td>27.2 ± 1.6</td>
</tr>
<tr>
<td>16,000 ft</td>
<td>8.8 ± 5.6</td>
<td>222.2 ± 7.0</td>
<td>26.9 ± 1.6</td>
</tr>
<tr>
<td>22,000 ft</td>
<td>9.9 ± 6.7</td>
<td>279.6 ± 32.1</td>
<td>31.2 ± 1.3</td>
</tr>
<tr>
<td>−35°C</td>
<td>3.7 ± 2.6</td>
<td>96.6 ± 30.5</td>
<td>24.0 ± 7.9</td>
</tr>
<tr>
<td>60°C</td>
<td>6.5 ± 3.5</td>
<td>122.2 ± 9.4</td>
<td>18.7 ± 0.2</td>
</tr>
<tr>
<td>BUDI</td>
<td>Room Temperature/Sea Level</td>
<td>0.8 ± 1.0</td>
<td>23.4 ± 5.6</td>
</tr>
<tr>
<td>8,000 ft</td>
<td>1.8 ± 1.5</td>
<td>29.9 ± 5.1</td>
<td>16.6 ± 3.0</td>
</tr>
<tr>
<td>16,000 ft</td>
<td>3.1 ± 2.9</td>
<td>55.2 ± 15.7</td>
<td>17.5 ± 2.1</td>
</tr>
<tr>
<td>22,000 ft</td>
<td>3.9 ± 3.1</td>
<td>95.0 ± 28.7</td>
<td>20.3 ± 2.7</td>
</tr>
<tr>
<td>−35°C</td>
<td>1.5 ± 1.0</td>
<td>36.3 ± 4.5</td>
<td>24.0 ± 3.0</td>
</tr>
<tr>
<td>60°C</td>
<td>1.7 ± 1.2</td>
<td>34.2 ± 2.3</td>
<td>19.2 ± 1.2</td>
</tr>
</tbody>
</table>

**TABLE II. Mean Battery Life of Each Concentrator at the Highest Bolus and Continuous Flow Settings**
in the home in the late 1970s. The devices were an attractive alternative because of the ability to supply unlimited oxygen, lower cost, and improved logistics compared to oxygen cylinders. Early concentrators were large and heavy, weighing as much as 143 lbs. Six of these early devices (DevilbissDeVO2, Rimer-Alco Dom 10, Mountain Medical Econo 2, Ventronics Hudson 6200, Dragerwerk Permox, and Cryogenic Associates Roomate) were evaluated at continuous flows of 1 to 4 liters by Johns et al and found that all the devices at 1 and 2 L/min produced oxygen concentrations of >90%, but began to fall at 3 L/min with one concentrator. Half of the devices produced oxygen concentrations ≥90% at the 4 L/min setting. Gould et al also conducted a study using three of the same concentrators as Johns (Mountain Medical Econo 2, De Vilbiss DeVO2, and Cryogenic Associates Roomate) in addition to Mountain Medical Mini 02 and Oxygen Enrichment Company OE-4E, producing similar results. Oxygen concentrators have also shown to be an effective and economical substitute for compressed oxygen cylinders in remote high altitude areas.

Although these early concentrators performed adequately as stationary units in the home, they were too large for ambulatory use, so smaller cylinders were used for this purpose. POCs emerged in 2000 that were smaller, lighter devices with optional battery operation capable of producing up to 3 L/min of continuous flow, making ambulation possible without switching to a cylinder. Fischer et al conducted a study in an altitude chamber with volunteers having chronic obstructive pulmonary disease using 5 commercially available Federal Aviation Administration-approved POCs (Invacare XPO2, Invacare, Elyria, Ohio; Freestyle AirSep C., Buffalo, New York; Evergo Philips Healthcare, Hamburg, Germany; Inogen One Inogen, Goleta, California; Eclipse 3, Sequal, Ball Ground, Georgia) using bolus mode or if not available, continuous flow mode, at a simulated altitude of 2,650 m (8,694 ft). The authors found that each POC was able to provide enough oxygen to the subjects to increase partial pressure of oxygen in arterial blood ≥ 10 mm Hg. POCs have also been evaluated as the oxygen source for chronic obstructive pulmonary disease patients during a 6 minute walk test and were found to be a suitable alternative to portable oxygen cylinders or liquid oxygen for ambulation. Rodriguez et al recently performed a bench study of using a POC as the primary oxygen source for a portable ventilator that could be used during transport. The study showed that the integrated system was capable of producing an FIO2 of up to 0.7 during selected settings.

Our study is the first to evaluate the performance of POCs at altitudes above normal commercial airline cabin altitude and after storage at extreme temperatures. With the exception of the Eclipse, the flow rates and bolus volumes were within or slightly less than the reported range. These differences are not clinically significant. The iGo would not operate at the 2 highest test altitudes, which were above the altitude limit stated in the owner’s manual. The Eclipse and Saros operated above the operator’s manual stated altitude limit. The POCs are designed to deliver a total volume of 3 L of oxygen/minute, whether in bolus mode or continuous flow mode. In the military setting, especially far forward, mass casualty, and austere environments, pulsed dose technology will provide a higher FIO2 and is more energy efficient than continuous flow, which would be helpful in those resource-constrained environments. When the combination of respiratory rate and set bolus volume exceeds the 3 L threshold, the Eclipse and Saros mitigate the effect by decreasing the bolus volume, whereas the iGo skips breaths to maintain an acceptable FIO2. Our study design did not go above the reported maximum respiratory rate for any bolus volume with the Saros and iGo. The maximum bolus volume for these 2 devices was 96 and 84 mL, respectively and was 192 mL with the Eclipse. To maintain an FIO2 in the specified range, at a respiratory rate of 20 breaths/min with the 192 mL bolus and a respiratory rate of 30 breaths/min with the 128, 160, and 192 mL bolus, the Eclipse decreases the bolus size. This strategy maintained the FIO2 at sea level/room temperature and all altitudes but not after storage at temperature extremes. The measured FIO2 range was 83% to 86%, but the bolus volumes were 1% to 13% larger after temperature extreme storage, which may help to explain the lower FIO2. The storage temperatures may have affected the device’s ability to effectively regulate the bolus volume and/or generate the target oxygen concentrations at the higher bolus volume settings. Although the liter flow with the iGo in continuous flow mode after storage at −35°C was within the reported range, the FIO2 was 15% to 20% lower than at room temperature, demonstrating the effect of extreme cold temperatures on oxygen generation during continuous flow mode. The POCs were tested at altitudes greater than recommended in the operator’s manual. While the iGo ceased to operate at 16,000 ft, the Eclipse and Saros operated within specified performance ranges at all study altitudes.

**Chemical Oxygen Generators**

Chemical oxygen generation is not a new concept. It is the method by which Joseph Priestly discovered oxygen during his work with mercuric oxide. Priestly published his findings in 1775. In 1902, the “Lancet” reported on Kamm’s oxygen generator invention for medical use. The device used chlorate cakes and manganese oxide and when heated by a spirit lamp produced approximately 4 cubic ft of oxygen before needing to be replenished with ingredients. More recently there has been interest in employing this technology in areas where providing oxygen in cylinders or in liquid form is logistically difficult or economically prohibitive such as during combat casualty care, disaster situations, and in extreme rural environments in undeveloped countries. To our knowledge, our study is the first to evaluate COGs at altitude and temperature extremes. Pollock and associates evaluated emOx and SysO2 COGs at sea level. These devices were similar to the BOB included in our study and
were similar in performance. No other publications of COG evaluations were found. The O2PAK and Traumaid were similar in design and function but the O2PAK produced more oxygen volume because of a higher flow rate and duration of operation at all conditions. After storage at ~35°C, the output of these 2 devices decreased, but increased after storage at 60°C as compared to room temperature. Oxygen is produced by an exothermic reaction and the temperature of the device ingredients at time of ignition and throughout operation affects the device output. The total output of the BOB was much less than the O2PAK and Traumaid due to a slower reaction. (Fig. 6) Peak flow occurred toward the end of the reaction with the BOB as compared to the beginning of the reaction with the other 2 COGs. Unlike the O2PAK and Traumaid, the BOB was unaffected by storage at temperature extremes. The device uses 2 dry granular chemicals which were stored at temperature extremes and uses tap water as a catalyst, which was not stored with the chemicals. This would be the practice during use in the field. The consistent water temperature may have allowed for a reaction that was comparable to the performance at room temperature.

Oxygen volume and flow rate increased with increases in altitude with each device. With the O2PAK and Traumaid, the atmospheric pressure impacts the rate of creation and/or expansion of the gas but without change in duration of operation. For a given mass of gas produced, a larger volume will be produced at altitude. Gas is dissolved in a liquid with the BOB and at altitude; Henry’s law may explain the increase in oxygen production and flow. Henry’s law states that the amount of gas dissolved in a liquid is directly proportional to the partial pressure of the gas above the surface of the solution. When ambient pressure decreases at altitude, the dissolved oxygen is released in greater quantity and because of the impact of altitude on gas density; a larger volume will be released.

Because of device design and use of dry chemicals to create oxygen, both the O2PAK and Traumaid can be operated in any orientation. The devices are small and easy to use, requiring the pulling of 2 pins to activate and start oxygen flow. The BOB is more time consuming to prepare for use. The device requires mixing of 2 dry chemicals with tap water and must be operated in an upright position either sitting on the ground or hanging by the handle due to the use of water as the catalyst, which would spill and/or clog the oxygen outlet. In addition, the cap filled with desiccant through which the oxygen exits the bag is heavy and during operation tends to fall over and crimp the bag, diminishing or ceasing the flow of oxygen. Modifications to positioning of the bag must be made to mitigate this problem. There are safety concerns related to the external temperatures during operation of all 3 COGs. The surface temperature of the devices reached 167 to 173°F, which could easily cause burns if positioned against a patient. Based on the total volume of oxygen produced at sea level, the cost per liter was $3.73 for the O2PAK, $6.44 with the Traumaid, and $1.73 with the BOB.

**CONCLUSIONS**

POCs and COGs have been proposed for use in far forward military operations and in disaster and mass casualty scenarios as alternatives to liquid and pressurized gaseous oxygen systems because of the logistics, weight, and explosive risks.

**FIGURE 6.** Sample run with all 3 chemical oxygen generators showing flow rates and duration of operation.
inherent in these systems. Although POCs and COGs are not meant to be a 100% solution, in an environment where there is no oxygen available, these systems may be viable alternatives. Although POCs may be used in a clinical environment because of the endless supply of oxygen produced if electrical power is available, COGs exhaust in 30 minutes or less depending on the manufacturer and design and the inability to adjust output, makes the devices unsuitable for continuous clinical care. COGs are more suitable for short missions from point if injury to Role 1 care or for temporary relief of altitude-induced hypoxemia experienced in the Special Forces environment. Because of the limitations of both of these types of devices, alternate oxygen systems such as liquid oxygen or oxygen cylinders should be available when appropriate. The austere environments in which the devices may be deployed may have an effect on performance. Storage at extremely cold temperatures had the greatest negative effect on the performance of the POCs. Allowing additional time for the devices to acclimate to room temperature before start up may improve device performance. POCs should not be operated at altitudes above that stated in the operator’s manual.

POCs are an attractive option because of their small size but the output is finite, performance is unpredictable at altitude and temperature extremes, and they may be cost prohibitive to use on a larger scale. The limited flow rate and total oxygen yield with the BOB does not supply an adequate amount of oxygen to be useful in emergency situations and the logistics of maintaining the system is cumbersome. As with the COGs, storage at extremely cold temperatures decreased the output of the O2PAK and Traumaid. All the devices tested may benefit from a longer time to acclimate to room temperature before use. Since the intended use of all these devices for military and disaster operations may require that both POCs and COGs be stored and operated in environments that are outside the manufacturer’s published thresholds, users must be aware of the limitations of each and mitigated as much as possible.

ACKNOWLEDGMENT

This study was funded by the U.S. Air Force.

REFERENCES

Recognition and Treatment of Nerve Agent Casualties: Evidence of Reduced Learner Engagement During Video-based Training

Alex Bukoski, PhD, DVM, DACVAA†; Rindi Uhlich, BS†; Johnny Tucker, MHA, MSIE†; Chris Cooper, MD†; Steve Barnes, MD, FACS†

ABSTRACT Changes in electrodermal activity (EDA) correlate with arousal and stress during stimulating experiences. We hypothesized that associations exist between short-term performance gains and changes in EDA. A total of 187 combat medics were randomly assigned to simulation (S), live tissue (L), or video (V) based training in the recognition and treatment of nerve agent casualties. Change in EDA from baseline to training was quantified for tonic and phasic responses and was categorized as positive (>+10%), no change (±10%), or negative (<–10%). Cognitive and psychomotor skills assessments were applied before and after the baseline/training period to quantify short-term performance changes. Statistically significant differences in both EDA arousal measures between training modalities (p < 0.001 with L>S>V) were observed. Notably, larger proportions of trainees experienced negative changes in tonic (67%) and phasic (21%) EDA measures in the V group when compared to the L and S groups. Regardless of training modality, negative tonic and phasic EDA responses were associated with lower psychomotor performance gains and this finding approached statistical significance (tonic: p = 0.056, phasic: p = 0.08). No significant differences were noted in pre- to post-training cognitive performance between EDA response categories. As quantified by EDA response to training, reduced arousal was associated with lower short-term psychomotor, but not cognitive, performance gains.

INTRODUCTION Recent increases in the use of chemical warfare in worldwide conflicts have emphasized the need to equip combat medics for the early recognition and treatment of nerve agent casualties. Given the logistical and fiscal challenges of sustaining combat medic readiness, educational opportunities in chemical, biological, radiological, nuclear, and explosive (CBRNE) casualty care have been increasingly provided through online training modules utilizing videos and graphics to engage the learner. Although these delivery methods are cost-effective and readily accessible, little is known regarding their educational effectiveness when compared to traditional inanimate and animate simulation-based training methodologies.

As defined in Bloom’s taxonomy, stress, emotion, and arousal are important aspects of cognitive engagement and are significant factors during the learning and retention of new information.1,2 Increased learner arousal is known to narrow learner focus, influence attention selectivity, and can substantially enhance all aspects of memory formation from initial coding and consolidation to memory retrieval.3,4 Decreased learner arousal can lead to disengagement and poor learner outcomes.5 Although the effect of stress on performance is often modelled as a bell-shaped curve, indicating the existence of an ideal stress level that maximizes performance, the validity of this concept as applied to training medical procedures has never been established. Although obtaining data that addresses the affective dimension of learning and performance may be particularly relevant to the military medic, whose domain of practice contains unique stressors, quantifying this component is hampered by the lack of a gold standard assessment methodology.

Change in surface skin conductance, i.e., electrodermal activity (EDA), reflects activity in the sympathetic nervous system, is a well-described method for quantifying emotional arousal in response to stimulating events, and has been used to measure learner engagement.7–9 Although the central and peripheral mechanisms that underlie the EDA phenomenon are complex, arousal-induced sympathetic nerve activity resulting in enhanced conductivity via the filling of sweat ducts is the essential physiological basis of the connection between EDA and stress. Significantly, the dependence of skin conductance on multiple physiological as well as environmental factors gives rise to interindividual variation in EDA recorded under resting conditions and must be taken into account when assessing the effects of interventions. However, as a noninvasive and unobtrusive modality for measuring affective response to intervention, EDA response to training may provide insight into what methods and environments best engage the learner and promote knowledge gain and retention during medical training.

While different skin conductance measurement techniques exist, EDA time series are generally composed of rapidly varying phasic activity bursts superimposed on a more slowly varying tonic activity level. Decomposition of the EDA signal into phasic and tonic activities is made possible by an order of magnitude difference in their characteristic...
timescales: phasic responses rise and fall over 1 to 5 seconds whereas the tonic conductance undergoes measurable change over 10 to 30 seconds. Given this separation, it is not surprising that most published literature using EDA employs a stimulus-response paradigm and relies on quantifying phasic responses within a predefined poststimulus window using a baseline-to-peak type analysis to gauge the magnitude of arousal. Although this methodology is well established, when discrete stimuli exposure times are not known, there is less agreement as to the best method of quantifying nonspecific EDA responses via time series analysis.

This article quantifies the affective response of the military medic to training in the recognition and treatment of nerve agent casualties (cholinergic crisis) utilizing EDA as the measure of affective response to three different training modalities: live tissue-based training using a nonhuman primate model, inanimate simulator-based training, and video-based training using a first person, high-definition recording of the live tissue training exercise. Using a pre- and post-training assessment design, short-term changes in cognitive knowledge and psychomotor skills were investigated as a function of EDA response to training. We hypothesized that statistically significant associations would exist between short-term psychomotor and cognitive performance gains, training modality, and changes in global measures of EDA response.

METHODS
This study involved the use of both human and animal subjects. All methods and procedures related to human subject use were reviewed and approved by the University of Missouri Health Sciences Institutional Review Board and the Human Research Protections Office of the U.S. Army Medical Research and Materiel Command, Office of Research Protections. Voluntary informed consent was obtained from all participants before data collection. All methods and procedures related to animal subject use were reviewed and approved by the University of Missouri Animal Care and Use Committee and the Animal Care and Use Review Office of the U.S. Army Medical Research and Materiel Command Office of Research Protections.

The experimental design utilized a pre- to post-training assessment schema to investigate the influence of 3 training modalities on short-term cognitive knowledge and psychomotor performance gains in the recognition and treatment of mild, moderate, and severe nerve agent exposed casualties with an emphasis on physical examination findings. The study was conducted as a randomized prospective trial over 10 months. Study participants included 212 voluntary active duty, guard, and reserve military medical personnel with a requirement for nerve agent casualty training from all branches of service. Only data sets of participants who completed all phases of the training and assessment exercises were included in the final analysis. If, for any reason, data sets were incomplete, those participants were excluded from analysis.

All phases of study design were guided by a multi-institutional consortium of subject matter experts in trauma and combat casualty care (Stephan C, Cooper C, Sobel A, et al: Acceleration of simulation-based training of medical skills through multidisciplinary collaborative research and implementation. Presentation at 2013 Association of Military Surgeons of the U.S. Meeting/AMSUS). The educational objectives and training curriculum defining the procedure of recognition and treatment of a nerve agent exposed casualty were developed and adapted from educational materials obtained from the Chemical Casualty Care Division (CCCD) of the U.S. Army Medical Research Institute of Chemical Defense (USAMRICD) (Hurst G, Tourinsky S, Madsen J, et al [editors.]: Field Management of Chemical Casualties Handbook, Ed 3. Aberdeen Proving Ground, MD: USAMRICD, CCCD, 2007; Nerve Academy 1.0: SIMapse Nerve Agent Laboratory. CD-ROM. Aberdeen Proving Ground, MD: USAMRICD, CCCD, 2009; Medical Management of Nerve Agent Casualties and Virtual Nerve Agent Casualty Assessment Exercise. CD-ROM. Aberdeen Proving Ground, MD: USAMRICD, CCCD, 2005).

Research Design
Educational interventions and the time ordering of their application are illustrated in Figure 1. The overall design can be...
separated into pre-training, training, and post-training periods. Following informed consent, demographic data (including age, gender, time in service, and relevant medical and training experience) were collected. The training portion of the design was standardized and consisted of a narrated didactic presentation followed by randomization of each subject into one of three training modalities: live tissue, simulation, or video. Training in each modality occurred in small groups of 5 to 8 participants consistent with the nerve agent casualty training formerly provided by the CCCD and approved animal use. Pre-training assessments to establish each participant’s baseline cognitive and psychomotor performance and post-training assessments to evaluate cognitive and psychomotor gains occurred immediately before and after the randomized training events.

The live tissue training exercise was based on the ketamine-anesthetized African green monkey model with physostigmine as the nerve agent simulant. The inanimate simulator training exercise used the Laerdal SimMan 3G (Laerdal, Wappingers Falls, New York, http://www.laerdal.com) simulator. These model choices reproduced the nerve agent casualty training formerly provided by the CCCD. The video-based training exercise used a trainer narrated first-person, high-definition video of the live tissue exercise described above. For all training modalities, participants were encouraged to ask questions and, for the live tissue and simulation modalities, participants were given the opportunity to acquire first-hand experience with the modality through direct interaction. Each training exercise was standardized via an instructor script and directed by an experienced trainer.

Psychomotor and Cognitive Assessments

The psychomotor component of recognizing and treating a nerve agent exposed casualty was decomposed into 13 readily identifiable and observable tasks, with emphasis on physical examination findings, as shown in Table I. This checklist was used for both the pre- and post-training psychomotor assessments, each of which consisted of three separate sub-assessments: one for each of the mild, moderate, and severe categories of nerve agent exposure. The Table I psychomotor checklist was applied to each of these sub-assessments, which were presented to each participant in random order. All psychomotor checklist tasks were designed for ease of observation and supported by clearly defined criteria for successful task completion that allowed for accurate and consistent assessment by trained observers. To further ensure consistency, inter-rater concordance was established among all observers before and during the study period (Hendricks M, Baez N, Honold E, et al: Alternative approach to establish inter-rater concordance. Presentation at the 2014 Medical Health System Research Symposium). Each psychomotor checklist task was scored as either complete or incomplete, and psychomotor gain was quantified as the pre- to post-training change in the total number of completed tasks.

The substrate for the pre- and post-training psychomotor assessments was the METIman simulator equipped with the Trauma/Disaster Casualty Kit (CAE Healthcare, Sarasota, Florida, http://www.caehospital.com).

The pre- and post-training cognitive assessments were constructed to assess participant knowledge related to the recognition and treatment of nerve agent exposed casualties. Each assessment consisted of 20 multiple choice questions, each with a single correct answer. Of these questions, 75% (15) were unchanged between the pre- and post-training administrations while 25% (5) were different. Short-term cognitive gain was quantified as the pre- to post-training change in the total number of correctly answered cognitive assessment items.

Affective Response

During the training phase of the research design (Fig. 1), the EDA of each consenting participant was assessed using the Affectiva Q Sensor (Affectiva, Waltham, Massachusetts, http://qsensor-support.affectiva.com/). This wristwatch-sized device was worn on the anterior aspect of the nondominant wrist and sampled EDA at 8 Hz. The Q Sensor measures exosomatic direct current EDA using two 12-mm electrodes spaced 4 mm apart. The small size of this device allowed for an unobstructed free range of motion of the hand. No skin preparation at the site of measurement was used and all measurements occurred in indoor, room temperature environments. Q Sensors were time-synched to a common clock and, at the conclusion of the training experience, raw data were downloaded for analysis in Matlab (MathWorks, Natick, Massachusetts, www.mathworks.com) versions R2013b and R2014a. A sample raw EDA time series is given in Figure 2 for illustrative purposes.

From the time-stamped raw data, baseline and training EDA epochs were extracted for further analysis. To compensate for pressure- and motion-induced measurement artifact,

<table>
<thead>
<tr>
<th>Task</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Assess Level of Consciousness</td>
</tr>
<tr>
<td>2</td>
<td>Rapidly Asses for Signs of Severe Exposure</td>
</tr>
<tr>
<td>3</td>
<td>Assess the Central Nervous System</td>
</tr>
<tr>
<td>4</td>
<td>Assess respiration</td>
</tr>
<tr>
<td>5</td>
<td>Assess the Eyes</td>
</tr>
<tr>
<td>6</td>
<td>Assess the Nose and Mouth</td>
</tr>
<tr>
<td>7</td>
<td>Assess the Musculoskeletal System</td>
</tr>
<tr>
<td>8</td>
<td>Assess the Gastrointestinal Tract</td>
</tr>
<tr>
<td>9</td>
<td>Assess the Cardiovascular System</td>
</tr>
<tr>
<td>10</td>
<td>Assess Sweat Glands</td>
</tr>
<tr>
<td>11</td>
<td>Assess Effectiveness of Treatment</td>
</tr>
<tr>
<td>12</td>
<td>Document Treatment on Field Medical Card</td>
</tr>
<tr>
<td>13</td>
<td>Classify Casualty into Symptomatic Category</td>
</tr>
</tbody>
</table>
which produce high-frequency noise in the EDA time series, a 1-second moving window average was applied to all data. Each participant’s baseline EDA was established during the standard didactic presentation and prior to participant knowledge of their training modality randomization. The baseline EDA point estimate was taken as the mean value of the baseline EDA time series following application of a 20-second moving window average.

Affective response to training modality was quantified using two different strategies. The first training EDA point estimate was designed to capture the more slowly varying tonic component of EDA response, which exhibits measurable change over 10 to 30 seconds, over the entire training modality exposure. Analogous to the baseline point estimate described above, the tonic training point estimate was computed as the mean value of the training EDA time series following application of a 20-second moving window average. The second training EDA point estimate was designed to characterize the more rapidly varying phasic component of EDA response, which rises and falls over 1 to 5 seconds. In this way, the phasic training point estimate was taken as the maximum observed EDA response after initial filtering of the training EDA time series. Note that this peak phasic EDA response could occur at any point during the training exercise and specific stimulus times were not recorded.

For exploration of associations between EDA response to training and pre- to post-training psychomotor and cognitive gains, the fractional change in EDA (FCEDA) response from the baseline (B) to training (T) conditions was computed as FCEDA = (T - B)/B for each of the training period tonic and peak phasic point estimates described above. Note that FCEDA, as a ratio of EDA metrics, is a dimensionless quantity expressing the magnitude and direction of each subject’s response to training relative to their baseline response. Each participant’s fractional change metrics were further classified into one of three mutually exclusive arousal categories characterized by FCEDA values that correspond to increased (>0.10), unchanged (±0.10), or decreased (<−0.10) baseline-to-training EDA responses. These arousal category assignments were used as the dependent variable in all subsequent analyses.

**Data Analysis**

The primary focus of the statistical analysis was on the interaction between the affective response to training modality and pre- to post-training psychomotor and cognitive performance changes. Analysis of variance (ANOVA) was used as the primary method for hypothesis testing and Tukey’s post hoc test was used for detection of between group differences. Statistical significance was set at $p < 0.05$. Descriptive statistics are also presented. When presented graphically, box and whiskers plots are used for data visualization: box edges give the 25th and 75th percentiles, whiskers extend to the most extreme data points, and notches represent comparison intervals for data variability about the central horizontal line representing the median.

**RESULTS**

Of 212 total participants, complete data sets were available for $N = 187$. The most common reason for data set exclusion was incomplete EDA data from the baseline period, training period, or both due to Q Sensor malfunction or excessive artifact contaminating the signal. Of the 187 participants: 47% (88) were female and 53% (99) were male; age range was 20 to 64 years (mean = 33.8 years, SD = 8.7 years); and 94% (175) claimed at least 1 to 4 years of service whereas only 6% (12) claimed less than 1 year of service. Regarding previous relevant experience, 63% (118) claimed some form of prior experience with any type of live tissue model, simulator, or human casualty in the recognition and treatment of nerve agent exposure while 37% (69) claimed no such prior experience. Of those participants that claimed some form of prior experience, only 12 (6%) claimed prior experience with nerve agent exposed humans. For the 187 participants with complete data sets, training modality randomization resulted in 32% (55) in the live tissue group.
36% (67) in the inanimate simulation group, and 33% (61) in the video group. Cross tabulations of the demographic data by training modality failed to reveal any inadequate representations of participant groupings.

**Psychomotor and Cognitive Assessments**

Statistically significant differences (ANOVA; \( p < 0.001 \)) were observed in the pre- to post-training change in psychomotor assessment scores as a function of training modality. Psychomotor gains were significantly greater in the live tissue group compared to either the simulation (\( p = 0.007 \)) or video (\( p < 0.001 \)) groups. Gains in the simulation group were not statistically different than those in the video group. The mean change (±SD) in total psychomotor assessment scores from pre- to post-training were 12.9 (±6.1) for the live tissue group, 9.5 (±6.6) for the simulation group, and 7.6 (±7.3) for the video group. In contrast to these findings, no statistically significant differences were detected between training modalities for pre- to post-training cognitive performance gains. Regardless of training modality, there was an approximately 35% increase in cognitive assessment scores from baseline.

**Affective Response**

For each of the three training modalities, tonic and peak phasic FC\(_{\text{EDA}}\) metrics were computed as described in the Methods section and the results were illustrated in Figure 3 using box and whiskers plots to summarize the data. As illustrated in Figure 3A, statistically significant differences in both tonic and peak phasic EDA responses from baseline to training were identified (ANOVA; \( p < 0.001 \) for both metrics). Tonic and peak phasic FC\(_{\text{EDA}}\) were significantly greater in the live tissue group compared to either the simulation (\( p < 0.001 \)) or video (\( p < 0.001 \)) groups. No statistically significant difference in affective response was noted between the simulation and video groups.

Table II gives the results of stratification of each participant’s tonic and peak phasic FC\(_{\text{EDA}}\) metrics into one of the three mutually exclusive FC\(_{\text{EDA}}\) categories defined as decrease (<=-0.10), no change (±0.10), and increase (>0.10). Note that no further distinction between the relative magnitudes of the FC\(_{\text{EDA}}\) metrics is made such that, e.g., all participants with FC\(_{\text{EDA}}\) > 0.10 are treated simply as having displayed a positive change in EDA response from baseline to training. From Table II, we observe that for tonic FC\(_{\text{EDA}}\), the majority of participants with increased arousal were in the live tissue group whereas the majority with decreased arousal were in the video group. Although less clearly delineated, similar trends were seen for the peak phasic FC\(_{\text{EDA}}\) metric. Looking across Table II, the vast majority of participants trained with live tissue experienced increased arousal regardless of the FC\(_{\text{EDA}}\) metric used: 54 of 59 for tonic and 59 of 59 for peak phasic. A similar trend was observed for participants trained with inanimate simulation. For those trained with video, 41 of 61 showed decreased arousal as quantified by the tonic metric compared to 13 of 61 using the peak phasic metric.

Mean pre- to post-training psychomotor assessment gains as a function of arousal category for each FC\(_{\text{EDA}}\) metric are presented in Table III. Although statistically significant differences in psychomotor gains between arousal categories were not observed for either FC\(_{\text{EDA}}\) metric, results approached statistical significance: ANOVA gave \( p = 0.056 \) for the tonic and \( p = 0.08 \) for the peak phasic FC\(_{\text{EDA}}\) metrics. Post hoc analysis using Tukey’s test revealed the source of this near significance to be differences between the decrease and increase arousal categories. In other words, regardless of training group assignment, participants with FC\(_{\text{EDA}}\) < −0.10 tended to have lower pre- to post-training...
TABLE II. Arousal Category Stratification of Participant FC\textsubscript{EDA} for the Tonic and Peak Phasic Metrics

<table>
<thead>
<tr>
<th>Training Modality</th>
<th>N</th>
<th>Tonic FC\textsubscript{EDA}</th>
<th>Peak Phasic FC\textsubscript{EDA}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Decrease No Change Increase</td>
<td>Decrease No Change Increase</td>
</tr>
<tr>
<td>Live Tissue</td>
<td>59</td>
<td>2 3 54</td>
<td>0 0 59</td>
</tr>
<tr>
<td>Simulation</td>
<td>67</td>
<td>13 16 38</td>
<td>3 4 60</td>
</tr>
<tr>
<td>Video</td>
<td>61</td>
<td>41 7 13</td>
<td>13 8 40</td>
</tr>
<tr>
<td>Combined</td>
<td>187</td>
<td>56 26 105</td>
<td>16 12 159</td>
</tr>
</tbody>
</table>

Mutually Exclusive Arousal Categories Defined as Decrease (\(<=0.10\)), No Change (\(\pm0.10\)), and Increase (\(>0.10\)). All Entries Refer to Numbers of Participants.

DISCUSSION

As reviewed by Lineberry et al,\textsuperscript{11} there is a relatively small body of published research addressing the comparative effectiveness of different simulation modalities in medical education. In addition, these studies tend to suffer from various methodological constraints, which may limit the validity of inferences based on their results. To this end, one of the specific recommendations made by Lineberry et al is to “...use criterion performance test beds that correspond highly to the conditions and demands, that trainees will face on the job.”

Given that environments in which medical skills are applied often contain significant stressors, this guideline suggests that the influence of the affective dimension of learning should be taken into consideration when evaluating medical skills training and performance assessment methodologies. As a first step toward addressing this issue, we have evaluated EDA response in a group of military medics exposed to one of three different training modalities—live tissue, inanimate simulation, or video-based instruction—aimed at teaching the recognition and initial treatment of cholinergic crisis induced by nerve agent exposure. Using a prospective and randomized pre- to post-training research design, we investigated associations between two EDA response metrics and changes in cognitive and psychomotor performance assessments. Our primary observations include: (1) live tissue training produced statistically significant increases in both tonic and peak phasic FC\textsubscript{EDA} metrics when compared to inanimate simulation or video training, (2) live tissue training produced statistically significant increases in psychomotor gains when compared to inanimate simulation or video training, (3) an association between arousal category and psychomotor gain that approached, but failed to meet the criterion for, statistical significance, and (4) no statistically significant associations between cognitive gains and either training modality or arousal category.

We note that the majority of participants with reduced arousal were from the video group and that video group participants displayed the lowest psychomotor gains.

Although engagement, arousal, and stress are significant aspects of learning psychomotor skills and gaining cognitive understanding; to the best of the authors’ knowledge there are no publications that directly address the influence of these factors, as quantified by EDA response to training, on the effectiveness of medical education. With the development of unobtrusive and wearable devices capable of continuously monitoring EDA, the ability to quantify arousal during hands-on training and performance is possible and provides the opportunity to quantitatively investigate the relationship between arousal and intervention outcome.

TABLE III. Mean (\(\pm\) SD) Pre to Post-training Psychomotor Gains for Each of the Tonic and Peak Phasic FC\textsubscript{EDA} Metrics According to Arousal Category

<table>
<thead>
<tr>
<th>Arousal Category</th>
<th>N</th>
<th>Tonic FC\textsubscript{EDA}</th>
<th>Peak Phasic FC\textsubscript{EDA}</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean Psychomotor Gain</td>
<td>Mean Psychomotor Gain</td>
</tr>
<tr>
<td>Increase</td>
<td>105</td>
<td>11.1 ((\pm)6.3)</td>
<td>159</td>
</tr>
<tr>
<td>No Change</td>
<td>26</td>
<td>9.9 ((\pm)7.9)</td>
<td>12</td>
</tr>
<tr>
<td>Decrease</td>
<td>56</td>
<td>8.3 ((\pm)7.6)</td>
<td>16</td>
</tr>
</tbody>
</table>
collapse EDA time series into unique point estimates that are compared to a baseline point estimate to account for individual variation in EDA under resting conditions. The tonic metric was designed to characterize overall drift in the underlying level of skin conductance, and the phasic metric was designed to capture maximal EDA response during training modality exposure. To augment the robustness of these fractional change metrics, we further classified each participant into one of three mutually exclusive arousal categories before linking the EDA and performance data. Although this stratification represents a relatively coarse-grained approach that may be limited by the application of arbitrary category boundaries, we find that our reported results are robust under small variation in the selected boundary values of ±0.10.

One of the primary drawbacks of the Affectiva Q Sensor in the training environments used in this study was the presence of high-frequency noise in the measured signal due to both motion and pressure artifact. Motion artifact arises when one or both electrodes lose contact with the skin surface resulting in signal dropout. Pressure artifact results from brief increases in the contact pressure between the electrodes and skin surface causing momentary spikes in measured conductance.10 We observed these artifacts in our measured EDA time series and compensated for their presence by using a moving window average to act as a rudimentary low-pass filter. Although more sophisticated filtering methods have been applied,10 we chose this method because of its ease of application and conceptual simplicity. In the future, it would be beneficial to develop data analysis algorithms that can automatically detect and correct for the appearance of these artifacts. Another limitation of EDA as a metric of affective response is the influence of environmental conditions on skin conductance. This study was performed in temperature-controlled classrooms and laboratories to mitigate this influence. Caution should be applied when attempting to use this methodology in military training exercises performed in variable environments.

Although published literature related to the study reported here is sparse, some publications are relevant to components of the results. For example, in a series of articles published between 2008 and 2014, Johnson et al. studied the influence of training modality on cognitive and performance measures using a pre- to post-training assessment design similar to that used here.12-15 The training modalities in these studies were an inanimate simulator and a CD-ROM-based presentation of clinical scenarios. Two of these publications address the management of casualties exposed to chemical agents and two address combat trauma care. Using purpose-designed cognitive and performance measures, these studies found that simulation was more effective than the CD-ROM; results that are broadly consistent with those reported here. However, these investigations did not employ live tissue-based training and also did not report any measures of affective response to training modality. In an unrelated publication, Subbarao et al.16 compared the use of high-fidelity, mannequin-based simulation to video clinical vignettes for instructing CBRNE first responder physicians, nurses, and paramedics in the recognition, triage, and resuscitation of contaminated victims. This study also utilized a pre- to post-training format but did not report any statistically significant differences between the two training modalities, did not employ a live tissue-based training modality, and did not assess the affective domain of learning.

In conclusion, for military medic nerve agent casualty training, affective response to training as quantified by tonic and peak phasic $FC_{EDA}$ metrics showed a statistically significant association with training modality. An association between arousal category and short-term psychomotor assessment gains that approached, but did not meet the criterion for, statistical significance was also observed. Because the majority of those participants who experienced reduced arousal during training were from the video group, and the majority who experienced increased arousal were from the live tissue and simulation groups, this suggests the presence of gaps in reliable learner engagement during video-based training and possibly superior learner engagement during live tissue and/or inanimate simulation-based training. Although further investigation is needed, we consider this an interesting and potentially significant finding given that current CBRNE training has a significant online-based component in both the civilian and military medical communities. We contend that advancing our understanding of arousal-based learning and the influence of training modality on learning, performance, and retention is essential for the future of combat medic skills training and readiness.

ACKNOWLEDGMENTS
Funded through USAMRMC grant no. W81XWH-10-JPC-MEDSIM-CCTC.

REFERENCES

System Design Verification for Closed Loop Control of Oxygenation With Concentrator Integration

2d Lt Matthew M. Gangidine, USAF MSC†; Thomas C. Blakeman, MSc, RRT†; Richard D. Branson, MSc, RRT†; Col Jay A. Johannigman, USAFR MC†

ABSTRACT Background: Addition of an oxygen concentrator into a control loop furthers previous work in autonomous control of oxygenation. Software integrates concentrator and ventilator function from a single control point, ensuring maximum efficiency by placing a pulse of oxygen at the beginning of the breath. We sought to verify this system. Methods: In a test lung, fraction of inspired oxygen (FiO₂) levels and additional data were monitored. Tests were run across a range of clinically relevant ventilator settings in volume control mode, for both continuous flow and pulse dose oxygenation. Results: Results showed the oxygen concentrator could maintain maximum pulse output (192 mL) up to 16 breaths per minute. Functionality was verified across ranges of tidal volumes and respiratory rates, with and without positive end-expiratory pressure, in continuous flow and pulse dose modes. For a representative test at respiratory rate 16 breaths per minute, tidal volume 550 mL, without positive end-expiratory pressure, pulse dose oxygenation delivered peak FiO₂ of 76.83 ± 1.41%, and continuous flow 47.81 ± 0.08%; pulse dose flow provided a higher FiO₂ at all tested setting combinations compared to continuous flow (p < 0.001). Conclusions: These tests verify a system that provides closed loop control of oxygenation while integrating time-coordinated pulse-doses from an oxygen concentrator. This allows the most efficient use of resources in austere environments.

INTRODUCTION Achieving adequate oxygenation is one of the primary goals of mechanical ventilation. Techniques and devices for achieving this goal—via adjustment of fraction of inspired oxygen (FiO₂) concentration, positive end-expiratory pressure (PEEP), and mean airway pressure—vary greatly.1 In adults, adequate oxygenation is typically considered an SaO₂ (arterial oxygen saturation) >90% and PaO₂ (arterial oxygen pressure) >60 mm Hg.2 However, oxygen delivery goals can be more easily monitored by the noninvasive and ubiquitous pulse oximeter, with adequate oxygenation goals having been defined as SpO₂ (peripheral oxygen saturation) of 94% ± 2%.3

In the normal hospital setting, oxygen usage to achieve these goals is typically of little concern, as the supply is virtually limitless. In far-forward military medical operations, however, oxygen becomes a limited resource to be conserved. The burdens of oxygen procurement are significant, with estimates quantifying it as up to 30% of the entire logistical footprint necessary to provide medical care during combat operations.2

In addition, recent experiences of asymmetric warfare in Operation Iraq Freedom and Operation Enduring Freedom have emphasized the need for lightweight and mobile options that are still able to provide meaningful support to the critically ill or wounded patient before, during, and after surgical intervention.4 Similar concerns over oxygen availability are applicable on the domestic front in possible incidences of disaster management that would require mass casualty care.5

A possible solution that has been explored more thoroughly in recent years is that of closed loop or autonomous oxygenation (and ventilation in general), which allows for computer control of ventilator settings in order to achieve predetermined oxygenation goals. Studies have presented a growing body of evidence that closed loop systems are more effective at both maintaining a goal oxygenation level and doing so while using less oxygen, as compared to manual clinician care, and patient outcomes have been equal or improved.2,3,5–7 Such systems allow for a more precise and gradual maintenance of SpO₂ goals, while also providing for rapid correction mechanisms in the instance of a hypoxic event.3,8 In the midst of a conflict with a characteristic injury of traumatic brain injury, this constant maintenance is particularly significant since even a single hypoxic event in patients with head injury is associated with poor outcome.9 Furthermore, such fine tuning also addresses the occurrence of hyperoxemia (usually only monitored in the neonatal population), decreasing its prevalence by avoiding clinician bias toward over-oxygenation, and reducing FiO₂ to nontoxic levels (<0.60).2,7

Portable oxygen concentrators (POC) have also come to the forefront as a means of supplying oxygen in austere settings. In the immature military theater, electricity is often the first aspect of a more established infrastructure that becomes available. With POCs running off batteries and being able to be plugged in for indefinite use, oxygen delivery is ensured while eliminating the logistic burden of cylinders or liquid oxygen.10 Air transport of critical patients has similar logistic and additional safety restraints in the use of oxygenation.
support equipment. Along with the ability to concentrate and provide oxygen in a continuous flow, POCs have also been developed that allow for the collection of concentrated oxygen in an internal reservoir and a following periodic release in the form of a pulse dose of oxygen. As early as 1990, this method of delivery was shown to be clinically effective and to utilize substantially less oxygen.\textsuperscript{11} In addition, by administering the pulse dose at the beginning of a breath cycle, one can ensure that the oxygen-rich gas enters first and travels to the sites of actual alveolar exchange, being “pushed” in by room air for the remainder of the breath, which will remain unutilized in the anatomic dead space (illustrated in Fig. 1).\textsuperscript{12} Operation in pulse dose as opposed to continuous flow mode also results in significantly less power consumption.\textsuperscript{10}

This study seeks to begin to integrate the aforementioned needs and advances into a single system that will be able to more effectively and efficiently provide for patient oxygen needs. Using the autonomous FIO\textsubscript{2}/SpO\textsubscript{2} control system developed and demonstrated by Johannigman et al\textsuperscript{3} as a basis, this new system integrates the use of an oxygen concentrator into the control loop as well (Fig. 2). The objective of this study was a proof-of-concept for the design validation of such a system, verifying successful functioning of a circuit integrating both ventilator and concentrator into a coordinated system controlled entirely by computer, providing adequate oxygenation while consuming minimal resources. It was hypothesized that in the functional system, pulsed dosed delivery of oxygen would prove more effective and efficient compared to continuous flow.

**METHODS**

The experimental setup was run entirely through a coordinating computer program on a personal computer (PC); from here, component devices were controlled and data were stored. The ventilator and oxygen concentrator system was connected to a test lung (TTL, Michigan Instruments, Grand Rapids, Michigan).

**Equipment**

All equipment used for experimentation was unmodified. The SeQual Eclipse 3 POC was used (Chart SeQual

**FIGURE 1.** A rough illustration of the oxygen distribution strategies in regular/continuous flow oxygenation versus pulse dose oxygenation. An example fraction of inspired oxygen/FIO\textsubscript{2} of 0.50 is shown in the diagram on the left. With pulse dose (right), the same amount of oxygen is used, but more of it is delivered to the part of the lungs where it is used.

**FIGURE 2.** Closed Loop Control Diagram with Concentrator Integration. FIO\textsubscript{2}, fraction of inspired oxygen; SpO\textsubscript{2}, saturation level of O\textsubscript{2} in hemoglobin; VT= tidal volume.
The Eclipse 3 was selected due to its oxygen generating capabilities, and due to the fact that ruggedized versions are available for applications in austere/military settings. The mechanical ventilator used was the Impact 731 (Impact Instruments, West Caldwell, New Jersey). The Impact 731 was also selected due to its propensity for use in austere settings, such as its employment by U.S. Air Force Critical Care Air Transport Teams (CCATT). These devices were connected to a PC and controlled externally through a program on the computer developed by Sparx Engineering (Manvel, Texas). The program ensured that these devices (in addition to other measurement devices) were all able to work together in a coordinated fashion controlled by a central entity, but it did not alter the way that each individual component functioned.

**Measurement**

Oxygen readings were collected by an O2Cap Oxygen Analyzer (Oxigraf, Mountain View, California). The sampling tubing for the instrument was positioned just before the test lung inlet in the ventilator circuit. A pneumotachometer (Hans Rudolph, Shawnee, Kansas) was also utilized before the test lung in order to record pressure and flow data. Both devices recorded data continuously, and the data collection program saved files to the PC for later analysis. Although oxygen data (in terms of FIO₂) was of primary interest, pressure and flow readings, as well as recordings of internal device settings and metrics, were also collected.

**Experimental Factors**

The experiment was designed in order to verify function across a full range of clinically-relevant ventilator settings. In particular, end-points were drawn from previous study of observed values during recent CCATT flights. Tidal volume (Vₚ) was examined at three levels: 350 mL (“min”), 550 mL (“mid”), and 750 mL (“max”). These Vₚ were paired with an appropriately inverse respiratory rate (RR): 22 breaths per minute (bmp), 16 bmp, and 10 bmp, respectively. These pairs were tested in a range of outputs for both continuous flow (3, 2, 1 Lpm) and pulse dose (192, 128, 64 mL). For continuous flow, oxygen was allowed to collect in a reservoir connected to the ventilator inlet. Additionally, tests were performed both with the absence of PEEP (0 cmH₂O), and with the presence of PEEP (10 cmH₂O). The test lung was set to a constant compliance of 0.03 L/cmH₂O. All tests were run at an inhalation:exhalation (I:E) ratio of 1:2.8. The ventilator was operated in volume control mode. When reported in pulse dose groups, Vₚ represent total Vₚ; the pulse dose volume given from the concentrator is accommodated for so that the ventilator delivers proportionately less air in order to achieve to total set Vₚ (min, mid, or max). For pulse dose mode, the burst of concentrated oxygen was administered a set amount of time before the start of each breath as defined by the ventilator. Larger doses were given a longer period of time: 1,000 ms before start of ventilator breath for 192 mL pulse, 750 ms prior for 128 mL, and 500 ms prior for 64 mL. This timing allowed a sufficient period for the pulse dose to be administered before the ventilator breath and then primarily be “pushed in” in front of it, rather than primarily mixing with the air from the ventilator. Values tested are summarized in Table I.

The system was allowed to stabilize at each new group of settings before measurements were used. Each data point represents the results from three consecutive breaths over three separate trials for each combination of settings. FIO₂ was the metric of chief interest.

A small separate set of trials was performed to measure the accumulated bolus volume delivered by the concentrator at different RRs. This allowed for quantification of which rates the concentrator was able to “keep up with” when set to deliver a bolus of 192 mL. Measurements were taken from 10 to 26 bmp, increasing by two. The system was given time to stabilize at each new setting. The data points each represent the average of three consecutive breaths during three separate runs.

**Statistical Analysis**

All data are expressed as mean ± SD. Comparisons between pulse dose and continuous flow concentrator modes at a given group of settings were done by two-tailed Student’s t test. Comparisons between multiple settings within a given group were accomplished via analysis of variance. A p value < 0.05 was considered significant.

**RESULTS**

The volume of data generated by the study precludes the comprehensive inclusion of all results. As the highest concentrator settings in both modes (3 Lpm for continuous, 192 mL for pulse) resulted in the greatest oxygen delivery, we will focus on the presentation of these results when applicable.

**TABLE I.** Experimental Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>350 mL</th>
<th>550 mL</th>
<th>750 mL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vₚ Total</td>
<td>350 mL</td>
<td>550 mL</td>
<td>750 mL</td>
</tr>
<tr>
<td>RR</td>
<td>22 bmp</td>
<td>16 bmp</td>
<td>10 bmp</td>
</tr>
<tr>
<td>PEEP</td>
<td>0 cmH₂O</td>
<td>10 cmH₂O</td>
<td></td>
</tr>
<tr>
<td>Continuous Flow</td>
<td>3 Lpm</td>
<td>2 Lpm</td>
<td>1 Lpm</td>
</tr>
<tr>
<td>Pulse Dose</td>
<td>192 mL</td>
<td>128 mL</td>
<td>64 mL</td>
</tr>
<tr>
<td>Pulse Timing</td>
<td>−1,000 ms</td>
<td>−750 ms</td>
<td>−500 ms</td>
</tr>
<tr>
<td>Compliance</td>
<td>0.03 L/cmH₂O</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I:E Ratio</td>
<td>1:2.8</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Summary of values tested for various experimental factors. Tidal volume (Vₚ) and respiratory rates (RR) are specifically paired; pulse dose and timing are specifically paired; all other factors were tested in all combinations. Bpm, breaths per minute; I:E, inhalation:expiration; PEEP, positive end-expiratory pressure.
Concentrator Rate Testing
When set to deliver 192 mL, the accumulated bolus volume output by the concentrator averaged within 5 mL of the set amount up through a rate a 16 bpm. Volume delivered at each rate was very consistent, with a standard deviation of less than 2 mL in all groups. After 16 bpm, the true volume delivered began to drop off linearly as RR increased. At a RR of 22 (rate used for the low-volume pulse-dose experimental group), true volume of concentrated oxygen delivered was 135.44 ± 1.01 mL; for a RR of 26 (maximum rate tested/clinically relevant), volume was 114.67 ± 1.37 mL. Compared to the set volume goal of 192 mL, all groups at 16 bpm and higher delivered an actual volume that was significantly less (p < 0.05 for all). At 16 bpm, the drop in actual volume delivered was about 4 mL; although this was significant statistically, it is not likely to be significant clinically. Results are illustrated in Figure 3.

System Function Verification
Although all data and settings can not be fully presented and analyzed here because of their large scope, the function or dysfunction of the system may still be reported across all settings. The system design was indeed able to operate as intended and deliver a time-coordinated FIO2 > 0.21 at all settings and combinations tested. Function was verified for RR of 10, 16, and 22 bpm; VT of 350, 550, and 750 mL; PEEP of 0 and 10 cmH2O; continuous flow of 3, 2, and 1 Lpm; and pulse dose of 192, 128, and 64 mL.

Delivered FIO2
The two different concentrator modes produced distinct patterns of oxygenation (Fig. 4). The continuous flow mode produced a much more steady-state type oxygen delivery overall. The pulse dose mode demonstrated more cyclic behavior, with periods of markedly high FIO2 immediately preceding the start of the ventilator breath, and then falling off into more distinct lows near 0.21 as the room air is administered behind the pulse. The placement of the FIO2 spike just before the start of the ventilator breath verifies that the pulse was being administered at the time it was programmed to be.

FIO2 results are highlighted here for the most clinically average ventilator settings studied: RR of 16 bpm and total VT of 550 mL. The peak FIO2 delivered in pulse dose mode was 76.83 ± 1.41% without PEEP and 70.95 ± 8.49% with PEEP. In continuous flow mode, the highest FIO2 delivered was 47.81 ± 0.08% without PEEP and 47.18 ± 0.07% with PEEP. For this setting—and all others examined—pulse flow provided decisively increased peak FIO2 values when compared to continuous flow at paired ventilatory factors (p < 0.001 in all settings).
**Closed Loop Control of Oxygenation With Concentrator Integration**

**Table II.** Oxygenation (Measured via FIO₂) Produced at Various Setting Combinations

<table>
<thead>
<tr>
<th></th>
<th>Average FIO₂ at Max Output</th>
<th>Average ± SD</th>
<th>Peak FIO₂ at Max Output</th>
<th>Average ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 PEEP</td>
<td>Min Cont 49.39 ± 0.54</td>
<td></td>
<td>0 PEEP Min Cont 49.61 ± 0.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PD 42.45 ± 2.03</td>
<td></td>
<td>PD 76.19 ± 3.20</td>
<td></td>
</tr>
<tr>
<td>Mid</td>
<td>Cont 44.95 ± 0.32</td>
<td></td>
<td>Mid Cont 47.81 ± 0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PD 34.30 ± 2.40</td>
<td></td>
<td>PD 76.83 ± 1.41</td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td>Cont 47.50 ± 0.21</td>
<td></td>
<td>Max Cont 57.20 ± 0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PD 32.07 ± 6.12</td>
<td></td>
<td>PD 76.57 ± 2.81</td>
<td></td>
</tr>
<tr>
<td>10 PEEP</td>
<td>Min Cont 48.50 ± 0.06</td>
<td>10 PEEP Min Cont 49.01 ± 0.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PD 39.49 ± 2.47</td>
<td></td>
<td>PD 72.21 ± 3.76</td>
<td></td>
</tr>
<tr>
<td>Mid</td>
<td>Cont 44.42 ± 0.13</td>
<td></td>
<td>Mid Cont 47.18 ± 0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PD 34.59 ± 3.97</td>
<td></td>
<td>PD 70.95 ± 8.49</td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td>Cont 50.77 ± 0.12</td>
<td></td>
<td>Max Cont 58.17 ± 0.13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PD 34.02 ± 1.25</td>
<td></td>
<td>PD 73.47 ± 3.32</td>
<td></td>
</tr>
</tbody>
</table>

A representive data set is shown in the table depicting the results of the maximum outputs of both oxygenation modes: 3 Lpm for continuous flow and 192 mL for pulse dose. In the table, Min = 350 mL/22 bpm, Mid = 550 mL/16 bpm, and Max = 750 mL/10 bpm with a $p < 0.001$ for continuous flow (Cont) versus pulse dose (PD) in each comparison. PEEP, positive end-expiratory pressure.

For FIO₂ over the course of the entire breath, pulse dose averaged $34.30 ± 2.04\%$ without PEEP and $34.59 ± 3.97\%$ with PEEP. Continuous flow averaged $44.95 ± 0.32\%$ without PEEP and $44.42 ± 0.13\%$ with PEEP. This difference was statistically significant for both PEEP groups ($p < 0.001$).

The highest peak FIO₂ delivered by the system was $76.83 ± 1.41\%$: this occurred with 192 mL pulse dose, no PEEP, RR 16 bpm, and $V_T$ 550 mL. The lowest peak FIO₂ delivered by the system was $31.57 ± 0.14\%$: this occurred with 1 Lpm continuous flow, no PEEP, RR 16 bpm, and $V_T$ 550 mL. The highest peak FIO₂ delivered by continuous flow was $58.17 ± 0.13\%$, occurring at a 3 Lpm flow, RR of 22, $V_T$ of 10, PEEP of 10 cmH₂O. For all groups, as concentrator output decreased, FIO₂ decreased.

These and other values for FIO₂ across various settings for the maximum output of each concentrator mode (3 Lpm flow, 192 mL pulse) are shown in Table II.

**DISCUSSION**

The study was able to successfully evaluate the oxygen provision capabilities of a novel ventilatory system. The closed loop control system was able to operate effectively across a full range of ventilator settings reflective of those encountered in the military critical care environment.

The oxygen concentrator was effectively integrated into the system, providing either sustained continuous flow or time-coordinated, computer-triggered pulse doses at the beginning of a breath cycle. In contrast to past work, which relied on positive pressure from the ventilator to initiate a pulse dose, this system’s oxygen concentrator was operated independently, and thus allowed for the use of PEEP as well, which should virtually always be present. Because this study was designed primarily to be a proof-of-concept for the system, the mere fact that the system functioned properly and produced meaningful FIO₂ results is a distinct attainment in itself. This project was designed to be able to take recent positive achievements in closed loop ventilation and oxygenation as well as with POCs and pulse dose oxygenation, and to begin to merge it all together into a comprehensive and autonomous respiratory care system. The successful operation of this ventilator/concentrator set-up was a significant milestone in achieving that goal.

The oxygen-generating capabilities of the system were found to be quite robust in both modes. This is significant for a number of reasons. First, the medical logistical burden of providing oxygen in austere locations has already been stressed. The advantages of being able to have an electric/battery-run device that can provide a patient with oxygen indefinitely are obvious; the necessary electric infrastructure to accomplish this is typically present, even in most far-forward settings. Second, it has previously been shown in research on Air Force CCATT patients that 68% of patients require an oxygen flow of less than 3 Lpm, and that an average FIO₂ of 49% corresponded to a fully healthy SpO₂ of 98%, with a majority of patients being managed in the 40 to 50% range.

Our system was either on par with or exceeding these values, suggesting that the POC represents a viable method of oxygen procurement, and is a good choice for inclusion in the closed loop system.

Pulse dose delivery of oxygen, in particular, was shown to generate markedly higher capabilities in terms of maximum FIO₂ provision, routinely providing oxygen in excess of 75%. In prior work with acute lung injury, pulse dose oxygenation has been shown to lead to significantly improved PaO₂:FIO₂ ratio when compared to continuous flow in volume control mode. Additionally, power consumption of the SeQual Eclipse POC has been previously measured, consuming an average of 151 W at a continuous flow of 3 Lpm and 103 W at a pulse dose setting of 192 mL. This means that in pulse mode, the concentrator consumes 68% as much power, while providing an FIO₂ up to 161% greater (computed at 3 Lpm flow, 192 mL pulse, middle RR and $V_T$, no PEEP); this equates to a 237% increase in efficiency of oxygen delivery by choosing pulse dose mode.
Our system allows for full effectiveness by being able to appropriately apply this superior efficiency. This is done by being able to use the developed computer program to coordinate the control of the ventilator and concentrator, ensuring that the ventilator compensates for the delivered volume from the concentrator (such that $V_T$ is not in excess of that set by the clinician), and that the concentrator pulse is timed to be automatically administered just before the start of the ventilator breath. In this way, the gas at the start of the inhalation sequence is essentially supplied by the concentrator rather than the ventilator, and the most oxygen-rich gas is what is utilized for exchange at the alveolar level. Pulse dose oxygenation allows for the utilization of the oxygen-rich gas—which can be a precious commodity in far-forward conditions—only in active respiratory space, and avoids supplying “superfluous” oxygen to the anatomic dead space where exchange does not occur (Fig. 1).

Next steps for the project include the creation of a full “lookup” table of provided FIO2 values at given settings. This information will be used to create more robust programing, which the software can draw upon in order to fulfill given oxygenation/ventilation goals (i.e., the program will have options of how to increase or decrease FIO2 in order to adjust for changes in SpO2 while simultaneously satisfying other ventilatory settings such as $V_T$ or RR). This is largely encompassed by the data generated from this study, but it could be filled in and expanded to provide greater resolution and range if desired. Such a lookup table would thus eliminate any potential issues caused by the decreasing amount of oxygen provided at higher RRs (as seen in Fig. 3) by having already accounted for the FIO2 that will actually be delivered.

This also provokes thought on how the concept of FIO2 is viewed. FIO2 is regarded mainly as a therapeutic value, determining the oxygen content provided to a sick patient. However, it may be more useful to in fact consider FIO2 in a diagnostic sense—or to consider it not at all in the case of autonomous control. For instance, a patient being on 70% oxygen may be more indicative of his level of lung injury than of the quality of his care. What’s more, the virtually ubiquitous report of oxygen conservation under closed loop control indicates that patients were likely hyper-oxygenated to begin with.\(^2,3\) The clinician drive to prevent the well-known and serious deleterious effects of hypoxia eschews the murky fact that hyperoxemia may have noxious effects at well, and possibly at FIO2 above only 0.40.\(^4,14\) A possibility for improved patient care exists if the decisions are put in the unbiased hands of the computer program, which can adjust FIO2 to whatever means necessary to achieve and maintain normoxia ($\text{SpO}_2 = 94\% \pm 2\%$). This autonomous integrator—satisfied by this design—could thus both improve patient care and conserve resources, without the care provider having to get wrapped up in the process. This offers a significant freedom to tend to other clinical responsibilities, as FIO2 was found to be the most frequently adjusted ventilation parameter in the management of critically ill patients under military care.\(^1,3\)

The current study also has several limitations. First, it is of course only a model, having been performed on a test lung. In vivo studies will be needed, likely first with a porcine model of acute lung injury, then moving on to clinical studies. The critical addition here will be the monitoring of the actual effect of the system on SpO2 and blood gases; and using the SpO2 reading to be able to provide active feedback and thus let the closed loop control operate freely and fully, as studied previously in the absence of the concentrator.\(^3\)

Being a passive test lung, the model also did not incorporate the addition of any spontaneous breathing; a fully capable system would have to be able to adjust for this. Testing may also be desired in different ventilation modes: this study only considered volume control mode, not pressure control or other variations. Likewise, testing was done at only a single standard lung compliance; this choice was made mainly to eliminate the inclusion of an additional variable of negligible significance at this point in system verification. However, consideration may be paid to the effect of this value in future experiments, as it could potentially impact the performance of the system in various disease states that would alter pulmonary compliance/resistance. Also, when adding in a pulse dose from the concentrator at the beginning of the breath, the system currently adjusts for volume, but not for inspiratory time ($T_i$); this results in longer inspiratory times than initially set when in pulse mode. The system must be made to either adjust for this, or to at least have a way of indicating the true resultant $T_i$. The former option is likely preferential, because the changed $T_i$ could otherwise alter the resultant I:E ratio, whose value can be important in the ventilatory management of a sick patient.

CONCLUSIONS

This study demonstrates functionality for a ventilation system that incorporates closed loop control of oxygenation and oxygen concentrator integration. The system was shown to provide viable amounts of oxygen across a range of clinical settings; and, especially when using coordinated pulse dose ventilation, to do so in a manner that potentially maximizes effect and certainly minimizes resource consumption. Such technology is of particular interest in austere settings such as far-forward military operations and disaster relief scenarios. Further testing and development is needed to eventually create and validate a single device capable of providing the level and type of care whose vision originates with this study.

ACKNOWLEDGMENTS

Software development for system by Rick Pettys, PE, and the team at Sparx Engineering, Houston, TX.

AUTHOR CONTRIBUTIONS: MMG – study design, system verification, experimentation/data acquisition, data analysis, and manuscript draft.

TCB – initial study design and system troubleshooting.

RDB – study design
and interpretation, system concept, and manuscript revision. JAJ – long-term project design/direction and author of initial proposal.

REFERENCES
An Overview of the Efficacy of a Next Generation Electroceutical Wound Care Device

Hosan Kim, PhD*; Soon Park, PhD†; Greggory Housler, MS‡; Vanessa Marcel, MS†; Sue Cross, PhD†; Mina Izadjoo, PhD*

ABSTRACT Novel approaches including nonpharmacological methodologies for prevention and control of microbial pathogens and emerging antibiotic resistance are urgently needed. Procellera is a wound care device consisting of a matrix of alternating silver (Ag) and zinc (Zn) dots held in position on a polyester substrate with a biocompatible binder. This electroceutical medical device is capable of generating a direct current voltage (0.5–0.9 Volts). Wound dressings containing metals such as Ag and/or Zn as active ingredients are being used for control of colonized and infected wounds. Reports on the presence of electric potential field across epithelium and wound current on wounding have shown that wound healing is enhanced in the presence of an external electrical field. However, majority of the electrical devices require an external power source for delivering pulsed or continuous electric power at the wound site. A micro-electric potential-generating system without an external power source is an ideal treatment modality for application in both clinical and field settings. The research presented herein describes efficacy evaluation of a wireless bioelectric dressing against both planktonic and biofilm forms of wound pathogens including multidrug resistant organisms.

INTRODUCTION

Skin protects the body from infection and maintains moisture balance.1 The human skin also has an innate ability to regenerate itself after trauma. Acceptable practice in treatment of skin lesions after injury or surgery is to protect the site from further trauma or infection by covering the injured site with a clean gauze. This barrier method has been the standard of care for hundreds of years, and in modern times, silver (Ag)-based wound care products are used. A recent critical finding is the presence of a primitive, physiologic electrical signal generated immediately at the time of skin injury. This electrical signal begins within the layers of skin and with skin injury, the electrical field lines project into the wound space directing the migration of cells and stimulating additional energy production required for cell migration and proliferation.2–4 This measurable signal can externally be replicated to enhance the normal healing process or to jump start a stalled healing event.4–7 In addition, electricity plays an important role in bioburden inhibition.8–10

As a result of the emergence of antibiotic and multidrug resistant (MDR) wound pathogens, there is a growing need for development of novel and effective wound care products. Next generation approaches, such as nonpharmacologic strategies against hard-to-eradicate wound pathogens are being introduced in clinical settings. There is growing recognition that energy-based technologies (electroceuticals) can have a diverse transformative impact on the health care field including wound care. Procellera (Vomaris Innovations, Tempe, Arizona) is a Food and Drug Administration-cleared, microcurrent generating antimicrobial wound dressing consisting of a matrix of alternating Ag and zinc (Zn) dots held in position on a polyester substrate with a biocompatible binder (Fig. 1). The antimicrobial activity existing in Ag-coated wound dressings is due to Ag’s ability to block the energy metabolism functioning across bacterial membranes.11 The broad-spectrum antimicrobial activity of Ag and Zn impregnated on the polyester has great activity against wound pathogens including antibiotic-sensitive or -resistant bacterial strains. This next generation electroceutical device is easily activated in presence of a conductive fluid, such as wound exudate or exogenous fluids like saline, and generates a physiologic level of electrical energy. This microelectric field can augment the natural electric field of injury initiated following skin wounding.

It has been known that a physiological current is necessary for initiation of the wound healing and transport of cells to the wound site.12,13 An electrical stimulus is essential for skin repair and regeneration,14,15 as it is the earliest guidance signal on tissue wounding to initiate cell migration and reepithelialization.14,15 In addition, electricity plays an important role in bioburden inhibition.8–10

In human skin, multicellular epithelia have dynamic barriers to protect host from potential threats such as infections and toxic materials and organize to pass selective ion transport across epithelial barriers.16,17 The polarized ion transport generates endogenous transepithelial electric potentials (TEPs) at millivolt levels.18,19 Since TEPs play multiple roles in biological events during development and regeneration of damaged tissues,15–17 once they collapse at the wound site, it...
results in generation of wound currents toward wound center from wound edge because of differences in electric potentials (EPs) between sites. These electrical properties in living organisms are commonly called bioelectricity. Bioelectricities play overriding roles in directional migrations of various cell types to organize the wound healing processes. Through various reports on accurate measurement of bioelectric currents, microelectric stimulation on wound healing, microcurrent therapy for wound healing, and damaged cornea, the EPs and currents demonstrated improved healing effects. However, most electrical wound care devices require an external power source that delivers pulsed or continuous electric power at the wound site. Therefore, the microelectric potential-generating system without an external power source will have significant application in both clinical and military field settings. Our service members are injured and disabled in the line of duty. In battlefield the speed and effectiveness of emergency treatment can mean the difference between life and death. A wound treatment paradigm aimed at accelerating tissue repair, reduced pain, dysregulated inflammatory response, and bioburden precisely suits the unique requirements of our Military Health System.

The current standard of care to repair tissue damage due to burn, trauma, or surgery relies mainly on the body’s own regenerative ability. The increase in infection-related injury, amputations, and death requires a fresh look at the regenerative process and how external intervention may play a significant role in reducing time to heal and in the prevention and treatment of wound-related infections. In this review, we will describe the efficacy of a novel bioelectric wound care device against clinical wound pathogens including MDR organisms, generation of the electrical signals, and the wound healing properties of an electroceutical wound care technology.

**Antibacterial Efficacy Against Clinical Wound Pathogens**

We postulated that the bioelectric dressing, consisting of a discrete Ag and Zn matrix pattern of microcell batteries could decrease wound infection by exerting an electricidal antimicrobial effect. Our testing method and procedures are described as follows:

Swatches of test and control textile materials were tested quantitatively for antibacterial activity by Method 147 entitled “Antibacterial Activity Assessment of Textile Materials: Parallel Streak Method (http://www.aatcc.org)” from the American Association of Textile Chemists and Colorists. In brief, an overnight bacterial culture was diluted to 10² colony forming units (CFUs) and applied on the sample and control fabrics each for 0 hour and 24 hours contact times at 37°C to compare the survival rate as CFUs. All of the sample (Procellera 2” × 2”) and control (gauze and blank polyester) fabrics were inactivated with quencher, including sodium thioglycolate in phosphate buffered saline (PBS) solution and washed off with PBS. Bacteria were mechanically separated from the fabric by vigorous vortexing and sonication to count surviving bacterial colonies using the agar plating method.

To examine the in vitro broad-spectrum antibacterial efficacy of the composite bioelectric dressing against clinically important bacterial wound pathogens, we tested the efficacy of the bioelectric dressing against various antibiotic-sensitive Gram-positive and -negative bacteria such as *Acinetobacter baumannii*, *Acinetobacter calcoaceticus*, *Enterococcus faecalis*, *Escherichia coli*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Staphylococcus epidermidis*, and controls. *P. aeruginosa* American Type Culture Collection (ATCC) 27853, and *S. aureus* ATCC 25923. A total of 28 bacterial isolates were tested from which 20 strains were Gram-positive and eight were Gram-negative. This dressing showed superior in vitro bactericidal activities against most of the nosocomial bacterial pathogens. However, it showed bacteriostatic activity against the *E. faecalis* isolate. Later, we tested the efficacy of this bioelectric dressing against 6 antibiotic-resistant clinical pathogens, such as extended spectrum beta-lactamase-resistant *K. pneumoniae*, MDR *P. aeruginosa*, and methicillin-resistant *S. aureus* strains. We observed 100% kill (Fig. 2) against all of the MDR organisms tested. In addition, seven vancomycin-resistance strains such as vancomycin-intermediate (VISA) and -resistant *S. aureus* (VRSA) species were tested to determine any associations between bacteriostatic activity and vancomycin resistance. As shown in Figure 3, all of the bacteria tested were killed by the bioelectric dressing at 24 hours. These findings demonstrated bactericidal efficacy against VISA and VRSA isolates.

This bioelectric dressing demonstrated strong bacterial killing effects reaching 100% kill at 24 hours, except for *E. faecalis* isolates, which showed some efficacy after 48 hours incubation period indicating that prolonged incubation may be required to eradicate the *Enterococcus* species. Membrane modification or efflux pump overexpression in the cell wall structure or other cell membrane of *Enterococcus* species may be associated with the bacteriostatic activity. Clearly, the antimicrobial properties of this dressing is derived from wound edge because of differences in electric potentials.
from the effect of Ag and Zn antimicrobials in addition to the bioelectric currents.

**Anti-Biofilm Efficacy Against Clinical Wound Pathogens**

Clinical bacterial wound pathogens in chronic infections are mostly associated with the formation of mono- or multispecies biofilms leading to difficult-to-treat infections, antibiotic resistance, and recurrent infections.\textsuperscript{27–29} Chronic wound pathogenic bacteria are mostly engaged in the wound biofilm formation, therefore, the eradication and treatment for infection control and prevention becomes complicated.

The presence of biofilms in chronic and nonhealing wounds is a major clinical concern.\textsuperscript{30} A number of approaches and testing methodologies for anti-biofilm efficacy assessment are being used in the wound care field. Surviving bacteria could form biofilms on a number of abiotic surfaces under static, continuous agitating, and hydrodynamic conditions. The bioelectric dressing was tested for its anti-biofilm efficacy against biofilms formed by both the poloxamer and colony drip-flow reactor (DFR) biofilm models, which are described below.

**Poloxamer Biofilm Model**

We evaluated the anti-biofilm efficacy of the bioelectric dressing in poloxamer biofilms generated under static conditions. We used poloxamer since it can be used for antimicrobial efficacy testing and supports growth of bacteria in a biofilm state.\textsuperscript{31} We established a poloxamer biofilm model using glass coverslips against both mono- and multispecies. In the mono-species biofilm testing against 10 clinical wound pathogens, we observed approximately 2- or 3-fold log\textsubscript{10} reduction in bacterial growth after 24 hours incubation compared to those of controls such as no treatment, gauze, and blank polyester without Ag and Zn metals. In addition, we developed multispecies biofilms in the poloxamer model employing a mix of 4 bacterial pathogens for efficacy testing against polymicrobial biofilms. The bioelectric dressing demonstrated approximately 1- or 2-fold log\textsubscript{10} reduction in chromogenic agar plates, which is an alternative approach for isolation of several bacterial strains (DRG International Inc., Springfield, New Jersey) compared to those of controls. Our poloxamer hydrogel biofilm model is appropriate for anti-biofilm efficacy evaluation of this dressing and demonstrated anti-biofilm activity against not only mono- but also multispecies biofilms formed by MDR clinical...
pathogens compared to those of controls such as gauze and blank polyester.\textsuperscript{32}

**DFR Biofilm Model**

We also employed an in vitro colony DFR biofilm model for making biofilms under low shear condition (5 mL/h) similar to natural environments (Fig. 4).\textsuperscript{33,34} The bioelectric dressing and controls such as gauze and blank polyester as controls were applied directly on the biofilms that are continuously deposited onto hydrophobic filter membranes for 72 hours at room temperature. Biofilm formation was confirmed by crystal violet staining and subsequent microscopic observation. Through vigorous shaking and sonication processes, the released bacteria were serially diluted and plated onto bacterial agar plates. The surviving bacterial colonies were counted after 24 hours incubation at 37°C. During the 72 hours incubation time period, it was shown that the *A. baumannii* biofilms were well deposited onto the blank polyesters but not onto the bioelectric dressing. We observed inhibition in bacterial growth on the bioelectric dressing when compared to those of blank polyester treatments. Crystal violet staining and microscopic examination of the blank polyester showed development of large and fully grown biofilms. The anti-biofilm activity of the bioelectric dressing against *A. baumannii* biofilms in colony DFR was more than 10-fold effective in reducing bacterial numbers compared to that of blank polyester which showed accumulation of more than 10\textsuperscript{9} CFUs/mL (Kim H, Bower B, Izadjoo M: In vitro Efficacy Testing of a Novel Wound Dressing against Clinical Bacterial Biofilms. Presentation at the Symposium of Advanced Wound Care, Las Vegas, NV. September 2013).

**Bioelectric Measurement**

There are growing number of reports on the beneficial effects of microelectric currents on wound healing, including anti-microbial effects and impact on common cellular functions including development and physiology.\textsuperscript{16,17,19,35} The (Ag–Zn)-printed bioelectric dressing as a wound care device was designed to accommodate (Ag–Zn) half-cell potentials by alternatively printing them on a woven polyester material in a well-characterized dot-matrix pattern. This wound care device can accommodate conceptually 200 embedded microcell batteries per square inch among neighboring Ag and Zn elements tightly printed elemental grains (2–10 microns) of Ag (900 mg/cm\textsuperscript{2}) and Zn (300 mg/cm\textsuperscript{2}) dots on a polyester sheet as a test specimen under various conditions (Fig. 5).\textsuperscript{36}

We measured the amount of generated relative EPs on the device (2” × 2”) using a calibrated microprobing system on a three-axis micrometer stage (Fig. 6) following presoaking with various conductive solutions (500 µL) such as 0.85% saline solution, culture media, or bacterial culture suspensions (approximately 10\textsuperscript{5} CFUs of *E. coli* to mimic infections), respectively. We observed constant generation of microelectric potential of the Ag- and Zn-printed dressing under all tested conductive solutions.\textsuperscript{37}

Taken together, the (Ag-Zn)-printed device could generate and successfully sustain EPs without changing electrical
properties including stable polarities under the presence of various conductive fluids, which can be encountered in wound environments. Numerous microelectrical circuits are expected to be created between these two neighboring Ag and Zn dots tightly printed on the polyester sheet without requiring any external power source. Thus, the consistently generated EPs at each battery couple with (Ag–Zn)-based dressing would restore disrupted physiologic bioelectric signals on wound sites leading to improved wound healing with antimicrobial activity.

Wound Healing Efficacy
Recent studies assessing the bioelectric dressing have yielded encouraging findings, including a rapid increase in degree and rate of epithelial migration of deep partial-thickness wounds versus controls in an in vivo porcine study. Accelerated human keratinocyte migration, improved mitochondrial function, and improved integrin expression were reported in a series of in vitro studies. A growing body of evidence supports the use of this technology in the reepithelialization of acute and chronic wounds of varying complexities and etiologies, and has consistently demonstrated significantly faster healing times versus controls, sharper wound closure trajectory and a more robust wound healing trend, and improved scar appearance and subjective patient outcomes. Procellera is an FDA-cleared bioelectrical dressing, which has been used for various acute and chronic wound indications.

DISCUSSION
Recent research indicates that pathogens exist as discrete species but can coinhabit a space, communicate intra- and interspecies, and set up defenses against certain microbial species. In an early state, these microorganisms exist as free floating, planktonic organisms looking for a place to latch onto. These organisms are in their weakest state in this form. Once the organisms have latched onto a site they begin to replicate. When the number of organisms reach a threshold or “quorum” they are able to function as multicellular communities and can communicate between each individual organism as well as with the entire mass. One common result of this communication is the excretion of a film to coat all the organisms, called a biofilm. This biofilm becomes a protective shield allowing the inhabitants to proliferate under the shield.
and, through genetic manipulation, modify their resistance to antibiotics and antimicrobials.\textsuperscript{39}

Infections with MDR pathogens within the military environments are spreading at an alarming rate,\textsuperscript{40} with combat-related injuries at greater inherent risk of infectious complications, whether from colonization before injury, environmental contamination, or nosocomial infections.\textsuperscript{41,42} The debilitating nature of wound infections in service members has potentially devastating consequences. Despite best infection control methods, failure of traditional antibiotic therapies in combating MDR organism colonization and infection in military treatment facilities remain a topic of significant concern. The overuse of antimicrobial prophylaxis, associated complications, and consequent microbial resistance has brought increased focus to the unmet need for innovative and aggressive nonpharmacologic approaches to control, reduce bioburden.

Energy is produced within human cells and is essential for cellular metabolism. The energy must increase during the wound repair cycle, to stimulate cell migration, deposition of an extracellular matrix, and cell “proliferation.”\textsuperscript{43} In the sterile environments, this energy is used exclusively by the cells. However, when pathogenic microorganisms are present, there is competition for the energy. Although various energy-based modalities with differing methods of action have been developed for treatment of soft-tissue wounds, many are complex in use, capital intensive, require tethered access, and myriad consumables, posing considerable limitations for clinical or home use. Challenges with prolonged wound-healing times and multifactorial sources of microbial contamination have created a need for portable, antimicrobial wound dressing, and treatment solutions.

“Procellera,” a microcurrent generating antimicrobial dressing, is currently being used in the United States for various acute and chronic wound indications, including partial and full thickness wounds such as pressure ulcers, venous ulcers, diabetic ulcers, burns, surgical incisions, and donor and/or recipient graft sites. A growing body of preclinical and clinical data support the efficacy of this device in treating soft tissue wounds, burns, and inflammatory conditions.

Unique in its method and mode of action, the portable, wireless nature of this bioelectric dressing is the evolution of an electrical stimulation device into a dressing-like form. This electrostimulus was previously shown to nonpharmacologically downregulate quorum-sensing genes and downregulate the antibiotic resistant genes within a biofilm environment.\textsuperscript{36} An alternative method of action also observed was the down-regulation of Glycerol-3-phosphate dehydrogenase (GPDH), an enzyme within the organisms required for metabolism. Its versatility, portability, and ease of use make it ideally suited for military field applications as well as clinical and home applications. A growing body of research on the bioelectric dressing demonstrates significant benefits in the treatment of partial and full thickness, acute and chronic wounds, including electroceutical antimicrobial efficacy and enhanced cellular migration impacting wound healing. Laboratory studies conducted by our team successfully demonstrated significant efficacy of the bioelectric dressing against various MDR wound pathogens and anti-biofilm properties of this dressing in both mono- and poly-microbial biofilm models.

Given the findings presented in this article, the use of a next generation, close proximity electrically active technology to optimize the wound-healing environment may better control bioburden, potentially reducing the need for local and/or systemic antibiotics, and improving the quality of life of injured military personnel. This dressing has significant utility to treat wounds and wounds infected with antibiotic and MDR pathogens. Results from our findings point to the bioelectric dressing as a viable and useful option for military caregivers to add as an anti-infection modality in clinical practice and may serve as a promising treatment modality for battlefield-acquired wounds.

REFERENCES


Telehealth Coaching: Impact on Dietary and Physical Activity Contributions to Bone Health During a Military Deployment

Laura L. Frank, PhD*; MAJ Mary S. McCarthy, AN USA, (Ret.)†

ABSTRACT  Purpose: To examine the difference in bone health and body composition via blood biomarkers, bone mineral density, anthropometrics and dietary intake following deployment to Afghanistan among soldiers randomized to receive telehealth coaching promoting nutrition and exercise. Methods: This was a prospective, longitudinal, cluster-randomized, controlled trial with repeated measures in 234 soldiers. Measures included heel bone scan for bone mineral density, blood biomarkers for bone formation, resorption, and turnover, body composition via FatPal, resting metabolic rate via MedGem, physical activity using the Baecke Habitual Physical Activity Questionnaire, and dietary intake obtained from the Block Food Frequency Questionnaire. Results: There were significant increases in body fat (p = 0.00035), osteocalcin (0.0152), and sports index (p = 0.0152) for the telehealth group. No other statistically significant differences were observed between groups. Vitamin D intake among soldiers was ≤ 35% of the suggested Dietary Reference Intakes for age. Conclusions: A 9-month deployment to Afghanistan increased body fat, bone turnover, and physical activity among soldiers randomized to receive telehealth strategies to build bone with nutrition and exercise.

INTRODUCTION

The Army has provided the majority of U.S. troops to Iraq and Afghanistan—over 1.5 million troop-years between September 2001 and December 2011, more than all other services combined. Many soldiers deploy for 9 to 12 months at a time and on multiple occasions. The repetitive physical demands required of deployed soldiers in their military occupational specialty often places them at risk for musculoskeletal injury. Indeed, two of four main postdeployment health concerns endorsed by all service members are musculoskeletal in nature.

Risk factors for bone-related injuries include increased age, female gender, smoking, inadequate calcium intake, poor fitness level, and over-training. Moreover, the hot, dry climates associated with deployments promote dermal calcium losses, which, when combined with inadequate dietary calcium and vitamin D intake, may place the soldier at greater risk for musculoskeletal injuries. With limited duty, lost workdays, and much discomfort, musculoskeletal injuries have a greater impact on the health and readiness of U.S. Army service members than any other medical condition in peacetime or conflict.

Although the impact of deployment on musculoskeletal injuries is clear, the impact of deployment on body composition and bone mineral density (BMD) is less clear. A 9-month deployment to Afghanistan among 110 infantry soldiers had a negative impact on strength, aerobic fitness, and body composition resulting in decreased percent fat-free mass and increased percent fat mass. In a similar study, Lester et al reported a decline in aerobic performance and increases in fat mass among 73 combat arms soldiers after a 13-month deployment to Iraq. However, improvements in upper and lower strength, upper body power, and increased lean mass were also reported.

Genetic factors account for 60% to 80% of peak bone mass with the remainder influenced by hormonal status, diet, environmental factors, and physical activity/exercise. Enlisted soldiers comprise 82% of the total Army force, with 69% of them between 17 and 30 years of age. This age range coincides with the period of peak bone mass, when the growth in the size of bones and the accumulation of bone mineral has stabilized. Evans et al reported that military training increased bone formation and resorption markers, suggesting rapid onset of strenuous exercise accelerates bone turnover similarly in men and women. These authors surmised that although bone turnover markers were higher in men than women, bone formation status may be related to aerobic fitness and serum calcium independent of gender. Furthermore, bone turnover markers may be affected by small changes in endocrine regulators related to nutrition. Previous findings have demonstrated that inadequate consumption of calcium and vitamin D as well as a decrease in exercise while deployed may be detrimental to bone health. Results from a recently published study examining the impact of diet, physical activity, and bone density in 53 soldiers before and after a 12-month deployment in Iraq showed that soldiers did not meet the recommended daily allowance (RDA) for vitamin D. No significant relationships between change in diet or physical activity and BMD were observed in this study; however, the study was underpowered for these outcomes. Therefore, more studies are needed to investigate...
the relationships between diet, physical activity, and BMD in a larger military population.

The purpose of the present study was to determine if a telehealth coaching initiative is superior to a one-time nutrition and fitness education class regarding: (a) dietary contributions to bone health and (b) exercise contributions to bone health, assessed before and after deployment. Dietary contributions included the intake of calcium and vitamin D from food, beverages, dietary supplements, and medications. Exercise contributions were assessed by a physical activity questionnaire, biomarkers of bone status, and heel bone density using portable ultrasound.

METHODS

Study Design and Participants

This study was approved by the Madigan Army Medical Center Institutional Review Board (no. 122909). Grant funding was provided by the TriService Nursing Research Program (no. HU0001-10-1-TS15, N10-C02). Informed consent was obtained from each soldier before participation in the study. 234 soldiers volunteered from two separate combat arms brigades on Joint Base Lewis-McChord, Washington. Volunteers were eligible to participate if they were active duty male or female soldiers, aged 18 to 30 years, in good physical and mental health, with the ability to read and write, scheduled to deploy to Afghanistan for at least 9 months. Soldiers with a diagnosis of bone disease, current stress fracture, history of electrolyte imbalance, or eating disorder were excluded from the study. The control group (CG) and telehealth group (TG) were scheduled to deploy for 12 months; thereafter, operations were curtailed and the TG returned after 9 months. Having left several months before the TG, the CG remained in theater for a full 12-month deployment period.

In this prospective, longitudinal, cluster-randomized, controlled trial with repeated measures, soldiers who were randomized to the TG received on-demand health-related messages via the Army Milbook and Outlook Mail platforms in an attempt to reach as many participants as possible throughout the study. The website was maintained and tracked by the Project Director (a registered dietitian) and the Research Assistant, both of whom have Master’s degrees in nutrition. The team also included an exercise physiologist, a nurse scientist, and an endocrinologist to address questions that may arise within their areas of expertise. The soldiers in the TG had 24/7 access to the website, with new content posted weekly on exercise, nutrition, and bone health. Any e-mail queries to the study team were addressed via direct e-mail within 24 hours. Soldiers have demonstrated that they are adept at using all methods of internet communication.

Data Collection

Baseline and follow-up measurements were taken within 30 days of deployment and redeployment. Once enrolled, the CG attended a 1-hour educational session on healthy eating and exercise to support bone health. Following the session, baseline anthropometric measurements took place with soldiers wearing the army physical fitness uniform of shorts, t-shirt, socks, and running shoes. Soldiers rotated through six stations, taking approximately 2 hours to complete all measures.

Demographic and Anthropometric Measurements

A 12-item demographic tool captured relevant personal and family history related to bone health. This included: age, gender, ethnicity, past and current use of tobacco products, alcohol consumption, caffeine intake, current list of medications/supplements, history of endocrine condition, history of bone disorders/stress fractures, and family history of osteoporosis.

Body composition included height (inch) measured using a stadiometer (Seca 213, Portable Stadiometer Height Rod, Chino, California) and body weight (lbs) measured using a digital scale (Detecto Model DR400 electronic scale, Webb City, Missouri); no shoes were worn for these measurements. Waist circumference was obtained using a nonelastic, coated fabric measuring tape of standard length (120 inches), which measured waist circumference at the minimal abdominal circumference (females) or level of the navel (males) rounded down to the nearest 0.1 inch. We used the MedGem indirect calorimeter (Microlife Medical Home Solutions, Inc., Golden, Colorado) to perform testing for resting metabolic rate (RMR). The MedGem handheld indirect calorimeter has been clinically validated and is a Food and Drug Administration 510K-cleared, class II, medical device. Based on a systematic review, the MedGem device is a valid and reliable indirect calorimeter for energy assessment in most adults. Body fat (BF) analysis was measured using near infrared reactance (Futrex, Inc., Hagerstown, Maryland) and recorded as percent BF; body mass index was also obtained via this method. Calcaneal bone density was measured by ultrasound using a Hologic Sahara device (Foremost Medical Equipment, Rochester, New York) on the dominant heel. The Hologic Sahara heel densitometer boasts the lowest ionizing radiation exposure of any competing densitometer. Results display BMD with an extensive reference database of >12,000 National Health and Nutrition Examination Survey subjects, T-score, Z-score, fracture risk assessment, and patient trending. Several studies have reported the advantages of using portable heel densitometry as it is fast, reliable, and field expedient.

Physical Activity Assessment

Physical activity was assessed at baseline and following deployment using the Baecke Habitual Physical Activity Questionnaire. This 16-item tool was validated on males and females of similar age as our study population and was designed to assess physical activity during work, sport, and leisure-time activities. The Baecke Habitual Physical Activity Questionnaire allows for computation of a total activity score and individual subscale scores for work, sport, and leisure, with mean scores representing indices of physical activity in each of these areas. Work activity measures frequency of sitting, standing, walking, lifting
heavy loads, and sweating, as well as degree of fatigue. Questions involving sports measure frequency, type, and duration of individual and/or team sports (golf, aerobic exercise, running, walking, basketball, tennis, weight/strength training, etc.). Leisure (off duty) activity time is assessed using a Likert scale (1 = never, 5 = very often) for sweating, participating in sports, watching TV, walking, and bicycle riding. Overall physical activity was determined by calculating work, leisure, and sport subscale scores, which were then summed and described on a scale of 1 to 15 (1 = low, 7.5 = moderate, 15 = high). The baseline survey assessed activity in the year before deployment, whereas the postdeployment survey assessed activity during deployment.

**Dietary Assessment**

Dietary intake was assessed pre- and postdeployment using the Block Food Frequency Questionnaire (FFQ) (Nutrition Quest, Berkeley, California). This questionnaire was previously validated in epidemiological studies and estimates the usual and customary intake of 110 food items. Respondents reported how often in the past year they typically consumed each food (1 = never, 9 = once per day). Pictures of individual portion sizes were provided to enhance accuracy of quantification (1/4 cup, 1/2 cup, 1 cup, and 2 cups). The FFQ contains questions about overall consumption of foods from each food group, beverage intake, and use of vitamin/mineral supplements. All completed FFQs were analyzed by NutritionQuest (Berkeley, California). Eleven nutrients (protein, fat, carbohydrate, calcium, sodium, vitamin C, D, and K, potassium, phosphorus, and cholesterol) and energy intake were included in this analysis. All nutrients were calculated as the sum of daily intake from food. Furthermore, supplemental intakes of vitamin C, vitamin D, and calcium were analyzed. Energy intake was compared to the U.S. Department of Agriculture’s Dietary Guidelines for Americans estimated energy intake for males and females 18 to 25 years of age with moderate physical activity levels.

All other nutrients were compared to the Institute of Medicine’s (IOM) dietary reference intakes (DRIs) for the 19 to 30 year life stage group; either as the RDA or the adequate intake (AI). According to the IOM, the RDA is the average daily dietary intake level; sufficient to meet the nutrient requirements of nearly all (97%–98%) healthy individuals in a group; calculated from an estimated average requirement. If sufficient scientific evidence is not available to establish an estimated average requirement, and thus calculate an RDA, an AI is used. The AI for other life stage and gender groups is believed to cover the needs of all healthy individuals in the groups, but lack of data or uncertainty in the data prevent being able to specify with confidence the percentage of individuals covered by this intake.

**Serum Biomarkers**

Biomarkers of bone health were drawn at baseline and follow-up; markers included calcium, 25-hydroxyvitamin D, osteocalcin, bone-specific alkaline phosphatase, and insulin-like growth factor (IGF)-1. We chose to include IGF-1 for its relevance to skeletal integrity and chronic stress to the body. Trained Army medics and/or nurses performed phlebotomy on site concurrent with study measurements in order to ensure all blood specimens were obtained, and processed in a timely manner. Specimens were placed in a cooler and transported to the hospital laboratory for processing every 2 hours. Fewer than 2% of specimens were lost to hemolysis or insufficient quantity of blood.

**Statistical Analysis**

Descriptive statistics and relational logic checks were utilized to identify missing and invalid values within the preliminary datasets. Continuous variables from dietary intake, body composition, and physical activity changes from pre- to postdeployment were analyzed using paired t tests. For categorical variables, the tests of difference in change over time were computed using generalized linear mixed models. The multivariate imputation by chained equations with classification and regression trees was used as the underlying modeling technique. Final analyses were conducted using the multivariate imputation by chained equations package in the statistical software R v3.0.1 (R Foundation for Statistical Computing, Vienna, Austria). False Discovery Rate (FDR) control was applied to correct for multiple comparisons.

**RESULTS**

Of 234 soldiers enrolled at baseline, 158 returned from deployment (TG, n = 66; CG, n = 92), yielding a 33% attrition rate. Reasons for not returning for follow-up included early redeployment to home station, injuries sustained in combat led to premature return, permanent change of station, or discharge from the service upon return. Demographic characteristics are shown in Table I. At baseline, the CG had significantly greater BF percentage compared to the TG (19.09 vs. 17.44, p = 0.04). Lifestyle factors (alcohol and tobacco use, history of stress fracture) were not statistically different throughout the study (Table I). Bone density remained stable over time. There was a significant difference in the proportion of soldiers with abnormal vs. normal calcaneal BMD, based on t score, between groups at baseline (p = 0.006) with the CG having a greater proportion of abnormal readings. Interestingly, the CG had an impressive rise in BMD upon return (+0.06 g) compared to the TG (+0.02 g) but this was not significant after applying FDR control (Table II).

**Anthropometrics and RMR**

Changes in anthropometrics and RMR from baseline to follow-up are shown in Table II. There was a significant change in postdeployment BF percent with an increase in the TG and a decrease in the CG (+3.94 ± 0.56% vs. −0.52 ± 0.42%,...
related to a second sport activity was higher in the TG than in the CG (+0.26 ± 0.12 vs. −0.14 ± 0.09; p = .07). Commonly reported sport activities included soccer, weight lifting, basketball, and running, all endorsed by the American Osteoporosis Foundation to build and maintain bone density. No significant changes in dietary or supplemental macro- or micronutrients were observed between groups. After comparing mean calories, macronutrients (protein, carbohydrate, and fat) and selected micronutrients to the IOM’s DRI, we determined that soldiers in this study closely met recommendations for caloric intake, and exceeded recommendations for macro- and micronutrients (Table IV). Dietary vitamin D intakes were low (range 31%–34% DRI) for both groups throughout the study.

**Physical Activity and Diet Outcomes**

Diet and physical activity indices pre- and postdeployment are shown in Table III. Sport index was significantly different between the TG and CG at baseline (2.68 ± 0.09 vs. 2.94 ± 0.05, p = .01) and trended in significance postdeployment (2.97 ± 0.09 vs. 2.76 ± 0.07, p = .08). Furthermore, a significant mean change from baseline to follow-up was observed, with the sport index higher in the TG and lower in the CG (+0.29 ± 0.11 vs. −0.17 ± 0.09, p = .02) on return from deployment. In addition, the index related to a second sport activity was higher in the TG than the index for the CG, showing a similar trend as for 1 sport activity (+0.26 ± 0.12 vs. −0.14 ± 0.09; p = .07). Commonly reported sport activities included soccer, weight lifting, basketball, and running, all endorsed by the American Osteoporosis Foundation to build and maintain bone density. No significant changes in dietary or supplemental macro- or micronutrients were observed between groups. After comparing mean calories, macronutrients (protein, carbohydrate, and fat) and selected micronutrients to the IOM’s DRI, we determined that soldiers in this study closely met recommendations for caloric intake, and exceeded recommendations for macro- and micronutrients (Table IV). Dietary vitamin D intakes were low (range 31%–34% DRI) for both groups throughout the study.

**Bone Health Outcomes**

Bone biomarkers are shown in Table V. Baseline 25(OH) vitamin D revealed a high rate of insufficiency (61%, level <30 ng/mL) and moderate level of deficiency (17%, level <20 ng/mL) in both groups. Participants significantly improved their vitamin D status postdeployment, with the CG achieving a “sufficient” level (mean = 34.35 ng/mL) and the TG remaining “insufficient” (mean = 26.02 ng/mL). Soldiers with 25(OH) vitamin D levels <20 ng/mL at baseline had a prescription written for them by the Unit Physician Assistant but we are unable to validate adherence to filling the prescription and taking the supplement. Osteocalcin, a bone turnover marker was significantly different between the TG and CG at baseline (21.12 ± 0.77 ng/mL vs. 18.48 ± 0.62 ng/mL, p = .01) and postdeployment (28.31 ± 1.08 ng/mL vs. 22.05 ± 0.99 ng/mL, p < .001). Furthermore, the TG showed a greater difference in mean-change for osteocalcin compared to the CG (+7.18 ± 0.78 ng/mL vs. +3.57 ± 0.77 ng/mL, p = .02). There were no statistical differences between groups in any other bone biomarkers.

**DISCUSSION**

This study indicates that diet and exercise coaching via telehealth methods to deployed soldiers is feasible but limited in its effectiveness for short-term overseas deployments. In this study, we did not see favorable changes in
TABLE III. Physical Activity and Diet Outcomes. The Data are Reported as Mean, (SE)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Baseline</th>
<th>Follow-Up</th>
<th>Change: Follow-Up to Baseline</th>
<th>FDR p</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CG n = 135</td>
<td>TG n = 85</td>
<td>CG n = 135</td>
<td>TG</td>
</tr>
<tr>
<td>Work Index</td>
<td>3.14 (0.03)</td>
<td>3.35 (0.05)</td>
<td>3.14 (0.03)</td>
<td>3.32 (0.04)</td>
</tr>
<tr>
<td>Sports Index</td>
<td>2.94 (0.05)</td>
<td>2.68 (0.09)</td>
<td>2.76 (0.07)</td>
<td>2.97 (0.09)</td>
</tr>
<tr>
<td>Sports Index 2</td>
<td>2.73 (0.05)</td>
<td>2.48 (0.08)</td>
<td>2.59 (0.07)</td>
<td>2.74 (0.09)</td>
</tr>
<tr>
<td>Leisure Index</td>
<td>3.28 (0.06)</td>
<td>3.14 (0.08)</td>
<td>3.24 (0.09)</td>
<td>3.37 (0.11)</td>
</tr>
<tr>
<td>Calories (kcal)</td>
<td>2,593.48 (112.61)</td>
<td>2,773.89 (161.11)</td>
<td>2,661.21 (112.95)</td>
<td>2,537.55 (142.78)</td>
</tr>
<tr>
<td>Protein (g)</td>
<td>101.94 (5.04)</td>
<td>108.59 (6.83)</td>
<td>107.71 (5.14)</td>
<td>101.82 (6.6)</td>
</tr>
<tr>
<td>Fat (g)</td>
<td>102.95 (5.15)</td>
<td>110.81 (7.27)</td>
<td>106.37 (4.96)</td>
<td>101.58 (6.43)</td>
</tr>
<tr>
<td>Carbohydrates (g)</td>
<td>316.83 (16.51)</td>
<td>332.45 (22.06)</td>
<td>326.35 (18.37)</td>
<td>288.97 (22.06)</td>
</tr>
<tr>
<td>Calcium (mg)</td>
<td>1,111.11 (54.82)</td>
<td>1,130.03 (70.61)</td>
<td>1,142.82 (53.91)</td>
<td>1,068.52 (65.07)</td>
</tr>
<tr>
<td>Phosphate (mg)</td>
<td>1,670.45 (82.18)</td>
<td>1,771.25 (117.54)</td>
<td>1,817.6 (94.28)</td>
<td>1,643.65 (115.42)</td>
</tr>
<tr>
<td>Sodium (mg)</td>
<td>4,227.83 (214.09)</td>
<td>4,442.03 (285.74)</td>
<td>4,551.85 (211.51)</td>
<td>4,199.16 (273.95)</td>
</tr>
<tr>
<td>Potassium (mg)</td>
<td>3,046.29 (147.75)</td>
<td>3,240.57 (185.69)</td>
<td>3,221.84 (155.19)</td>
<td>3,025.75 (168.86)</td>
</tr>
<tr>
<td>Vitamin C (mg)</td>
<td>147.41 (8.68)</td>
<td>166.25 (22.46)</td>
<td>169.3 (10.37)</td>
<td>141.89 (12.41)</td>
</tr>
<tr>
<td>Vitamin D IU</td>
<td>191.15 (11.75)</td>
<td>184.62 (12.88)</td>
<td>211.56 (14.61)</td>
<td>205.49 (17.03)</td>
</tr>
<tr>
<td>Vitamin K (mg)</td>
<td>137.71 (9.02)</td>
<td>146.62 (12.07)</td>
<td>176.16 (11.91)</td>
<td>154.3 (15.23)</td>
</tr>
<tr>
<td>Calcium Suppl (mg)</td>
<td>86.6 (18.56)</td>
<td>98.35 (22.86)</td>
<td>107.97 (21.4)</td>
<td>114.53 (29.25)</td>
</tr>
<tr>
<td>Vitamin D Suppl IU</td>
<td>69.42 (11.79)</td>
<td>65.55 (12.29)</td>
<td>60.74 (17.63)</td>
<td>123.53 (24.71)</td>
</tr>
<tr>
<td>Vitamin C Suppl (mg)</td>
<td>66.88 (19.87)</td>
<td>35.21 (8.34)</td>
<td>75.16 (26.7)</td>
<td>79.2 (33.8)</td>
</tr>
</tbody>
</table>

CG, control group; FDR, False Discovery Rate; g, grams; mg, milligrams; kcal, kilocalorie; IU, International Unit; Suppl, supplemental; TG, telehealth group; vit, vitamin.

body composition, or significant changes in BMD, RMR, or dietary/supplemental macro- or micronutrient intake.

Percent BF increased from 17.44% to 21.28% in the TG while percent BF decreased from 19.09% to 18.56% in the CG. While unexpected and disappointing, the increase in BF percent in the TG is in agreement with other studies showing a significant increase in BF among deployed soldiers.5,6 These previous studies also reported a decrease in aerobic exercise during deployment, which might explain the increases in BF. Similarly, studies in athletes have reported increases in BF when aerobic training from running is reduced or stopped.24,25 Our study did not measure aerobic fitness, however, there was a significant positive change in sport activity in the TG, which was not evident in the CG, both for 1 sport and 2 sport activities, leaving few explanations for the increased BF. The increase in percent BF was not explained by increases in caloric intake, as there were no significant changes in mean calories consumed between groups over time, and neither group exceeded the Dietary Guidelines for Americans for this age group.19 As concerning as it is that the TG increased their mean BF percent without an apparent cause, the CG results demonstrate that not all soldiers experience the same fate since they had a decline in BF percent, even if modest in its reduction.

Increases in weight and percentage of BF measurements continue to be of considerable concern in the military. Soldiers are required to meet body composition standards biannually.26 If soldiers are unable to meet these standards Commanders can refer the individual to a weight management program where sustained inability to meet BF standards can result in subsequent discharge from military service. Therefore, continued strategies targeting appropriate reduction of fat mass and an increased lean body mass are essential for a fit-and-ready force. Studies of deployed soldiers show the need to increase aerobic training and improve aerobic fitness with the specific aim of optimizing health and mission performance.

As stated, there were no significant increases in BMD in the TG. The greater increase in BMD for the CG, along with lower percent BF, is difficult to explain given that scores for all work, sport, and leisure activities were reduced during the deployed interval. Calcaneal measurements have not been correlated with more standard tests of BMD, namely dual energy x-ray absorptiometry, and therefore it is unknown how representative this reading is of overall BMD. Few studies have examined the impact of deployment on BMD. Sharp et al7 reported improvements in overall BMD as measured by dual-energy x-ray absorptiometry over a 9-month deployment to Afghanistan. Carlson et al11 reported mixed results with increases in spine BMD and decreases in femoral neck BMD after a 12-month deployment to Iraq. The results of this study did show improvement in postdeployment bone health, with nonsignificant but desirable increases in calcaneal BMD in both groups, and significantly greater mean-change in osteocalcin levels in the TG compared to the CG. To the present authors’ knowledge, this is the first study to examine osteocalcin levels before and after military deployment, especially as it pertains to telehealth intervention. Others have shown a positive relationship between physical training and osteocalcin levels.6 Therefore, because there was a significant increase in the sport index in our TG and the mean age of our cohort would presumably still be accruing bone mass, it is not surprising that a significant increase in this bone turnover marker was detected. Another hypothesis is that bone trauma, such as stress

Telehealth Coaching

MILITARY MEDICINE, Vol. 181, May Supplement 2016 195
No significant relationships were observed for dietary factors between groups. However, this study underscores the fact that male soldiers fail to meet the DRI for vitamin D, and even when supplemental vitamin D is consumed, soldiers meet ≤50 percent of the DRI. This study is in agreement with a recently published study reporting low vitamin D intake among deployed soldiers. Vitamin D is necessary for normal bone growth and maintenance. Low 25(OH) vitamin D levels have been associated with an increased risk in musculoskeletal injuries in active duty military, therefore more public health programs for military personnel are needed to raise awareness of the nutritional factors that influence change in bone density and stress fracture risk. A recent study showed that low-fat dairy products and the major nutrients in milk (calcium, vitamin D, and protein) were associated with greater bone gains and a lower stress fracture rate, and that potassium intake was also associated with greater gains in hip and whole-body BMD. Anecdotal evidence suggests there may be a real or perceived fear of renal calculi formation while deployed because of alterations in diet and hydration, this may explain dairy-avoidance behaviors. Offering vitamin D fortified non-dairy beverages (such as fruit juice) may help offset the low vitamin D status seen in so many returning soldiers. We are unable to offer any comment on the potential effects of sun exposure as these data were not collected but this may have contributed to increases noted in either or both groups. Choices regarding lifestyle are important for all young adults but the challenges to a balanced diet and exercise regimen in the deployed environment may have long-standing consequences for the soldier, as well as total force fitness. Health promotion efforts by Brigade-level public health nurses and dietitians during deployment and in peacetime can have a major impact on lifestyle behaviors and bone health of young soldiers who are developing peak bone mass. Early and aggressive educational outreach efforts can prevent chronic musculoskeletal conditions and disabling osteoporosis as well as reduce musculoskeletal-related injuries while deployed.

**LIMITATIONS**

There are several limitations to the present study. First, and foremost, is the low sample of females enrolled (n = 14) and their attrition rate of 93% in particular. For this reason, we did not include females in our pre-post deployment analyses. This is most unfortunate as it leaves many questions unanswered regarding their deployment health behaviors, rate of musculoskeletal injuries, and differences in bone density.
On-demand telehealth coaching occurred for the TG from April 2012 to December 2012 via e-mail and a study website. Telehealth strategies may not have been as effective as anticipated because of limitations in availability of and access to the internet. The study team anticipated challenges with internet access and developed strategies to avoid these pitfalls with information management experts. However, there are unknown confounding circumstances in the austere deployed environment, particularly when a unit is separated and sent to various forward operating bases. Indeed, recruitment and attrition issues in military clinical trials and health research studies have been previously published.33 Furthermore, a shorter interval of time for the intervention during deployment (from 12–9 months) occurred, which may have affected our anthropometric and bone health results.

Although the Block FFQ has been used by this research team in several other studies to collect frequency and quantity of food consumed by soldiers in various environments, one problem identified with its use in this population is that it often does not capture the full amount eaten in one sitting of a particular food item. The images showing portion sizes are inadequate and potentially confusing for soldiers, although they take great care to complete the instrument with assistance available from the research team. We cannot be certain the instrument captured the full extent of macronutrients and micronutrients. Recommendations for modifying the instrument specifically for young military service members and athletes have been forwarded to the tool developers.

We did not have an appropriate measure of bone resorption after learning from the laboratory that tartrate-resistant acid phosphatase was no longer an available test. We were not made aware of this until after many of the baseline serum specimens had been collected and therefore we were unable to consider an alternative biomarker. This may have offered greater insight into the hypothesis that stress fracture injury may have contributed to the significant change in osteocalcin for the TG.

These data may not be representative of all combat arms soldiers because of our small sample size. Additionally, changes in physical performance during deployment may vary with the types and intensity of missions required of individual units. Our findings should be generalized cautiously as the study limitations are substantial and fail to support a broad impact on deploying military units. However, several studies have reported similar trends with regard to BF and physical activity despite differences in unit composition, location, and mission.5,6,11

### TABLE V. Bone Biomarkers, Reported as Mean (SE)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Baseline</th>
<th>Follow-Up</th>
<th>Change: Follow Up-baseline</th>
<th>FDR p</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGF-1 (ng/mL)</td>
<td>CG n = 135</td>
<td>TG n = 85</td>
<td>CG n = 135</td>
<td>TG n = 85</td>
</tr>
<tr>
<td></td>
<td>189.1 (5.25)</td>
<td>205.24 (6.72)</td>
<td>178.74 (4.74)</td>
<td>184.18 (6.17)</td>
</tr>
<tr>
<td>25(OH) Vitamin D (ng/mL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27.75 (0.7)</td>
<td>21.65 (0.86)</td>
<td>34.35 (0.94)</td>
<td>26.02 (1.11)</td>
</tr>
<tr>
<td>Calcium (mg)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.59 (0.03)</td>
<td>9.58 (0.04)</td>
<td>9.45 (0.03)</td>
<td>9.44 (0.04)</td>
</tr>
<tr>
<td>BS Alk Phos (mcg/L)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15.09 (0.53)</td>
<td>18.73 (0.67)</td>
<td>17.12 (0.72)</td>
<td>20.22 (0.86)</td>
</tr>
<tr>
<td>Osteocalcin (ng/mL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18.48 (0.62)</td>
<td>21.12 (0.77)</td>
<td>22.05 (0.99)</td>
<td>28.31 (1.08)</td>
</tr>
</tbody>
</table>

25OH Vit D, 25-hydroxyvitamin D; BS alk phos, bone specific alkaline phosphatase; CG, control group; FDR, False Discovery Rate; IGF-1, insulin-like growth factor -1; mcg/L, micrograms per liter; mg = milligrams; ng/mL, nanograms per milliliter; TG, telehealth group.

### CONCLUSION

This study is unique in that telehealth technology was employed to coach deployed soldiers to sustain or improve nutritional factors and exercise in order to support bone health. More studies are warranted to examine the effectiveness of nutrition and exercise interventions among active duty service members on bone health and prevention of musculoskeletal injuries.

### ACKNOWLEDGMENTS

This study was supported by the TriService Nursing Research Program (grant no. Hu0001-10-1-TS15, N10-C02), Bethesda, MD, during the study period August 2010–2014.

### REFERENCES

8. Office of the Deputy Assistant Secretary of Defense (Military community and Family Policy): 2012 Demographics Profile of the Military
Evaluation of Miniature Wireless Vital Signs Monitor in a Trauma Intensive Care Unit

Jonathan P. Meizoso, MD; Casey J. Allen, MD; Juliet J. Ray, MD; Robert M. Van Haren, MD, MSPH; Laura F. Teisch, BS; Xiomara Ruiz Baez, MD; Alan S. Livingstone, MD; Nicholas Namias, MD, MBA; Carl I. Schulman, MD, PhD, MSPH; Kenneth G. Proctor, PhD

ABSTRACT  A previous study demonstrated basic proof of principle of the value of a miniature wireless vital signs monitor (MWVSM, MiniMedic, Athena GTX, Des Moines, Iowa) for battlefield triage. However, there were unanswered questions related to sensor reliability and uncontrolled conditions in the prehospital environment. This study determined whether MWVSM sensors track vital signs and allow for appropriate triage compared to a gold standard bedside monitor in trauma patients. This was a prospective study in 59 trauma intensive care unit patients. Systolic blood pressure, temperature, heart rate (HR), skin temperature, and pulse oximetry (SpO₂) were displayed on a bedside monitor for 60 minutes. Shock index (SI) was calculated. A separate MWVSM monitor was attached to the forehead and finger of each patient. Data from each included pulse wave transit time (PWTT), temperature, HR, SpO₂, and a summary status termed “Murphy Factor” (MF), which ranges from 0 to 5. Patients were classified as “routine” if MF = 0 to 1 or SI = 0 to 0.7, “priority” if MF = 2 to 3 or SI = 0.7 to 0.9, and “critical” if MF = 4 to 5 or SI ≥ 0.9. Forehead and finger MWVSM HRs both differed from the monitor (both p < 0.001), but the few beats per minute differences were clinically insignificant. Differences in MWVSM SpO₂ (1–7%) and temperature (6–13°F) from the monitor were site specific (all p < 0.001). Forehead PWTT (271 ± 50 ms) was less (p < 0.001) than finger PWTT (315 ± 42 ms); both were dissociated from systolic blood pressure (r² < 0.05). The SI distributed patients about equally as “routine,” “priority,” and “critical,” whereas MF overtriaged to “routine” and undertriaged to “critical” for both sensors (all p < 0.001). Our findings suggest that MF does not accurately predict the most critical patients, likely because erroneous PWTT values confound MF calculations. MF and the MWVSM are promising, but require fine-tuning before deployment.

INTRODUCTION
The vast majority of military trauma deaths occur prehospital and a significant proportion of in-hospital mortality occurs in patients with no vital signs in the field.¹⁻³ Thus, large opportunities exist for reducing trauma mortality through innovation in the prehospital setting.⁴⁻⁵

The United States Special Operations Command (SOCOM) and the Department of Defense designed and funded the development of a miniature wireless vital signs monitor (MWVSM) (MiniMedic, Athena GTX, Des Moines, Iowa) based on the success of its original wireless vital signs monitor.⁶ The new system has reduced weight and size and addressed an unmet need to acquire vital physiologic information from small surface sensors placed on up to 5 casualties simultaneously and then to wirelessly transmit this data to miniature monitors carried by any first responder within a 100-m range. This would allow for appropriate monitoring, triage, prioritization of transport, and tracking of changes in multiple casualties on the battlefield or in other austere conditions. This particular system contains two novel features. First, it incorporates an injury acuity algorithm termed the Murphy Factor (MF), which summarizes overall patient status, and is calculated from available vital signs (whether measured by the sensor, input by the first responder, or calculated by the device); additionally, it factors in the changes in these vital signs over the last 30 seconds. MF includes an adjustment to overcome data drops that are common during triage in chaotic environments. Second, the system is based on pulse wave transit time (PWTT) rather than systolic blood pressure (SBP). Various methods of PWTT measurement correlate well with SBP.⁷

A previous study suggested that a single numeric MF from a peripheral MWVSM could predict the need for lifesaving interventions during prehospital transport.⁸ Although this prehospital study provided basic proof of concept, there were multiple sources of variability including patient selection, injury severity, sensor placement, transport time, and missing or intermittent data.

The present study was conducted in the controlled setting of a trauma intensive care unit (TICU) for a fixed period of time with a bedside monitor used as the gold standard for paired comparison of vital signs. Furthermore, the shock index (SI) was used a standard to compare the MF as both methods provide a single number to assess overall patient status. The SI, which is calculated by dividing the heart
Evaluation of MWVSM in a Trauma ICU

MATERIALS AND METHODS

Study Protocol and Patient Selection

This study was conducted at the Ryder Trauma Center (University of Miami/Jackson Memorial Hospital) and approved by the Institutional Review Board at Jackson Memorial Hospital and the University of Miami Leonard M. Miller School of Medicine, Miami, Florida. A prospective observational trial was performed in a convenience sample of 59 TICU patients from October 2013 to July 2014. Eligible patients were adults (>17 years old) with an arterial catheter who were admitted to the TICU. Patients meeting inclusion criteria were enrolled from 8 a.m. to 5 p.m. on Mondays to Fridays during the study period.

Data from each MWVSM were compared to data simultaneously obtained with a bedside vital signs monitor (GE Solar 8000M multichannel monitor, GE Healthcare, Milwaukee, Wisconsin) in the TICU. This monitor collected the following standard vital signs: HR, SBP, core body temperature as measured by Foley catheter, and pulse oximetry (SpO₂). The test system is composed of two components: a patient sensor that is affixed to either the forehead or an extremity and a handheld unit. The patient sensor weighs 3.8 oz and uses two AAA alkaline batteries. The handheld monitor is 4.0 × 2.5 × 0.79 inches in size, weighs 4.2 oz, and uses two AAA alkaline batteries. Communication between MWVSM devices is achieved with Communication Zigbee (802.15.4) Wireless Protocol with a range of 100-m line of sight.

The patient sensor records physiologic variables (skin temperature, SpO₂, HR, and PWTT) every second and transmits every 5 seconds to the monitor. Other information is manually input with the handheld unit, including Glasgow Coma Score, medications, and fluids. In addition to recording and storing continuous data, the sensor computes MF, which is a proprietary algorithm that incorporates changes in multiple parameters obtained from the MWVSM correlated favorably with conventional vital signs and SI in critically ill trauma patients.

Rate (HR) by the SBP, has previously been shown to correlate with markers of cellular hypoxia and shock, such as lactate and central venous oxygenation. It is also an indicator of patients requiring transfusion and a predictor of mortality. We test the hypothesis that changes in multiple parameters obtained from the MWVSM correlate favorably with conventional vital signs and SI in critically ill trauma patients.

Comparisons were made between the MWVSM and conventional monitor measurements for HR, SpO₂, and temperature. Correlations were made between PWTT measured by the MWVSM and SBP measured by the conventional monitor.

SI was calculated using the conventional bedside HR and SBP. To compare the two indices, MF was classified using manufacturer specifications, where “routine” is MF = 0 to 1, “priority” is MF = 2 to 3, and “critical” is MF = 4 to 5. Normal SI ranges from 0.5 to 0.7 and was classified as “routine” if SI = 0 to 0.7, “priority” if SI = 0.7 to 0.9, and “critical” if SI ≥0.9. The study team determined values for SI a priori. Prior investigators have found that SI >0.9 is associated with injuries requiring immediate attention, while others have used a similar breakdown to the present study.

Statistical Analyses

Data were analyzed using SPSS version 22.0 (IBM Corporation, Armonk, New York). Data are reported as mean ± SD if normally distributed or as median (interquartile range) if not normally distributed. Independent data were compared with paired Student’s t test. Categorical data were compared using χ² test or Fisher’s exact test as appropriate. Significance was assessed at p < 0.05.

RESULTS

The study population was comprised of 59 TICU patients, aged 47 ± 20 years with a male preponderance (80%). Mechanism of injury was blunt in 60% and penetrating in 27%; nontraumatic cases made up 13%. These basic demographics are similar to our usual TICU population.

Actual SBP was 129 ± 20 mm Hg. Corresponding PWTT with the MWVSM averaged 314 ± 2 ms with the finger sensor and 271 ± 2 ms with the forehead sensor. Forehead PWTT (271 ± 50 ms) was significantly less (p < 0.0001) than peripheral PWTT (315 ± 42 ms), but this is logical as it takes longer for the pulse wave to transmit to the periphery. However, peripheral and forehead PWTT were dissociated from SBP. For SBP vs. peripheral PWTT the correlation, although weak, was statistically significant (r² = 0.0385, p < 0.001). There was no significant correlation between SBP and forehead PWTT (r² = 0.0007, p = 0.513) (Fig. 1).

Actual core temperature was 100 (3.2) °F. Corresponding skin temperature with the MWVSM finger sensor was 86.8 ± 8.4°F and with the forehead sensor was 93.1 ± 4.7°F. Clearly these measures were dependent on sensor location. Paired temperature differences between core temperature and skin temperature (with both finger and forehead sensors) were statistically significant (p < 0.001). Furthermore, the paired differences between forehead and finger sensors (6–13°F) were statistically significant (p < 0.001).

200

MILITARY MEDICINE, Vol. 181, May Supplement 2016
Actual HR was 102 ± 18 beats per minute (bpm), compared to 105 ± 22 bpm with the finger sensor and 104 ± 21 bpm with the forehead sensor. Although the differences between conventional and MWVSM HR were statistically significant (p < 0.001), a difference of only 2 bpm cannot be considered clinically significant. Measurements between the forehead and finger sensors were similar (p = 0.35). Finally, median SpO₂ was 100 (2)%. This was compared to 100 (3)% with the finger sensor and 93 ± 8.3% with the forehead sensor. Similar to HR, conventional and MWVSM measurements were statistically (p < 0.001) different, but clinically insignificant. Paired differences between the finger and forehead sensors were also significantly different (p < 0.001).

Table I shows the differences between HR, SpO₂, and temperature using the bedside monitor and the MWVSM.

Figure 2 compares MF and SI in all three severity categories and demonstrates inappropriate triage by both forehead and peripheral sensors. According to the SI, the patients were distributed about equally in the “routine,” “priority,” and “critical” categories; however, MF significantly overtriaged patients to the “routine” category and undertriaged patients to the “critical” category in both sensors (all p < 0.001).

**DISCUSSION**

The present study is the second from our group that evaluates the MWVSM for potential trauma triage. The major new findings are that peripheral and forehead PWTT were dissociated from SBP and that relative to SI, MF significantly overtriaged patients to the “routine” category and undertriaged patients to the “critical” category for either the forehead or peripheral sensor. These observations support the conclusion that MF does not accurately predict the most critical patients likely because erroneous PWTT values confound the MF calculation. The concept of MF and the MWVSM is promising, but requires further fine-tuning before deployment.

The military WVSMs before this miniature version were smaller than conventional monitors, but still bulky and difficult to transport for SOCOM. The MWVSM offers a (theoretical) logistic advantage for continuous monitoring of multiple casualties on the battlefield from remote locations. We demonstrated basic proof of concept that the MF can summarize overall patient status in the prehospital setting. This follow-up study aimed to critically evaluate the MWVSM in the controlled setting of a TICU to demonstrate whether data obtained from the MWVSM correlates with the conventional vital signs monitor. Furthermore, we aimed to correlate MF with SI, a previously validated measure that is able to accurately predict early mortality and need for massive transfusion after blunt trauma.

MWVSM HR is reasonably accurate. Although the values are statistically significant, a 2 to 3 bpm difference is clinically insignificant. Data from the peripheral MWVSM sensor, but not the forehead sensor, agrees with conventional SpO₂; this can be expected secondary to different perfusion patterns in the forehead and periphery. However, skin temperature from either sensor differs from the actual core temperature. It is unclear what, if any, additional vital information is provided by measuring skin temperature from either site.

The most significant finding is that PWTT is poorly correlated with SBP regardless of sensor placement. We believe that the inaccuracies in triaging patients using the MF are secondary to this discrepancy in PWTT as the MF uses
PWTT in its summary status alarm algorithm. Unfortunately, the algorithm for calculating the MF is proprietary and we are unaware of the weight that PWTT carries in the calculation. MF derived from both sensors overtriaged patients to the “routine” category and undertriaged patients to the “critical” category; overall, MF inappropriately triaged approximately 50% of patients relative to SI.

The MWVSM was developed for use in chaotic military mass casualty environments when triage and transport decisions are difficult. MF is an overall summary status alarm that would theoretically aid the combat medic to make life and death decisions. The MWVSM will never replace a well-trained medic, but it can provide useful information under times of stress. Woodford et al15 have also previously demonstrated the benefit of an automated continuous vital signs analysis system with decision support capabilities in prehospital care. Furthermore, a significant amount of work regarding WVSM and machine learning has been done by the U.S. Army Institute of Surgical Research. Liu and Holcomb studied the WVSM, which is the predecessor to the MWVSM, in 305 patients and found that the WVSM improved accuracy in identifying patients requiring lifesaving intervention.6 The same group has used automated systems incorporating nonstandard vital signs, including HR variability (HRV) and HR complexity (HRC), and found that these improve the prediction of mortality and lifesaving interventions in trauma patients.16,17

Several other noninvasive strategies have been proposed to triage trauma patients.6,15,16,18–21 King et al18 prospectively analyzed 75 trauma patients requiring prehospital helicopter transport and found that HRV predicted base excess and the requirement of lifesaving interventions; HRV more accurately identified critical patients than trauma center criteria or prehospital vital signs. Indeed, the current “standard” vital signs measures used to identify patients in shock have been questioned; Eastridge used the Joint Theater Trauma Registry to study 7,180 military combat casualties and found that SBP ≤ 100 mm Hg is a better indicator of true hypotension and hypoperfusion in military trauma than the traditional 90 mm Hg.22 Ryan et al19 found HRV to be a predictor of morbidity and mortality in the hemodynamically stable trauma patient. HRC also analyzes HR

<table>
<thead>
<tr>
<th>Table I. Summary of Conventional and miniature wireless vital signs monitor (MWVSM) Vital Signs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Heart Rate (bpm)</strong></td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>Bedside Monitor</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Peripheral MWVSM Sensor</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Paired Difference Sensor vs. Monitor</td>
</tr>
<tr>
<td>Forehead MWVSM Sensor</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Paired Difference Sensor vs. Monitor</td>
</tr>
<tr>
<td>Paired Difference Sensor vs. Sensor</td>
</tr>
</tbody>
</table>

**FIGURE 2.** (A) Percentage of patients with peripheral sensor Murphy Factor (MF) vs. Shock Index (SI) in each acuity category. (B) Percentage of patients with forehead sensor MF vs. SI in each acuity category.
time sequences and has been found by several groups to correlate with the need for lifesaving interventions in the prehospital setting. Continuous near-infrared spectroscopy can predict the need for transfusion in patients deemed hemodynamically stable with an SBP >90 mm Hg and may prove to be an excellent adjunct to physiologic monitoring in the future.

Limitations to this study include the use of a convenience sample, which is not randomized and thus subject to bias. No interventions were performed based on results from the MWVSM. Furthermore, we do not have data regarding whether any of these patients required blood product transfusion or other interventions. The MF depends on the reliability of input signals and these are sometimes lost when patients are in extremis; in this present study, some data were lost even when patients were not in extremis. If the overall purpose of the MWVSM is to identify those patients who are in shock (i.e., cold, clammy, hypotensive, and tachycardic) then the device needs further fine-tuning to avoid data drops. Additionally, the MWVSM is designed to collect continuous data however the data analyzed included only snapshots at every 5 minutes compared to the conventional monitor. Our previous study demonstrated that a MF >3 in the prehospital setting was associated with the need for lifesaving intervention upon arrival to the trauma center. The current findings suggest that MF is not a reliable indicator of shock because it undertriages a proportion of the most severely injured patients. Together, this suggests that a significant number of patients could be misclassified by MF.

In summary, in a TICU population at an urban level I trauma center, the MWVSM demonstrated poor correlation between PWTT and SBP. We believe this was responsible for the inaccurate triage using the MF when compared to the previously validated SI. Improvements in measuring PWTT are required to obtain more accurate MF values. The use of noninvasive blood pressure rather than PWTT may provide another option for improved identification of patients in shock using the MWVSM. Newer modalities such as HRV, HRC, and near-infrared spectroscopy will likely be beneficial in future wireless vital signs devices for more accurate and timely diagnosis and treatment of combat casualties. Further work is required to validate the MWVSM for use by SOCOM in the battlefield.

ACKNOWLEDGMENTS

This work was supported in part by the Office of Naval Research (Grant No. N140610670) and the U.S. Army Medical Research & Materiel Command (Grant No. 09078015).

JPM was directly responsible for all aspects of this study. He participated in the collection, analysis, and interpretation of data, and drafting and revision of the manuscript, figures, and tables. CJA, JR, RMVH, LFT, and XRB participated in the experimental design, collection of data, and revision of the manuscript, figures, and tables. ASL, NN, and CIS were medically responsible for the patients and participated in the review and revision of the manuscript, figures, and tables. KGP had overall responsibility for the study, including conception and experimental design; analysis and interpretation of the data; drafting and revision of the manuscript, figures, and tables; statistical expertise and evaluation; obtaining funding for this project; and supervision.

REFERENCES


Developing a Cognitive and Communications Tool for Burn Intensive Care Unit Clinicians

Christopher Nemeth, PhD†; Shilo Anders, PhD†; Robert Strouse, MFA†; Anna Grome, MS†; Beth Crandall, BA†; Jeremy Pamplin, MD†; Jose Salinas, PhD†; Elizabeth Mann-Salinas, PhD†

ABSTRACT Background: Burn Intensive Care Unit (BICU) work is necessarily complex and depends on clinician actions, resources, and variable patient responses to interventions. Clinicians use large volumes of data that are condensed in time, but separated across resources, to care for patients. Correctly designed health information technology (IT) systems may help clinicians to treat these patients more efficiently, accurately, and reliably. We report on a 3-year project to design and develop an ecologically valid IT system for use in a military BICU. Methods: We use a mixed methods Cognitive Systems Engineering approach for research and development. Observations, interviews, artifact analysis, survey, and thematic analysis methods were used to reveal underlying factors that mold the work environment and affect clinician decisions that may affect patient outcomes. Participatory design and prototyping methods have been used to develop solutions. Results: We developed 39 requirements for the IT system and used them to create three use cases to help developers better understand how the system might support clinician work to develop interface prototypes. We also incorporated data mining functions that offer the potential to aid clinicians by recognizing patterns of clinically significant events, such as incipient sepsis. The gaps between information sources and accurate, reliable, and efficient clinical decision that we have identified will enable us to create scenarios to evaluate prototype systems with BICU clinicians, to develop increasingly improved designs, and to measure outcomes. Conclusion: The link from data to analyses, requirements, prototypes, and their evaluation ensures that the solution will reflect and support work in the BICU as it actually occurs, improving staff efficiency and patient care quality.

BACKGROUND

Patients who are admitted to the Burn Intensive Care Unit (BICU) present health care teams with unique challenges as a result of their fragile and often unstable condition. Their complex combinations of life-threatening injuries and illnesses make trauma and surgical care for these patients necessarily complex. Clinicians from 15 specialties must work together to make effective decisions, develop treatment plans, assess patient progress, and refine care management over time. This team must also account for limited resources and must adjust their course of treatment according to variable patient responses to interventions.

Care also relies on clinician cognitive work, which includes decision-making and related activities such as problem detection, sense making, and building common ground among the care team members. Under time pressure, intensive care unit clinicians must rely on a large volume of data that is separated among multiple sources. The decisions clinicians make are only as good as the information that is available and important (salient) when the decisions are made. Because of this, the Institute of Medicine recommended improving access to accurate, timely information, and making relevant information available at the point of patient care.

Research and development for this project is being conducted by Applied Research Associates, Inc., an 1,100-member science and engineering consulting firm, which is creating a decision and communications support system that will serve a 16-bed military tertiary care BICU. This Cooperative Communication System (CCS) is expected to enable the health care team to remain connected to information about each patient and to each other across time and location as the team delivers care. The CCS will keep providers informed of a patient’s status, and of other health care providers’ patient care activities, enable the staff to understand goals, objectives and tasks related to each patient, and to reconcile differing points of view. Its decision and communication support and machine learning features will make it possible for clinicians to make more accurate and timely diagnoses, to perform more timely and appropriate tests, and to make better plans to optimize patient care. Use of the CCS is expected to improve the availability of information and the synchronization of care among BICU team members, which in turn are expected to improve patient outcomes.

This article describes rigorous field study, analysis, requirements, and information design and programming to design and develop an ecologically valid information technology (IT) system.
METHODS
The CCS research team is using a mixed methods Cognitive Systems Engineering\textsuperscript{2,3} (CSE) approach for this study. The CSE approach includes methods that are particularly well suited to both learn about behavior and cognition as humans confront complexity in work settings such as the BICU and to develop tools to support their cognitive work. The approach translates knowledge about human cognitive performance to develop solutions, including information system interface design.\textsuperscript{4} In this study, knowledge that clinicians need includes vital signs and laboratory values that one would expect would matter in trauma and surgical care decision-making. Knowledge also includes unexpected data patterns that matter, but are difficult to detect.

As a “systems engineering” methodology, the CSE approach includes all of the agents that can act in the work setting: clinician and support staff, tasks, information sources, the facility, and more. Figure 1 illustrates five phases in the approach and how the activities in each phase relate to phases of this project. As Figure 1 shows, CSE phases include data collection, data analysis, and solution development. Integration of these five phases ensures that the solution the CSE process produces is inherently valid by being grounded in worker and work setting data. Each element in the solution that the CSE approach produces can be traced back through requirements, through analyses, to the original data. The ability to identify each element among workers, work setting, and tools can also help designers to anticipate shifts and unintended consequences that can happen when new IT such as the CCS is introduced.\textsuperscript{5} The CSE approach has been proven to successfully study cognitive activity in complex field settings in high-hazard sectors such as defense, national security, nuclear power plants, and law enforcement. The project team has recently used CSE to perform work on behalf of the Department of the Army,\textsuperscript{6,7} Chief of Naval Operations (Nemeth C, Wiggins S, Crandall B, et al: C2 Upgrade for NECC Branch [OPNAV N857]. Contract N00024-10-C-6309. Washington, DC: Department of the Navy. 2011), Office of Naval Research (Anderson KR, Crandall B, Grome A, Nemeth C: Environmental and ship motion forecasting cognitive aid investigation: Decision and information requirements development. Contract N00014-11-C-0360. Washington, DC: Office of Naval Research. 2014), and Department of Homeland Security (Nemeth C, Grome A, Laufersweiler D, Crandall B, Strouse R: A research roadmap to improve screening performance through cognitive systems engineering. Contract GS-10F-0298K. Washington, DC: Department of Homeland Security, 2013).

Our project team studied clinicians who work in a 16-bed, American Burn Association accredited regional referral burn center that is a part of a 450 bed, academic, military, level I trauma center. The team obtained approval for human subject research from the funder and research site institutional review board and obtained informed consent from all participants.

In Year 1, the research team used data collection methods (observations, interviews, surveys, and artifact analysis) to go beyond surface descriptions (phenotypes) that revealed underlying patterns (genotypes) of systemic factors that mold the work environment and affect clinician decisions.

Data Collection
A team of 2 to 4 researchers made 4 week-long data collection visits to the research site, and coordinated additional collection with an on-site research nurse between visits.

FIGURE 1. Five phases of cognitive systems engineering.
During these visits, they performed the following data collection methods:

Observation of clinical teams as they provided patient care and managed the unit. Team members conducted 31 observations with the BICU staff, including bedside, charge and wound care nurses, residents, attending physicians, and physical, occupational, and respiratory therapists. These sessions involved shadowing a single person and asking them to talk aloud as they completed their work. Use of probe questions enabled researchers to request background and clarifying information in context to better understand motivations, information use, and decision-making.

Forty-nine semi-structured cognitive task analysis (CTA) interviews lasting between 30 to 90 minutes each with members of the BICU clinical staff elicited knowledge about their background, perspectives, work activity, information sources, and challenges they face.

Artifact analysis of computer-based and hard copy information sources that clinicians use in their work, including sign-out sheets, personal notes, status boards, and information system and equipment displays.

Brief surveys to identify patterns, such as work team relationships (usually conducted by the on-site research nurse in-between research team visits).

Data Analysis

The research team analyzed data collected from four week-long site visits and research nurse support at the site between visits. Through the following eight steps (Fig. 2), their analyses identified clinician goals and barriers to goal achievement. After the first site visit, the team performed an initial data review and extraction of emerging themes to review and analyze interview and observation notes. Following the second site visit, the team conducted a systematic data review and coding to reveal thematic categories developed during working sessions, and code interview sections to relate them to each theme. After the third visit, they reviewed and interpreted coded data, synthesized and merged findings, and reflected on newly collected data. Each of the steps used analyzed cognitive work to provide the basis for analyzing the cognitive work requirements of BICU clinical teams and distil a descriptive model, as well as artifact analysis of the forms and documents that the BICU clinical teams use, to more fully understand the kinds of information they seek, use, and share with one another. After developing initial requirements for the CCS, team members made another visit to present the challenges/barriers and initial requirements to a select set of BICU clinicians to obtain an initial appraisal of the findings by verifying accuracy and identifying possible gaps. The team used results from the data analysis to identify barriers to cognitive work, and develop final requirements for the CCS that would enable BICU staff to overcome those barriers.

Participatory Design

Research, software development, and machine learning team members met with the clinical co-principal investigator (JP) for a 2-day data analysis and design session to refine and revise design requirements. The team also held a similar design session a few weeks later at the research site to capture clinician insights. In these sessions, representatives from all of the clinician groups that work in the BICU proposed system design ideas that might facilitate timely, effective, and efficient patient care. The sessions provided the interface designer with beginning concepts for further development and refinement. The research team also updated and

![FIGURE 2. Data collection and analysis process.](image-url)
refined the use cases that the software development team would need.

**RESULTS**

One hundred fifty-one BICU clinicians and staff members representing all unit roles consented to participate in this research and many were subject of interviews and observation. Roles included attending physician (surgeon, intensivist), fellow, resident, physician assistant, respiratory therapist, occupational/rehabilitation therapist, wound care specialist, dietician, bedside nurse (registered nurse, licensed vocational nurse), unit nurse (e.g., infection control), care manager, ward clerk, chaplain, volunteer, other physician (e.g., anesthesiologist, consulting physician), ancillary services, and student (medical, nursing). Members of this sample and each of the roles also participated in design workshops.

Year 1 results showed that the IT solutions that are currently available to BICU clinicians are not sufficient for clinician information needs. This is because current solutions do not help clinicians to efficiently drive down uncertainty at the individual and the team level. This compels clinicians to exert cognitive effort find and model information that is stored within and across multiple health IT systems to make decisions. To counter this, we identified 21 barriers to effective clinical care and recommending 39 requirements for the CCS prototype (see Table I). These requirements were further developed into rough, then increasingly refined, information displays through creative design workgroups and repeated interviews and surveys. Data analysis identified problems that current health IT solutions present, 21 barriers to cognitive work on the BICU, and developed 39 CCS requirements.

**The Problem**

The following examples demonstrate difficulties using current health care IT, such as finding important (salient) information that the CCS is intended to address:

Example 1

Patient on insulin drip (which is tracked on the medication flow sheet and the in/out flow sheet) but the patient was not getting hourly blood glucose measurements (which are tracked on the labs and vital signs flow sheets). Small example, but the patient’s blood glucose on re-check after six hours was <30.

Example 2

Ok, I’m trying to identify what possible new medication might have caused a patient’s liver to start to fail (this same scenario could apply to any system). There is NO way for me to organize the data in such a way that I can see: Vital signs, Labs, Medication at the SAME time. I must do this manually. This is true in [commercial IT system] too. We should be able to do this, especially if we can assign a medication to a system, and potentially unassigned it.

Example 3

Ordered a right upper quadrant ultrasound yesterday. Turns out, the patient had several of these in the past, not necessarily in the last month (last was in July), all with similar results—difficult to see the gallbladder. [We] did a different study today. Probably would have saved at least the cost of the procedure yesterday had I known this . . . .

**Barriers and Requirements**

Each of the barriers that the team discovered presents an opportunity to learn how the CCS can support better care coordination. Using the barriers, the team created requirements for the CCS that would enable clinicians to overcome them (Table I). The first barrier provides an example:

No effective means to synchronize and adapt different aspects of patient care over the course of a shift, across caregiver team.

The requirement states how the CCS solution can help to overcome the barrier.

*System shall provide access to a plan of patient care, visible to all caregivers responsible for that patient, that includes:*

(a). Current patient status and top-level assessment.
(b). Goals and priorities for those goals.
(c). Changes/updates, such as indication that plan is being updated when one caregiver is working on it.
(d). Schedule of activities and any changes, timeline.
(e). Orders and their status.
(f). Identity and contact information for patient’s care team.

The collection of requirements supports development of a number of use cases. They also guide the interface designer’s configuration of display content and layout, and software developers planning for interactive features.

**Use Case**

A use case is a narrative description that suggests how a system might be used. By assembling requirements into a description, software developers can get a sense of how the system will operate to support cognitive work on the unit. The first paragraph of a use case for access to a patient care plan that was described above, describes how each of these features (shown in bold type) would serve clinician needs.

At 6:30 a.m., a bedside nurse has started his preparation for the day shift by reviewing information on the patient he is responsible for. Opening CCS, he can see a roster of patients on the unit, chooses his patient’s “at-a-glance” view that shows recent vital signs, current orders, medications, care plan, and notes from the night shift. He checks the patient’s standing care plan and treatment goals (from the electronic health care record), and reviews orders (from
### Problem/Barrier

- No effective means to synchronize and adapt different aspects of patient care over the course of a shift (e.g., among RN, OT/PT, wound care)
- Lack of awareness around activities/events that are tightly coupled
- No efficient communication of patient status change across disciplines
- Documentation requires significant time from key members of the clinical team (RNs, Residents, RTs etc.) and is often redundant
- Updated information is available but not readily accessible or visible to clinicians (e.g., cultures)
- Orders late, missing, or overtaken/replaced by other orders
- Reliance on verbal orders and no standardized way to share orders
- Trends are important information, but cannot get them from Essentris or other IT
- No ability to keep track of patient status over time > 24 hours
- How many clinical staff are currently on the unit?
- Is patient ready for upcoming surgical procedure
- OR RN does not know enough about upcoming procedure to prepare surgical suite properly
- Bedside RN does not know enough about surgery as it is being performed to prepare properly for patient’s return
- Rounding Checklist not readily available/accessible to all members of clinical team
- Impact of dropped tasks, gaps, and lapses not known or tracked
- Checklist management is unclear (responsibility for making sure items are completed is unclear).
- Reliance on clinician to mentally integrate data

### Needs/Requirements

- Need to determine optimal timing and sequence of activities
- Need awareness of planned/scheduled patient care activities (e.g., wound care, rehab, line changes, etc.)
- Means to share the plan
- Means to adapt the plan in real time and share changes across the team.
- Bedside nurse needs to shift the goals and priorities
- Means to know how changes in orders affect/change planned activities
- Means to know what planned events are and who needs to be there
- Practitioners need to understand what’s going on with their group of patients across the shift (whatever their group happens to be)
- Means to post checklist so all staff have ready/easy access
- Clinicians need to be aware that updated information is available, particularly regarding laboratory cultures
- Need efficient, accurate way to specify meds, procedures
- Physicians need access to orders from charge nurse’s checklist
- Physicians need prompts to enter orders
- Need indicator of status of order entry (has it been placed or not?)
- Need indicator of status of order (in process, completed)
- Physicians need to be aware when entering order that it is different from previously entered orders
- Changes to orders need to be disseminated to wider team so that team has common ground. Changes in orders need to be apparent to whole team
- Information Management tools and processes built around efficient use of staff time and effort
- Minimize staff time required to capture information by reducing redundant information gathering and entry
- Minimize staff time spent as the “system integrators” who move data from one system to another
- Need “user-friendly” interfaces/systems
- Means to indicate if patient is highly unstable (because information for unstable patients can become inaccurate in short time frame)
- Means to know whether information in system is up-to-date (e.g., is this an accurate reflection of the patient’s status right now?)
- Means to know whether orders are in process but results not entered into system yet (e.g., cultures, laboratory results)
- Means to know recency of information updates
- Means to capture and disseminate changes to orders that occur verbally within subteams
- Clinicians need trend information
- Need view of patient that is more than just this shift. Both macro level view of indicators and over longer time spans
- Need to know who is available, and where to find them
- Need access to nurse assignments by shift, by patient
- Means to access assistance, guidance, decision makers
- Need to know which specialty is assigned to each patient (e.g., RT) and patient acuity
- Need means to know whether patient is prepared for procedure (have they gotten blood products, antibiotics, consent, pregnancy test)
- OR nurse needs procedure specific description (need to know more about specific information needs)
- Bedside Nurse needs means to know what to expect re patient needs following procedure (e.g., what was worked on, how much blood given or lost, sedation?)
- Means to construct checklist in real time (during Rounds) or immediately after
- Means to post checklist so all staff have ready/easy access
- Means for incomplete items to “roll over” to populate next day’s check list and to be reviewed at next day Rounds
- Clinicians need a holistic/macro-view of the patient’s trajectory (e.g., are they getting better or getting worse over last 24 hours?)

---

**TABLE I.** Barriers and Requirements

<table>
<thead>
<tr>
<th>Problem/Barrier</th>
<th>Needs/Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Documentation requires significant time from key members of the clinical team (RNs, Residents, RTs etc.) and is often redundant</td>
<td>Clinicians need a holistic/macro-view of the patient’s trajectory (e.g., are they getting better or getting worse over last 24 hours?)</td>
</tr>
<tr>
<td>No ability to keep track of patient status over time &gt; 24 hours</td>
<td>Need to know who is available, and where to find them</td>
</tr>
<tr>
<td>How many clinical staff are currently on the unit?</td>
<td>Need access to nurse assignments by shift, by patient</td>
</tr>
<tr>
<td>Is patient ready for upcoming surgical procedure</td>
<td>Need to know which specialty is assigned to each patient (e.g., RT) and patient acuity</td>
</tr>
<tr>
<td>OR RN does not know enough about upcoming procedure to prepare surgical suite properly</td>
<td>OR nurse needs procedure specific description (need to know more about specific information needs)</td>
</tr>
<tr>
<td>Bedside RN does not know enough about surgery as it is being performed to prepare properly for patient’s return</td>
<td>Bedside Nurse needs means to know what to expect re patient needs following procedure (e.g., what was worked on, how much blood given or lost, sedation?)</td>
</tr>
<tr>
<td>Rounding Checklist not readily available/accessible to all members of clinical team</td>
<td>Means to construct checklist in real time (during Rounds) or immediately after</td>
</tr>
<tr>
<td>Impact of dropped tasks, gaps, and lapses not known or tracked</td>
<td>Means to post checklist so all staff have ready/easy access</td>
</tr>
<tr>
<td>Checklist management is unclear (responsibility for making sure items are completed is unclear).</td>
<td>Means for incomplete items to “roll over” to populate next day’s check list and to be reviewed at next day Rounds</td>
</tr>
<tr>
<td>Reliance on clinician to mentally integrate data</td>
<td>Clinicians need a holistic/macro-view of the patient’s trajectory (e.g., are they getting better or getting worse over last 24 hours?)</td>
</tr>
</tbody>
</table>
the laboratory test database) that are pending as well as the day’s care activities that the Wound Care team, Respiratory Therapists, and Physical Therapists have recommended and what times they can perform them.

Both software development and machine learning team members are using these requirements and use cases to develop, evaluate, and refine interface prototypes.

After translating analysis findings into concise problem statements and information system requirements, the team developed a number of visual representations to describe BICU cognitive work and key resources that clinicians use (model of cognitive work, care team, and information sources) and prototype information displays.

**Model of Cognitive Work**

Complexity can hide underlying systematic patterns of cognitive work that clinicians perform in the BICU. Figure 3 illustrates these patterns that our CSE approach revealed.

The top level of the model (at left) shows the unit’s primary role in cognitive work: synchronization of patient care both among clinicians and over time. The next level down includes activities that all unit members perform to accomplish synchronization: clarification, coordination, negotiation, and anticipation. Supporting tasks make each of those activities possible. Each task can be observed in the way that clinicians interact with each other and use information sources to minimize uncertainty. Requirements that the team developed from these tasks indicate possible leverage points, or opportunities, to improve synchronization.

**Patient Care Providers**

Knowing what to include and exclude is part of the challenge in the study of a complex system such as the BICU. To do that, the team asked 8 nurses, 5 respiratory therapists, 2 physical therapists/occupational therapists, 1 nutritionist, and 1 physician on the BICU “Who do you communicate with to do your work?” The resulting network is being used to guide development of role-specific screens in the prototype versions of the CCS.

**Information Sources**

Artifact analysis developed an inventory of the information sources shown in Figure 4 that clinicians rely on to provide patient care. Sources ranged from physical items (e.g., status boards) to communications (e.g., cell phones) to computer databases (e.g., the electronic health record) and paper and electronic sources (e.g., arterial blood gas monitor). Disconnection among most of these sources was one of the barriers the team’s inquiry revealed. The need for clinicians to transcribe and reenter data from one system to another detracts from time to care for patients, and also presents the opportunity for inaccurate transcriptions.

**Information Displays**

Based on the participatory design sessions, the design team developed several versions of the interface design. This resulted in an information design prototype that was based on Year 1 findings and requirements with views organized according to clinician needs.
Patient View (Fig. 5)

Makes salient information evident by showing critical variables for each patient organized by neural, cardiac, respiratory, gastrointestinal, pulmonary, and renal systems. A “parent-child” display tab feature serves as a kind of tab reference to see more detailed material. The view also includes a Wound Flow analysis of the patient’s skin and graft condition (developed by the research site), as well as the patient’s schedule for the day.

Multidisciplinary Rounds View

Provides a means for the charge nurse to document key details of the daily interdisciplinary rounds that are conducted each morning starting at 8:00 a.m. Entry of goals, medications, and orders captures patient care decisions, put them in motion, and makes it possible to track their progress through the day.

Unit Level View

Indicates the location and condition of each patient in the 16-bed unit, and the two operating rooms nearby. Provides a message window to share information that affects the whole unit, and staff members on the unit that shift.

DISCUSSION

Health care IT systems must reflect actual clinical practice to provide information that will effectively support decision-making and related cognitive work of patient care. We have shown how the CSE research approach can be used to identify barriers to decision-making, and develop potential solutions to overcome them.

Despite years of effort in medical informatics, a gap remains between the complexities of the clinical work setting and the information systems that are intended to support clinician cognitive work. This is true of the electronic health record (EHR) as well as other health care IT such as Computer Physician Order Entry. The difference has implications for clinician performance and, ultimately, patient care. The examples in the Problem section of this article demonstrate how a clinician’s inability to find salient information affects clinical decision-making. We contend that the reason for this is a failure to accurately reflect the work domain and behavior in the clinical setting.

During this research we have studied individual and team clinician work in actual and controlled settings. Among the findings mentioned above, we have also found issues with health care IT displays, including the EHR. The EHR is intended to serve as the central information source for clinicians to use while making patient care decisions. EHRs are often linked with other systems, including clinical decision support, and computerized physician order entry. Applications such as dispensing medications can also include interaction with other systems such as bar coding at medication dispensing, robot for medication dispensing, and automated dispensing machines. Administrative applications include electronic medication administration records and bar coding at medication administration. These interrelationships can have a widespread effect on the work that clinicians perform.

Clinician patient care decisions are based on information that is provided by various means, which increasingly include the EHR. While providing some benefits, the EHR’s rapid development has created “. . . digital piles grown so gigantic, unwieldy, and unreadable that sometimes we wind up working with no information at all.” Among all of these
data, where does the clinician look for what matters when assessing trends and making diagnostic and therapeutic decisions? Do data that matter stand out, or are they obscured by other elements? And how can system developers know what matters? What data matter most to a patient and clinician at the moment they are being considered? Machine learning features, we are including in the CCS can be used sort through the “digital” piles to make useful information salient (stands out or is prominent).

Automation has traditionally been employed in high-hazard settings to replace individuals in the performance of work that is considered to be inappropriate for humans. Rather than replace humans, though, automation needs to aid humans as they work to solve problems. The way that a problem is presented can improve or degrade the performance of cognitive work13 and aiding has typically been directed at the novice level. In fact, aiding is most needed on difficult problems, which are the type of problems that experts confront. As in other high hazard settings, expertise14 in health care is the ability to know what is—and what is not—important.

Health care activities rely on the acquisition, portrayal and analysis of therapeutic and diagnostic information as an integral part of individual patient care. The daily work of the clinician requires representations that serve as a map of the ever-changing territory of work that must be successfully navigated.15 What is represented, and how it is represented, depends on the individual and group cognitive work that it is intended to support? Individual elements of information vary enormously in the length of time that they remain reliable, and their weight depends a great deal on their context. The need for accurate, timely information also exists at the unit level, such as the operating room and intensive care unit, where the technical work of unit planning and management directs who will get care, what type of care will be provided, and when it will be provided.

Progress in improving health care IT to support patient care relies on going beyond the surface descriptions (phenotypes) of work domains to the underlying patterns (genotypes) of systemic factors.16 Understanding any work domain and the forces that shape it requires methods that are suited to their study. Human factors17 and CSE research methods within the naturalistic decision making model18 have proven value in revealing the key aspects of health care work domains such as the BICU in this study to develop valid information displays.

Improvement in IT support for health care cognitive work requires repeated, deep looks into the clinical work setting using methods that are suited to the study of individual and team cognitive work to find what data truly matter. Use of CSE’s decision-making approach to understand patient care settings can inform the development of effective IT support. The salience that results can begin to overcome embedded difficulties with records that, left unattended, will continue to impede clinical care for patients.

As a BICU IT system, CCS is a Force Protection resource to provide optimal support for military patients. Through CCS decision support, clinicians can make more accurate and timely diagnoses, perform more timely and appropriate treatments, and provide evidence-based care that reduces the time lag from “bench-to-bedside” care. As a team tool, CCS builds consensus and efficiency that can be expected to shorten patient length of stay and improve outcomes.

As a networked system, the CCS has the potential to extend beyond the fixed walls of a hospital to incorporate prehospital, contingency operations, and theater evacuations during military operations. Improved communication, the CCS affords, also facilitates hand off on arrival at the care facility.
CONCLUSION

The findings from our CSE study are being used to create an information display that presents salient information, which will spare clinicians from having to find and synthesize it as they do now. This is expected to improve staff efficiency and patient care quality by improving clinician decision-making and communication. Specific CCS views sort information according to BICU cognitive work, from preparing for and conducting rounds, to individual patient care, to managing the unit as a whole. The link from data to analyses, requirements, prototypes, and evaluation ensures that the CCS solution will reflect and support work in the BICU as it actually occurs.

The research team’s prototype, which can also mine data for relevant information, will be tested and validated using criteria from the first year of research. Use of the CCS is eventually expected to help to decrease missteps, lapses, delays in care, and the morbidities from causes such as wrong medication/dose, infections, and unanticipated emergencies such as cardiac arrest.

ACKNOWLEDGMENTS

The authors express their gratitude to the clinicians who participated in this effort, LTC Kevin Chung, MD, as well as colleagues Greg Rule, Dianne Hancock, LVN, and Nicole Caldwell, RN, for their work on this project. We are also grateful to Dr. Michael O’Connor for contributing insights on salience and the electronic health record. This work is supported by the U.S. Army Medical Research and Materiel Command under contract no. W81XWH-12-C-0126.

CN (the principal investigator, lead author) contributed in data collection/analysis and interface design; JP (the co-principal investigator) is the clinical subject matter expert; SA and AG contributed in data collection/analysis and requirements development; RS contributed in interface design; BC contributed in data analysis, requirements development; JS (the task area manager) contributed in software development standards; EMS (the task area manager) contributed in research standards and is the clinical subject matter expert.

REFERENCES

Adaptive Virtual Reality Training to Optimize Military Medical Skills Acquisition and Retention

Ka-Chun Siu, PhD*; Bradley J. Best, PhD†; Jong Wook Kim, PhD‡; Dmitry Oleynikov, MD§; Frank E. Ritter, PhD¶

ABSTRACT The Department of Defense has pursued the integration of virtual reality simulation into medical training and applications to fulfill the need to train 100,000 military health care personnel annually. Medical personnel transitions, both when entering an operational area and returning to the civilian theater, are characterized by the need to rapidly reacquire skills that are essential but have decayed through disuse or infrequent use. Improved efficiency in reacquiring such skills is critical to avoid the likelihood of mistakes that may result in mortality and morbidity. We focus here on a study testing a theory of how the skills required for minimally invasive surgery for military surgeons are learned and retained. Our adaptive virtual reality surgical training system will incorporate an intelligent mechanism for tracking performance that will recognize skill deficiencies and generate an optimal adaptive training schedule. Our design is modeling skill acquisition based on a skill retention theory. The complexity of appropriate training tasks is adjusted according to the level of retention and/or surgical experience. Based on preliminary work, our system will improve the capability to interactively assess the level of skills learning and decay, optimizes skill relearning across levels of surgical experience, and positively impact skill maintenance. Our system could eventually reduce mortality and morbidity by providing trainees with the reexperience they need to help make a transition between operating theaters. This article reports some data that will support adaptive tutoring of minimally invasive surgery and similar surgical skills.

INTRODUCTION

About 4,300 physicians of the U.S. Army Medical Command continuously rotate through deployments across primary care, combat casualty care, and host nation care, with 2,800 individual deployments and an average deployment of 113 days. (Buller JL, Presentation given at the Medicine Meets Virtual Reality meeting, February 2011). The nature of required skills varies dramatically by deployment. For example, in-theater care for high-velocity wounds often requires procedures such as debridement, cauterization, and ligation, whereas usual surgical care in the civilian setting emphasizes procedures such as laparoscopic cholecystectomy and hernia repair.

This constant shift of required skills confronts surgeons with the need for skills different than those they are employing before deployment, but that they must somehow train or retrain to expertise before use. Although they are deployed, their previously sharp skills required in other theaters then may decay through disuse unless they are able to somehow train those skills as well. The enormous challenge posed by this problem is to understand and quantify the nature of surgical skill decay and develop a set of methodologies for training interventions to prevent that decay that minimizes training time, maximizes efficacy, and reduces mistakes during the initial portion of deployments.

The Department of Defense estimates a need to train 100,000 military health care personnel annually, representing a profound educational challenge.1 The consequences of ineffective medical training are dire. In the United States, medical errors are estimated to result annually in at least 50,000 excess deaths and 1,000,000 avoidable injuries.2 The military has long pursued the integration of simulation and robotic technologies into medical training and applications, and these techniques may provide leverage to address this issue.

Specifically, the use of these intelligent technologies for training military medical personnel can help measure skills, minimize errors, schedule training, and control the duration and expense of training. Medical personnel transitions, both when entering an operational area, and when returning to the civilian theater, are characterized by the need to rapidly reacquire skills that have decayed through disuse.

During this period of skill reacquisition, there is an increased risk of mistakes that may result in death and injury to patients. It is urgent that we increase the speed of reacquisition of surgical skills, while avoiding regaining the necessary skills on actual patients. The use of immersive virtual reality (VR) techniques coupled with metric-driven scheduling of training has the potential to dramatically reduce the cost of training, and the cost both in lives and dollars of errors and mistakes caused by lack of fluency in necessary techniques.

Although we are situating our solution within a particular surgical domain, our focus is on generalizing a theory of
skill decay to account for differential decay across skills that vary in their dependence on cognitive and psychomotor elements that will apply more broadly to medical skills. The particular domain we will focus on for preliminary work is laparoscopic surgery (LS), which we have chosen for its widespread use in the civilian theater, its affinity for study through simulation, and the precision of data that can be collected during its practice.

The skills required for LS, a minimally invasive surgery, transfer poorly from proficiency at open surgery on the same procedures, and the need for special training to acquire the fundamentals of LS has long been recognized.3 Laparoscopic skills are difficult to acquire and maintain, however, because of both their technical complexity and the challenging physical environment in which they must be executed, which includes a spatially restricted monocular visual field, limited tactile feedback, and a confined working area. Training for LS in an operating theater, the traditional approach to surgical training, is both costly and time consuming,4 limiting the time that surgeons can spend learning and practicing.

As a result, designers of medical curricula for minimally invasive surgery have turned to simulated environments in an effort to reduce or prevent the attrition of these critical surgical skills. Introductory LS training through simulation is now widespread, but has been primarily limited to novice laparoscopic surgeons practicing on simple psychomotor tasks, such as suturing. However, simulation training could be useful for training more complex tasks for surgeons with a wider range of experience, and for maintaining skills over periods of disuse. Few studies have examined durability of simulator-based training generally.5,6 Moreover, little attention has been given to learning important cognitive skills or more complex tasks involving sequences of simple tasks.

To increase our understanding of learning cognitive skills during surgical training, we developed several simple surgical tasks (e.g., peg transfer [PT], needle passing [NP]) in our adaptive VR trainer for LS training.7 These fundamental surgical tasks were used to train important basic surgical skills for complex and advanced tasks learning; for example, bimanual coordination, precision, and manipulations are the simple, but fundamental surgical skills for suturing.

Our adaptive training framework (Fig. 1) consists of three levels of design (modeling, comparison, and optimal training). A modeling methodology was developed including: (1) cognitive task analysis, to derive an ontology of the knowledge and skills to be measured and trained, (2) mathematical modeling, to determine the domain- and individual-specific variation in skill acquisition and attrition, and (3) cognitive modeling, to embed the specific model of skill attrition within a more general model of learner behavior, which can then be combined with the ontology derived from the task analysis.

At the second level of performance comparison, human subject data were collected from learners with varying levels of surgical experience, novices to expert surgeons. Our cognitive model at the first level is derived from and subsequently tested against the empirical human learning and forgetting data. We used the Adaptive Control of Thought/Rational (ACT-R) cognitive architecture8,9 to model the learner’s learning and forgetting. While these architectures are labeled “cognitive,” cognitive architectures have been developed to provide complete processing models including the entire range of cognitive, perceptual, and psychomotor behavior,10 and ACT-R in particular is an implementation of a unified theory of cognition.11 As such, ACT-R includes distinct modules for perceptual processing (visual and auditory), motor behavior, memory, and skill acquisition. It is exactly
this breakdown that we intend to leverage in identifying the emphasis on cognitive and psychomotor elements within skills. The human data were used to test potential interventions for the third training level based on objective measures (kinematics and electromyography [EMG]).

We then implemented the third level of optimal training schedules for the cognitive and psychomotor tasks to be learned. Our adaptive VR trainer will incorporate a cognitive model of skill acquisition and retention. We based it on previous work on the user modeling of learning, combining a hierarchical representation of surgical skills, where more complex skills and tasks are represented at higher levels, and training is driven by an estimate of individual skills and ability, as well as the dependence of the task on individual processing channels (e.g., cognitive, perceptual, and psychomotor). The model would tailor the training session to the level of complexity appropriate for the trainee at that moment in time and to predict and prescribe the course of training needed to produce a desired future level of competence, based on both demonstrated competence expected decay.

The Skill Retention Theory

As mentioned before, we can investigate surgical performance by considering skill acquisition and decay. A consensus understanding has been proposed, which specifies that there are continuous stages of learning. Many theories propose a three-stage process of learning: (a) the first stage for acquiring declarative knowledge to perform a procedural task, (b) the second stage for consolidating the acquired knowledge, and (c) the final stage for tuning the knowledge toward overlearning. Based on this understanding, Figure 2 shows the three different stages of learning and forgetting, providing important insights about how forgetting would be different for the learners at each stage.

The First Stage: Declarative

In the first stage of learning, skill acquisition occurs and simple training focused on skill acquisition may be adequate. For this first stage of learning and forgetting, knowledge in declarative memory degrades with lack of use, perhaps catastrophically as indicated by X’s in Figure 2, leading to the inability to perform the task. In this stage, learning and forgetting are explained by the activation mechanism in ACT-R. With lack of use, the strength of declarative memory declines. Decreased memory strength leads to response time increasing and accuracy decreasing. In addition, the ACT-R theory explains that increases in working memory load leads to decrements of retrieval performance from memory based on the activation mechanism. Thus, performance with this level of knowledge decreases with increased working memory load.

The Second Stage: Associative

In the second stage of learning, task knowledge is represented with a mix of declarative and procedural memory. With lack of use, the declarative knowledge can be forgotten, leading to missed steps. Procedural memory, on the other hand, is basically immune to decay. Forgetting slopes in this stage could vary by subtasks because mixed knowledge decays at different rates. In the first and second stage, catastrophic memory failure can occur because the declarative knowledge is not fully activated. In this mixed stage, training should be

FIGURE 2. A theory of skill retention, showing the three stages of learning and forgetting. The solid line indicates a learning curve and the dashed line indicates a forgetting curve from each corresponding stage. At each stage, the learning and forgetting rates are different.
provided to keep declarative knowledge active and also to support further proceduralization.

The Third Stage: Procedural

In the third stage of learning, task knowledge is available in both declarative and procedural forms, but procedural knowledge predominantly drives performance. Practice will compile knowledge into procedural knowledge. We describe this type of task knowledge to be proceduralized skill. With lack of use, declarative knowledge may be degraded. Nevertheless, learners can still perform the task—if all the knowledge is proceduralized and thus not forgotten with time. Less well-known skills that are infrequently used, like recovery from unusual errors, may be degraded. This type of skill would require knowledge retrieval from declarative memory unless task knowledge is proceduralized. In this final stage, or to reach this final stage, practice for proceduralization should be provided. It also suggests that training should occur until trainees reach the crossing thresholds, noted as dashed horizontal lines for the stage thresholds in Figure 2.

To identify the possible solution to optimize the military medical skills acquisition and retention, the purpose of this project is to develop a methodology using our adaptive VR trainer to study surgical skill decay and test an intervention based on the integration of cognitive models of surgical skill decay within our adaptive training framework. This testing framework can predict the decay effect and maximize the training experience by monitoring the occurrence of mistakes during skill acquisition and retention.

METHODS

We collected data from 5 novices and 4 medical trainees. Each participant performed two basic surgical tasks (Figs. 3A and 3B), 5 times in three sessions: at baseline, 1 week after baseline, and 1 month after baseline. Kinematic data including time to task completion, total distance traveled, and average speed of both hands were recorded. Muscle effort of four muscles (upper trapezius, anterior deltoid, flexor carpi radialis, and extensor digitorum) was monitored using a wireless Trigno EMG system (Delsys, Boston, Massachusetts). Using methods from our previous studies, raw EMG signals were recorded with a sampling rate of 2,000 Hz using EMG works acquisition software based on the manufacturer’s recommendation, and were processed with a band-pass filter of 20–300 Hz and smoothed by a root-mean-square technique with a 150 ms moving window to compute the root-mean-square EMG data. To reduce the intersubject variation, the maximal voluntary contraction (MVC) was obtained from each muscle to normalize EMG signals. The EMG data are presented as the percentage of MVC (%MVC).

RESULTS

All participants were able to complete the task. We examined their performance in terms of kinematics and EMG.

Kinematics

On the PT task, participants completed tasks during the first session in 153.3 seconds on average, whereas the tasks during subsequent sessions were completed more quickly in 132.1 and 123.4 seconds, respectively. Similarly, collapsing across sessions, trial 1 was completed in 153.4 seconds while subsequent trials within the session were completed more rapidly (Fig. 4). The NP task showed a similar pattern, with participants completing the baseline session trials in 120.5 seconds on average, and more quickly in 96.1 and 91.7 seconds in subsequent sessions (Fig. 5).

Given the expectation that completion time during learning follows a power law, a statistical analysis was conducted using the log of completion time (LogTime). For task PT, we performed a regression of session and trial on LogTime, producing an $R^2$ of 0.13 (degrees of freedom [df] = 121). For the PT task, session 3 (at 4 weeks) was significantly faster than prior sessions ($p < 0.01$), whereas trials 3 and 4 were significantly faster than trials 1, 2, and 5 ($p < 0.05$). Similarly, for task NP, we performed a regression of session and trial on LogTime, producing an $R^2$ of 0.16 (df = 125). Unlike the PT task, for the NP task, both sessions 2 and 3...
(at 1 week and 4 weeks) were significantly faster than prior sessions ($p < 0.05$ and $p < 0.01$, respectively). Further, for the PT task, trials 3, 4, and 5 were significantly faster than trials 1 and 2 ($p < 0.05$, $p < 0.01$, and $p < 0.01$, respectively).

We fit a preliminary predictive performance model to these data using the power law of learning and forgetting to make predictions, using the regression model of session and trial on LogTime for comparison. The initial model for the PT task produces an $R^2$ of 0.72 and a coefficient of variation of 0.09, indicating the model accounts for the majority of the variance in the human data. The preliminary model of the NP task produces an $R^2$ of 0.78 and a coefficient of variation of 0.15, similarly indicating that the model predicts the majority of the variance in the human performance data. Graphs of the model predictions are presented in Figures 6 and 7.

Qualitatively, the NP and PT task models predict the appropriate range of variation, with both models predicting the greatest speedup during the first session and decay in learning across sessions that effectively rewind this learning. Further, they capture the appropriate range of performances with the PT model spanning 171 to 114 seconds, and the NP model ranging from 142 to 90 seconds, in line with the human performance.

**EMG**

The EMG data were further analyzed within task by converting EMG data (%MVC) to a z-score measure, thereby controlling for individual differences in overall muscle activation. The EMG z-scores were averaged to provide an overall indication of muscle activation during task performance. Because of the normalization transformation, the z-score means for tasks PT and NP are both 0, enabling the collapsing of the two tasks. We performed a regression analysis of session and trial on the z-score of the EMG muscle activation, and found that session 3 (week 4) was characterized by statistically significantly less muscle activation ($p < 0.001$, df = 255) than sessions 1 and 2 (weeks 0 and 1).
Further, trials 3, 4, and 5 had significantly less muscle activation than trials 1 and 2 \((p < 0.05, \text{df} = 255)\).

**DISCUSSION**

In the course of working through practice sessions on the NP and PT tasks, participants clearly demonstrated both learning and forgetting, demonstrating statistically significant patterns. Participants exhibited speedup during skill acquisition and decay of that acquired skill during periods of disuse. Participants also showed less decay due to disuse as the skill became more practiced. Although the pattern of reaction time could be expected to fit a logarithmic regression model, our efforts to fit such a model only resulted in an \(R^2\) of 0.16 and 0.13, thus failing to account for the majority of the variance in participant completion times.

These patterns were much more successfully predicted through a model employing the Power Laws of Learning and Forgetting, encapsulated within the ACT-R cognitive architecture. This modeling, although preliminary, captures the majority of the variance available in the data set, with an \(R^2\) of 0.72 for the PT task, and an \(R^2\) of 0.78 for the NP task. Thus, the ACT-R based model captured substantially more variance than a logarithmic regression model.

The data appear to suggest that there is also a fatigue effect at work, and participants slow down after three or four trials of either the PT or NP task, though this pattern did not reach statistical significance within the small sample we evaluated. Although we have previously modeled such slowdowns,\(^\text{22}\) we have not yet attempted to apply a fatigue component to our modeling work. We would, however, expect to capture even more of the variance in the human performance through such a mechanism.

The EMG analysis confirms that the muscle activation required for the task was decreasing, and thus the procedural aspects of the task were becoming simpler. That is, the low-level motor learning appeared to be durable, but the task performance still shows decrements that are not accounted for by the durability of the motor skill. Our theory, however, accounts for the forgetting that takes place at the declarative and procedural levels as well, and predicts this time course of change over skill disuse.

Our future work will use further data analysis at the motor, declarative, and procedural levels to make holistic performance predictions. These research efforts will eventually help to address the maintenance of surgical skills, especially for experienced surgeons, and combat surgery. More importantly, the use of our adaptive VR trainer could measure medical skills in military medical personnel, minimize errors, schedule training, and potentially control the duration and expense of military medical training.

While preliminary, this work demonstrates the ability to accurately predict the acquisition and decay of surgical skill within a VR training system by leveraging the laws of learning and forgetting embedded within a cognitive architecture. Thus, this study serves as a validation for both the VR platform and the cognitive modeling paradigm. Given these components, we expect to be able to plan surgical training remediation with the aim of making the best use of available training time and avoiding costly mistakes because of skills that fall into disuse.

**CONCLUSIONS**

Based on the learning theory and this and similar results, our adaptive VR training system can improve the capability to interactively assess the level of skills learning and decay, optimize skill relearning across levels of surgical experience, and positively impact skill maintenance. Our training system could eventually reduce patient injury and morbidity by providing trainees with the reexperience they need before operating in a new theater.

**ACKNOWLEDGMENTS**

This study was supported by a NASA Nebraska Space Grant and EPSCoR (KCS), a Nebraska Research Initiative grant (KCS and DO), and grants from ONR No. W911QY-07-01-0004, N000141010401, and N00015-01-2275 (FER).

**REFERENCES**

Face Transplantation in a Highly Sensitized Recipient

Anil Chandraker, MD*; Ramon Arscott, MD†; George Murphy, MD‡; Christine Lian, MD‡; Ericka Bueno, PhD§; Francisco Marty, MD‖; Helmut Rennke, MD¶; Edgar Milford, MD**; Stefan Tullius, MD, PhD††; Bodhan Pomahac, MD§

ABSTRACT  Face transplantation was performed in a highly sensitized recipient with positive preoperative crossmatch and subsequent antibody-mediated rejection. The recipient was a 45-year-old female with extensive conventional reconstructions after chemical burns over the majority of the body. Residual quality of life and facial functions were poor. Levels of circulating anti-human leukocyte antigen (HLA) antibodies were high, and panel reactive antibody score was 98%. A potential donor was identified; however, with positive T and B cell flow crossmatches. The transplant team proceeded with face transplantation from this donor, under tailored immune suppression and with available salvage options. The operation was successful. Plasmapheresis and induction immune suppression (i.e., thymoglobulin followed by mycophenolate mofetil, tacrolimus, and steroids) were provided. Five days later, there was significant facial swelling, rising anti-HLA antibody titers, and unprecedented evidence of C4d deposits on skin. High doses of steroids and thymoglobulin were provided; however, rejection increased such that by day 19 it was diagnosed grade III in the BANFF scale. After stopping thymoglobulin because of serum sickness, combination therapy of plasmapheresis, eculizumab, bortezomib, and alemtuzumab was provided. HLA antibody levels decreased while swelling and redness improved. At 3 months, there were no longer signs of rejection on biopsy.

INTRODUCTION

A positive crossmatch indicates the presence of donor-specific antibodies (DSA), and has been considered a contraindication to solid organ transplantation due to risks of hyperacute rejection and allograft loss.1,2 Vascularized composite allotransplantation (VCA) dates back 16 years,3 and up until now has closely followed donor–recipient matching practices established in solid organ transplantation. Thus, all reported hand and/or face allotransplantation cases to date had negative donor/recipient crossmatches.4,5 By extension, there have been no reports of antibody-mediated rejection (AMR) in VCA,4,6 in spite of a high incidence of cell-mediated acute rejection.4 Specifically, there have been no reports of circulating DSA in VCA recipients and no reports of C4d deposition during acute rejection in VCA. C4d is a complement degradation product and marker of AMR that is generated when DSA bind to antigen and activate the complement cascade.6,7 Owing to this lack of precedent for AMR in VCA, the histopathological scale used to grade the severity of VCA rejection, the 2007 BANFF scale is based on historical features of cell-mediated rejection.8 Also by extension, all episodes of acute rejection in VCA to date had been treated with high-dose steroids and/or antithymocyte globulins (basiliximab or alemtuzumab), which specifically target cell-mediated rejection.4,5,9

Despite lack of evidence for humoral rejection in VCA, clinical practice still mandates an assessment of just how “sensitized” candidates for hand and/or face allotransplantation are. This assessment is depicted by pretransplant panel reactive antibody (PRA) scores.8 A PRA score is a calculated percentage risk that a given recipient would have a positive crossmatch with a potential organ donor when comparing the human leukocyte antigen (HLA) antibodies found in the candidate’s serum with a panel representative of the HLA class I and II molecules found in the general population.10 Furthermore, standard transplant medicine practice also entails ruling out the presence of DSA against HLA class I and II molecules by both (i) complement-dependent cytotoxicity (CDC) crossmatch and (ii) flow cytometry crossmatch (FCXM).10

This is a report on the successful management of the first face transplantation in a recipient with high PRA (i.e., highly sensitized) and positive crossmatch. Evidence of AMR is demonstrated by histological changes and C4d deposition in allograft biopsies with concomitant elevated titers of DSA. Lastly, this report describes the immune-suppression regimen successfully used to abrogate the rejection episode.

*Renal Division, Schuster Family Transplantation Research Center, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA.
†Department of Surgery, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA.
‡Division of Dermatopathology, Department of Pathology, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA.
§Division of Plastic Surgery, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA.
¶Division of Infectious Diseases, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA.
**Division of Nephrology, Department of Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA.
††Division of Transplant Surgery, Department of Surgery, Brigham and Women’s Hospital, Harvard Medical School, Boston, MA.

This work was presented at the Military Health Systems Research Symposium in Fort Lauderdale, FL, August 18–21, 2014.
doi: 10.7205/MILMED-D-15-00136
METHODS
The transplant recipient was a 45-year-old female, who 6 years prior had sustained chemical burn injuries to 80% of the body surface area and endured over 50 conventional reconstructive procedures mostly involving split-thickness and full-thickness skin grafting as well as release of contractures. At the time of presentation for face transplant evaluation, the patient demonstrated significant functional impairments, including severe contraction and eversion of the lips, a proximally retracted nose with reduced bulk and volume, lack of functional eyelids, and extensive and painful neck contractures (Fig. 1). All of these contributed to poor quality of life.

This remarkable patient was evaluated and deemed eligible for face allotransplantation by the multidisciplinary VCA team at the Brigham and Women’s Hospital (BWH), after a lengthy screening period that involved close institutional review board oversight and active participation of the institution’s bioethics team. Some of the strengths this candidate was found to possess were her previous experience as a nurse working on a transplant floor, which made her fully aware of the risks of immunosuppression and her excellent support network. Given her high degree of sensitization, which placed her at increased risk for hyperacute rejection and allograft loss, we implemented a protocol that has been used by our transplant medicine coinvestigators with highly sensitized patients in kidney transplantation with excellent 5-year survival. The patient provided informed consent to enroll in a face transplantation research protocol approved by the Partners Human Research Committee (protocol no. 2008P000550), underwent full screening and was subsequently placed in the transplant wait list. As part of the pretransplant screening, HLA class I and II antibodies in the patient’s serum were determined using a flow cytometry-based Luminex 100/200 System (Luminex, Austin, Texas) and single antigen screening beads. The Panel Reactive Antibody (cPRA) score was thus calculated.

The New England Organ Bank identified and obtained consent from the next of kin of a brain-dead donor who matched the patient’s sex and skin color and texture. Both CDC and FCXM flow cytometric crossmatches between the donor and the recipient were performed. The CDC-crossmatch was performed using donor T and B lymphocytes, mixing recipient serum and donor cells followed by complement. Antihuman globulin was also added to increase sensitivity and the assays were done with serial dilutions of recipient’s serum to gain information on the strength of the antibodies detected. The CDC crossmatch was expressed as positive or negative, based on the percentage of dead cells. FCXM was performed by incubating the recipient’s serum with donor lymphocytes, fluorochrome-conjugated anti-IgG antibody, and CD3 and CD19 monoclonal antibodies to identify T and B cells. Stained cells were processed by flow cytometry to assess median fluorescence binding of IgG from the recipient’s serum.

An immunosuppressive regime was designed based on the crossmatch results that involved traditional induction with antithymocyte globulin (ATG) 1.5 mg/kg, mycophenolate mofetil (MMF) 1g intravenous twice daily, a steroid taper and tacrolimus (Prograf) at 2 mg twice daily (up-titrated quickly to a goal level of 10 ng/mL), and plasmapheresis (therapeutic plasma exchange [TPE]) every other day starting on postoperative day 1 (POD1), with each TPE followed by 10 mg intravenous immunoglobulin (IVIG) to prevent rebound antibody secretion. Postoperatively, the immunosuppression regimen was modified as informed by allograft biopsy results and circulating DSA levels.

Following previously published principles of facial allotransplantation, a robust salvage plan was outlined in the event of loss of the facial allograft. Specifically, efforts were made to preserve the functional units of the recipient’s face, such as the functional cartilage of the nose. Extensive lysis of contractures and the use of split-thickness skin grafting over a dermal substitute were deemed highly likely to restore pretransplant appearance and function in the unfortunate event of allograft loss. Potential donor sites for split-thickness skin grafts were identified on the recipient’s left arm, forearm, and back.

![Figure 1](image-url)  Photographs of the recipient before face transplant, and immediately (day 0) and 3 months after surgery. The clinical appearance during acute allograft rejection is also provided, inclusive of biopsy sites on lower right neck (post-operative day 6, POD6). The 3-month time point corresponds to both clinical and histological resolution of the allograft rejection (post-operative day 95, POD95).
Postoperatively, biopsies from the skin of the allograft were obtained every time there were clinical signs of possible allograft rejection. Biopsies were assessed according to the BANFF classification of skin-containing composite tissues. The presence or absence of C4d in the allograft was determined by both immunoperoxidase and direct immunofluorescence staining of biopsy samples collected in formalin and Zeus transport solution, respectively. All biopsy specimens were received, processed, and tested by the Department of Pathology at the BWH; in particular, the immunofluorescence microscopy was developed and performance characteristics were determined by the Immunohistochemistry Laboratories in the Department of Pathology at the BWH.

Serum samples were acquired concomitant to allograft skin biopsies and assessed for DSA levels using flow cytometry-based Luminex 100/200 System and single antigen screening beads as described above.

RESULTS

Luminex solid phase prescreening of the stored recipient’s serum while on the wait list revealed antibodies to a significant number of HLA class I and II antigens, and a cPRA of 98. The patient waited in the list for 14 months, which at that time was 3 to 4 times longer than any of the prior VCA recipients at BWH. After the contingency plans outlined above were put in place, the transplant team decided to transplant the face of a donor against whom the recipient may have had DSA (i.e., positive crossmatch). A donor was identified shortly later and was accepted based on matching blood group, age, sex, and skin color.

A sample of recipient’s serum taken on the same day—before the transplant operation confirmed the presence of DSA, specifically, anti-HLA A2, A32, B57, BW4, DQ7, DQ9, and DR4. Anti-HLA DQ7, DQ9, and DR7 were present in the undiluted serum only, but anti-HLA A32, B57, and BW4 were still present at 1:32 dilution. The T cell flow crossmatch was positive at DFU 1,428 (cutoff of positivity is >60) and the B cell crossmatch was 1,850 (cut off > 1,000). Initial T cell CDC tests conducted in serum taken from the recipient 3 months prior and stored since yielded negative results, whereas when conducted using sera drawn on the day of transplantation the results were weakly positive with a cell death score of 20%.

The operation was performed uneventfully (Fig. 1). Peri- and postoperative immunosuppressive management was informed by DSA levels and allograft biopsy results. The standard induction regimen (centered on abrogating cell-mediated rejection) was bolstered with TPE and IVIG every other day, starting POD1, in an effort to reduce DSA burden and associated risks of AMR.

The allograft underwent one lengthy and complex rejection episode in the immediate postoperative period. This episode started by POD5, when the patient presented with significant lower facial swelling and erythema (Fig. 1). Induction therapy and three rounds of PTE had been completed. Biopsies of the allograft skin showed no evidence of cellular rejection. However, circulating DSA were stronger when compared with the pretransplant results. Anti-HLA A2, A32, B57, and DR7 remained present at 1:32 dilution (Fig. 2). Out of concern for humoral rejection, TPE was stopped after the fourth run and switched to complement blockade with eculizumab once per week, as well as administered a second steroid pulse and taper. MMF and tacrolimus remained unchanged.

By POD12, biopsies from the allograft skin showed perifollicular lymphocytic infiltration consistent with BANFF grade I rejection. As the patient was already undergoing a second steroid pulse, the team made no alterations in management. By POD15, however, redness and swelling were unchanged and there was more pronounced lymphocytic infiltration with exocytosis into epithelium, consistent with BANFF grade II rejection. In addition, for the first time signs of possible AMR became evident, as suggested by intraluminal neutrophils on the specimens of allograft skin and further supported by findings on immunofluorescence microscopy of capillaries in the papillary dermis and around the eccrine glands. Small arteries and arterioles were also reactive for C4d (3 to 4+/4+). Of note, the tissue was negative for C3 and C1q deposits and there was no evidence for immune complex deposition (Figs. 3 and 4). By POD19, allograft skin biopsies were graded as BANFF grade III rejection (Fig. 5). There were foci of epidermal lymphoid exocytosis and early primarily follicular apoptosis that were slightly more prominent. Again, occasional intraluminal neutrophils without frank necrotizing leukocytoclastic vasculitis were noted and the strength and distribution of C4d
staining remained just as prominent as that seen on POD15. There was also a further increase in circulating DSA levels: anti-HLA A2, A32, B18, B57, Bw4, DR7, DQ7, and DQ9 were all detected in the neat and 1:8 serum samples, and anti-HLA A2, B57, DR7, DQ7, and DQ9 in the 1:32 dilution samples. At this point, the immunosuppression regimen was modified. Specifically, the following interventions were implemented in an effort to counteract the humoral component of the acute rejection: (i) 6 additional runs of TPE and IVIG over the course of 8 days, (ii) eculizumab administration following TPE on POD20, 22, and 27, and (iii) bortezomib administration following TPE on POD22 and 25. Other interventions were carried out to address the cell-mediated component of the acute rejection, namely: (i) a third steroid pulse and taper over the course of POD16 to 25, (ii) another 6.5 mg/kg of ATG over POD19 to 24, and (iii) extracorporeal photopheresis daily on POD27 to 29. Of note, ATG was held on POD25 because of signs of serum sickness in the patient’s knees and ankles. The immunosuppression regimen was bolstered with a one-time dose of 15 mg of alemtuzumab on POD29 (Fig. 5).

These adjustments in immunosuppression propitiated a slow reduction in erythema of the allograft. By POD29, there was significant decrease in circulating DSA levels. Although anti-HLA A2, A32, B18, B57, Bw4, DR7, DQ7, and DQ9 were found in the neat serum samples, only anti-HLA B57 and DQ7 were detectable at 1:8 dilution and no DSAs were detectable at 1:32 dilution. Allograft biopsies were downgraded to Banff grade II rejection and showed less C4d immunoreactivity (Fig. 5). The above described immunosuppressive regime was continued for 2 further weeks: a fourth round of TPE and IVIG was performed on POD33 to 36, eculizumab was administered weekly, and 2.2 mg of Bortezomib were administered every 4 days. Maintenance with tacrolimus (target levels of 8–12 ng/mL), MMF, and steroids continued as well (Fig. 5).

By POD39, there were no DSA detected in the 1:32 or 1:8 sera dilutions, and allograft skin biopsies were unchanged in terms of Banff grade and C4d deposition. The recipient was discharged from the hospital on POD41 on a maintenance regime of tacrolimus, MMF, and low-dose steroids, as well as TPE and IVIG twice per week. On POD51, allograft biopsies showed no evidence of overt vasculopathy, and only superficial/mid-dermal perivascular lymphocytic infiltrate suggestive of Banff grade I. C4d staining persisted at that time, but the only detectable DSA were anti-HLA A2, B57, Bw4, DR7, and DQ7 found in the neat serum samples (Fig. 2). Although the team stopped TPE, circulating DSA levels continued their decline (anti-HLA B57,
showed only sparse supraventricular lymphocytic infiltrate, positive C4d staining only on the vessels around the eccrine glands of the dermis, with the papillary dermis being largely negative for C4d immunoreactivity. By POD116, there was no longer evidence of active cellular or AMR. At this point, this rejection episode was deemed resolved. Routine allograft biopsies taken subsequently yielded normal (i.e., rejection) results, until a biopsy taken on day 358 demonstrated perivascular chronic inflammation with minute focus of epidermal spongiosis associated with vacuolar interface change and exocytosis, consistent with BANFF grade II allograft rejection. The patient was admitted to the hospital for steroid bolus treatment (solumedrol 500 mg intravenous every day x3); however, 1 month later, the findings of BANFF II rejection persisted. Topical steroids were added. Clinical and histopathological presentation improved slowly and rejection was considered resolved by the beginning of the 14th postoperative month. On the 21st month, an allograft biopsy showed grade II rejection, which was treated with oral steroids, 100 mg for 5 days. Finally, on the 24th postoperative month there was another acute rejection episode, graded Banff II/III which resolved after 1 week of topical treatment with clobetasol ointment.

DISCUSSION

With regards to VCA, this case yielded the first observations of: (i) transplantation in a highly sensitized recipient with positive donor–recipient crossmatch, (ii) evidence of AMR, and (iii) successful management of AMR. AMR is therefore both possible and relevant in VCA, a point that has been refuted based on the absence of DSA and/or C4d deposition in prior VCAs. It is important to note, however, that all prior VCAs were performed in patients with negative donor–recipient crossmatches.

The hereby described findings call for a revision of the current BANFF working classification for VCA. When the BANFF classification was drafted in 2007, there was consensus that “several pieces of histologic and clinical information” needed to be gathered to define AMR in VCA, including “the presence of C4d deposition and its relationship with donor-HLA-specific antibodies, the presence of vasculitis, neutrophilic margination, thrombi and necrosis, a complete history of sensitization (e.g., PRA, crossmatch results, transfusions, pregnancies, and previous allografts), and the presence or absence of autoantibodies and T- and B-cell crossmatch performed before transplantation.” All of these pieces of information have been outlined in this case report, and as such they may help the eventual revision of the BANFF classification so as to define and incorporate AMR in the context of VCA rejection.

Sensitization is a common scenario in burn patients and in those with a history of allograft failure, and up until now had been considered a contraindication to VCA. This report may provide a step toward revisiting and expanding the criteria of eligibility for face allotransplantation. Face transplantation in a sensitized patient with a positive crossmatch was possible and could be managed by careful adjustments of immunosuppression with drugs currently used in solid organ transplantation; however, the rescue drug protocol was complex and costly, and needs further refinement. Although the patient continues to do well clinically with no evidence of rejection on biopsies 6 months postoperative, time will better reveal the eventual prognosis and help inform management and course.

CONCLUSION

Face transplantation was performed in a highly sensitized recipient with positive donor–recipient crossmatch. The 2-year postoperative outcomes suggest that face transplantation in this patient population is possible and manageable. The incidence of episodes of acute rejection in this patient has been comparable to those reported in the literature for other cases of vascularized composite allotransplantation, and have been managed successfully. There is no evidence of chronic rejection. However, further refinements to the immunosuppression protocol and longer followup are needed.

ACKNOWLEDGMENT

The authors wish to acknowledge the funding source, U.S. Department of Defense’s Biomedical Translational Initiative research contract no. W911QY-09-C-0216. The authors also acknowledge the New England
Organ Bank and the donor families for the gift of the facial allografts, without which this research would have never been possible.

REFERENCES
Development of Lyophilized Loop-Mediated Isothermal Amplification Reagents for the Detection of *Leptospira*

Hua-Wei Chen, PhD†; Giulia Weissenberger, BS†; Wei-Mei Ching, PhD†

**ABSTRACT** Leptospirosis is considered to be the most widespread zoonosis. This worldwide emerging infectious disease is caused by the pathogenic species belonging to the genus *Leptospira*. Polymerase chain reaction (PCR)-based diagnostic assays have been developed for detecting *Leptospira* DNA in cell cultures and clinical samples. Because PCR requires specialized equipment and extensive end-user training, it is not suitable for routine work in resource-limited areas. We have developed a loop-mediated isothermal amplification (LAMP) assay to detect the presence of *Leptospira* in patient, animal and environmental samples using lyophilized reagents at a single temperature of around 63°C with a heating block. The sensitivity of this LAMP assay is very similar to the PCR method. The amplified DNA products can be visualized with the naked eyes using hydroxy naphthol blue or detected by the fluorescence signal of SYBR green dye in the reaction when an ultraviolet lamp or compact fluorescence tube scanner is used. This LAMP assay is simple, rapid, and can be performed with a water bath or heating block. The lyophilized LAMP reagents are stable for 3 months when stored at 4°C and 1 month when stored at 25°C, respectively. It is ideal for resource-limited settings where leptospirosis is endemic.

**INTRODUCTION**

Spirochetes of the genus *Leptospira* cause leptospirosis, which is considered the most widespread zoonotic disease in the world. Pathogenic *Leptospira* species colonize the renal tubules of chronically infected hosts such as dogs, rats, and cattle. These maintenance hosts typically remain clinically asymptomatic and shed leptospires into the environment in their urine. Transmission of leptospirosis in humans and non-maintenance host animals occurs as a result of incidental contact when infected animals contaminate water or soil or by direct contact. The disease represents a potential hazard to U.S. military personnel deployed or training in disease-endemic regions. Sporadic outbreaks occurred in the Vietnam conflict. U.S. troops training in Panama since the 1960s were involved in outbreaks and have experienced sporadic infections. Traditionally, disease transmission is related to certain socioeconomic or climate conditions that favor endemcity in animal reservoirs and human exposure that are generally confined to developing countries. In recent years, human outbreaks have been observed to be related to outdoor recreational activities and army expeditions. Because of deployment, there is a high risk among U.S. military and civilian personnel overseas. Recent studies showed that leptospirosis is prevalent in the Caribbean, Latin America, the Indian subcontinent, southeast Asia, and Oceania. A seroepidemiological survey of rodents collected at a U.S. military installation in South Korea indicated that the seroconversion rate of leptospirosis is higher than scrub typhus or murine typhus. The Armed Forces Medical Intelligence Center ranked leptospirosis as the 7th highest global risk-severity index disease among 53 infectious diseases of military significance in 2007. In April 2010, a panel was held to prioritize infectious diseases of military importance and created the Infectious Disease Threats Prioritization Panel, ranked leptospirosis 10th on the list. It is again among the top tier in the gap analysis conducted by MIDRP Program L based on the end-user feedback in 2012.

Symptoms of leptospirosis are easily confused with other febrile illnesses such as dengue and malaria, which require different treatment regimens. The acute phase of the disease lasts approximately 1 week. Untreated patients may suffer kidney and liver damage. The mortality rate in the severe form can be as high as 15%. Therefore, timely diagnosis is essential as antibiotic therapy provides the greatest benefit when initiated early in the course of illness. Despite the high risk of leptospirosis infection in developing countries, it is often underreported. The lack of rapid diagnostic tests often presents a barrier to the early diagnosis of leptospirosis, especially in areas where there is a high prevalence of malaria and dengue. In these areas, most cases of acute febrile illness are clinically diagnosed as malaria or dengue instead of using laboratory methods that differentiate infectious pathogens that may be the cause of infection or may coinfect individuals.

Microscopic agglutination test is considered the gold standard of serological diagnosis. This assay is time consuming as it works by detecting the antibody titer increase.

---

*Naval Medical Research Center, 503 Robert Grant Avenue, Silver Spring, MD 20910.
†Uniformed Services University of the Health Sciences, 4301 Jones Bridge Road, Bethesda, MD 20814.
Presented at the Military Health Science Research Symposium, Fort Lauderdale, FL, August 2014.

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, or the U.S. Government.

This article was prepared as part of the author’s official duties. Title 17 U.S.C. §105 provides that “Copyright protection under this title is not available for any work of the U.S. Government.” Title 17 U.S.C. §101 defines a U.S. Government work as a work prepared by military service member or employee of the U.S. Government as part of that person’s official duties.

doi: 10.7205/MILMED-D-15-00149
in serum samples obtained weeks apart. Although this technique provides an efficient retrospective diagnosis, it does not provide early diagnosis.18 Other diagnostic tests include dark-field microscopy, enzyme-linked immunosorbent assay, and Western blot assays known to have low sensitivity.29 Because of the low sensitivity, the use of these serological assays for the initial management of acute leptospirosis in adult patients was found inferior to empirical treatment in a study conducted in Thailand.20 The need for rapid diagnostics at the time of admission has led to the development of numerous polymerase chain reaction (PCR) assays. PCR and real-time quantitative PCR (qPCR) can be used for detecting leptospirosis in cell culture and clinical samples within the first week of illness, before the detection of antibodies where initiation of treatment may be effective.21 PCR or qPCR is based on the detection of genes universally present in bacteria such as gyrB,22 rrs (16S ribosomal RNA gene),23 secY,24 or genes restricted to pathogenic Leptospira spp. such as lpl32,25 ligA, and ligB.26 Both PCR and qPCR are costly and often not readily available in many laboratories, especially those where leptospirosis is endemic.

Originally discovered by Notomi et al.27 loop-mediated isothermal amplification (LAMP) offers an alternative DNA amplification method. LAMP uses Bst DNA polymerase for strand displacement DNA synthesis along with primers that create cauliflower-like structures with multiple loops.27 The most significant advantage of LAMP is that amplification occurs under isothermal conditions. Therefore, only a heating block or a water bath is required. The reaction products can be seen by agarose gel electrophoresis, using a fluorescent dye such as SYBR green to be visualized under ultraviolet light28 or by hydroxy naphthol blue (HNB): a metal indicator that can be seen by the naked eye.29

We have developed a highly specific LAMP assay using a combination of primer sets targeting the lpl32 and lpl41 genes.30 Unlike PCR, which requires a thermocycler and multiple temperatures, LAMP only requires a water bath or heating block and at a single temperature. The limit of detection for this LAMP assay is about 12 copies and very sensitive. Our results showed that lyophilized LAMP reagents retain the same reactivity as freshly prepared reagents and were stable for 3 months at 4°C, 1 month at 25°C, and less than 1 week at 37°C. These lyophilized LAMP reagents are ideal for a resource-limited setting where leptospirosis is endemic because of their sensitivity and durability even when stored at 25°C.

MATERIALS AND METHODS

Primers

The primer sets used in this study were the same as the ones described previously in Chen et al.30 All primers were synthesized by Eurofins MWG Operon (Huntsville, Alabama).

Genomic DNA Template

Genomic DNA used in LAMP assay was extracted from Leptospira interrogans Copenhageni strain Fiocruz L1-130 by QiAamp DNA Mini Kit (Qiagen, Stockach, Germany) following the manufacturer’s instructions.

LAMP Reaction

LAMP reactions were carried out as described previously.27 In brief, a 25 μL reaction mixture contained 1.6 mM of each FIP and BIP primer, 0.8 mM of each LF and LB primer, 0.2 mM of each F3 and B3 primer, 20 mM Tris HCl (pH 8.8), 10 mM KCl, 8 mM MgSO4, 10 mM (NH4)2SO4, 0.1% Triton X-100, 0.8 M betaine (Sigma-Aldrich, St. Louis, Missouri), 1.4 mM dioxynucleoside triphosphate (dNTP) mixture (New England Biolabs, Beverly, Massachusetts), 8 U Bst DNA polymerase (New England Biolabs), and 5 μL of DNA template. The reaction mixture was incubated at 63°C for 60 minutes. Each reaction was terminated by adding 5 μL of 10X BlueJuice (Invitrogen, Carlsbad, California) for gel detection. The reaction products were examined by electrophoresis on a 2% agarose gel stained with a 1:10,000 dilution of GelRed (Phenix Research Products, Asheville, North Carolina). Other detection methods include addition of HNB into the reaction mixture35 to enable direct visual detection or inclusion of SYBR green to detect the reaction products by an ultraviolet light. Furthermore, the inclusion of SYBR green in the reaction mixture also allowed real-time detection with fluorescence measurement systems such as the ESEQuant Tube Scanner (Qiagen) and 7500 Fast Real-time PCR System (Applied Biosystems, Foster City, California).
Lyophilized Reagents

Bst DNA polymerase, dNTPs, and primers were supplemented with sugars and subjected to lyophilization in a freeze-dryer using a protocol based on the method described by Saleki-Gerhardt and Zografii.32 This service is provided by GeneReach USA (Lexington, Massachusetts). Tubes containing those reagents were packaged in a zipped aluminum foil bag (Fig. 1).

Reaction Mixture for Lyophilized Reagents

20 μL of reconstitution buffer containing 25 mM Tris-HCl (pH 8.8), 12.5 mM KCl, 10 mM MgSO4, 12.5 mM (NH4)2SO4, 0.125% Triton X-100, and 1 M betaine was added into tube to resuspend the lyophilized reagents in the individual vials. 5 μL of DNA template was added to start the reaction.

RESULTS

Evaluation of the Lyophilized Reagents

The LAMP reactions were performed side by side using the lyophilized reagents and freshly prepared reaction mixtures. Both preparations were able to detect 12 copies of genomic DNA (Fig. 2). The stability of the lyophilized reagents was also tested and followed for 3 months when the reagents were stored at 4°C, 25°C, and 37°C (Table I).

Detection of LAMP Products

Agarose gel electrophoresis of the LAMP products was detected using GelRed, which displayed the typical ladder-like pattern (Fig. 2). Alternative detection methods included the change of reaction mixture color due to the decrease of Mg2+ ions concentration by metal ion indicator HNB (Fig. 3A) or the detection of double-stranded LAMP products by using SYBR green (Fig. 3B). Monitoring the fluorescence signal in real-time with SYBR green presence in the reaction was also performed with a fluorescence measurement system (Fig. 3C). All these methods have the same detection limit at 12 copies per reaction. Both alternative methods allow the postreaction read out to be done without opening the tubes.

DISCUSSION

Leptospirosis is the most widespread zoonotic disease throughout the world and a military priority infectious disease for developing better detection methods. Because the clinical presentation of leptospirosis is very similar to malaria and dengue, it is not possible to reliably predict the cause of infection based on clinical signs and symptoms.16 Early detection of Leptospira infection is particularly important to help guide appropriate treatment as early as possible. Although culturing and microscopic agglutination test are the gold standard methods for leptospirosis diagnosis, they are not practical for early detection. Culturing Leptospira takes weeks and anti-Leptospira antibodies can only be detected by the end of first week after onset of symptoms. Molecular tests like PCR allow testing clinical samples during the first 5 days of disease onset. PCR-based diagnostic assays have been developed for detecting Leptospira DNA in cell cultures and clinical samples.22–26 As PCR requires

<table>
<thead>
<tr>
<th>Detection Limit (cp/rxn)</th>
<th>Day 0</th>
<th>Day 7</th>
<th>Day 14</th>
<th>Day 28</th>
<th>Day 56</th>
<th>Day 84</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stored at 4°C</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Stored at 25°C</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>12</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Stored at 37°C</td>
<td>12</td>
<td>50</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>200</td>
</tr>
</tbody>
</table>

*Reaction mixtures were incubated at 63°C for 60 minutes.
specialized equipment and extensive end-user training, it is not suitable in resource-limited areas for routine work.

Previously, we developed a sensitive and specific LAMP assay targeting the pathogenic spp. specific genes lipL32 and lipL41. As reported in a recent manuscript, the assay can detect 12 copies of L. interrogans and does not detect genomic DNA from other clinically encountered bacterial species, such as different strains of Orientia tsutsugamushi, Rickettsia typhi, Rickettsia conorii, Rickettsia rickettsii and Bartonella bacilliformis. In this study, Bst DNA polymerase, dNTPs, and mixture of primer sets were lyophilized and packaged in a zipped aluminum foil bag. The sensitivity of these lyophilized reagents remains like to demonstrate that this LAMP assay using lyophilized reagents can detect Leptospira with sensitivity similar to that of PCR in a remote setting.

ACKNOWLEDGMENT

This research was supported by Naval Medical Research Center, research work unit (grant no. 6000.RAD1.J.A0310).

REFERENCES


**Magnetic Nanoparticles as a Potential Vehicle for Corneal Endothelium Repair**

Lauren E. Cornell, MS*; Jenny L. Wehmeyer, PhD*; COL Anthony J. Johnson, MC USA*; Mauris N. Desilva, PhD†; David O. Zamora, PhD*

---

**ABSTRACT** The corneal endothelium is paramount to the health and function of the cornea as damage to this cell layer can lead to corneal edema, opacification, and ultimately vision loss. Transplantation of the corneal endothelium is associated with numerous limitations, including graft rejection, thus an alternative therapeutic treatment is needed to restore endothelial layer integrity. We hypothesize that a nanotechnology-based approach using superparamagnetic iron oxide nanoparticles (SPIONPs) can ultimately be used to guide corneal endothelial cells (CECs) to injured areas via an external magnetic force without changing their morphology or viability. In this feasibility study we examined the effects of SPIONPs on the morphology and viability of bovine CECs in the presence of a magnetic force. The CECs were exposed to increasing SPIONP concentrations and the viability and cytoskeletal structure assessed over 3 days via metabolic analysis and rhodamine phalloidin staining. Significant differences (p < .05) in the metabolic activity of the CECs (100×10⁶ SPIONP/cell) occurred in the presence of magnetic force versus those with no magnetic force. No differences were observed in the cytoskeleton of CECs in the presence or absence of magnetic force for all SPIONP concentrations. These SPIONPs will next be evaluated with human CECs for future applications.

---

**INTRODUCTION**

The corneal endothelium aids in maintaining corneal transparency and hydration, but has little to no regenerative potential.¹ Eye injuries that involve damage to the corneal endothelium pose a serious problem for visual acuity of the patient. Cellular dysfunction of this layer is the second leading cause of corneal blindness.² In modern military operations, such as Iraq and Afghanistan, ocular trauma is responsible for 16% of battlefield evacuations, with only 20% of those injured returning to duty.³ Current standards of care for corneal endothelium repair involve surgical techniques such as keratoplasty which involve surgery to the cornea, most notably corneal transplants. Penetrating keratoplasty, or a full thickness corneal graft, has been the standard of care for endothelial decomposition for the past 50 years. It is associated with several limitations such as permanent weakening of the eye, postsurgical complications in relation to graft rejection, and increased intraocular pressure.⁴,⁵ While rejection rates have improved with more tissue specific keratoplasty procedures,⁶ surgical keratoplasty methods are plagued by a ∼35 to 50% drop in endothelial cell counts within the first 5 years, with continued gradual decline of endothelial cells every year afterwards regardless of technique. Current treatment options are limited to replacement of the entire graft.⁷–¹¹ Successive grafts however pose issues of earlier onset of allograft reactions than first time grafts and have a higher chance of eventual failure especially if operative complications such as persistent inflammation or neovascularization occurred with the first procedure.¹² Thus a tissue specific form of therapy which can be accomplished without replacing the entire tissue is needed to maximize long term viability of treatments for endothelial dysfunctions.

Cellular replacement therapies, combined with advanced tissue engineering methodologies, may play a vital role in the treatment of endothelial disorders. Some emerging research strategies involve (1) the production of transplantable endothelial cell sheets or (2) the injection of endothelial cell colonies into the damaged area.¹³ For both strategies, corneal endothelial cells (CECs) can be cultured and expanded in the laboratory and then transferred to the patient. However, invasive surgical procedures similar to keratoplasty are still required to implant the cell sheets.¹⁴ Furthermore, cells delivered via these methods continue to have poor cellular attachment to the corneal posterior stroma.

Mimura et al sought to overcome this limitation by magnetically guiding the attachment of rabbit endothelial cells loaded with ferromagnetic (spherical iron powder) particles. They demonstrated that magnetically guided, ferromagnetic-loaded endothelial cells could reduce corneal edema and form a monolayer on the Descemet’s membrane when injected in vivo into an injured rabbit cornea. Upon investigation for intraocular complications, there was no rise in intraocular pressure; however iron leakage from the loaded cells showed potential to damage surrounding ocular tissue and long term safety needed to be confirmed.¹⁵ While iron powder was a useful mechanism for providing feasibility for magnetic

---

*Institute of Surgical Research, Ocular Trauma 3698 Chambers Pass STE B JBSA Fort Sam Houston, TX 78234-7767.
†Naval Medical Research Unit San Antonio 3650 Chambers Pass, JBSA Fort Sam Houston, TX 78234-7767.

This information was presented at the Military Health System Research Symposium, Fort Lauderdale, FL, August 20, 2014, and at the San Antonio Postdoctoral Research Forum, University of Texas Health Science Center at San Antonio, San Antonio, TX, September 15, 2014.

The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

doi: 10.7205/MILMED-D-15-00151
movement of cells, it left room for improvement as a therapeutic option. Magnetic nanoparticles embody unique physical properties which enable them to function at the micrometer and submicrometer levels of biological interactions. For these reasons they are actively investigated for application in medical research for a variety of maladies such as cardiovascular disease, neurological disease, and cancer. Superparamagnetic iron oxide nanoparticles (SPIONP) hold a specifically valuable characteristic in that they can be manipulated with external magnetic fields, but they do not retain magnetic properties once the field is removed, thus reducing the problem of clumping. Patel et al. utilized this property of superparamagnetic spheres (100–900 nm in diameter) and were able to facilitate endothelial cell attachment to the posterior human corneal stroma in conjunction with a magnetic field.

It has been established that nanoparticle size and composition impact cellular uptake and cytotoxicity, with these results varying across cell lines and cell culture conditions. To date, there is a lack of in-depth cellular analysis of SPIONPs exposure on CEC function. Additionally, SPIONPs as a preferred nanoparticle and method of cellular “magnetization” for endothelial cell movement should be considered due to its favorable size, magnetic properties, and current use in disease research such as targeted cancer therapy. Therefore, this investigation aims to evaluate the interaction between SPIONPs and CECs.

While SPIONPs hold great promise as a tool for CEC guidance, their potential impact on CEC viability must first be investigated. Thus, this work examined the biocompatibility of SPIONPs on bovine CEC. Specifically, the viability and cytoskeletal structure of CECs exposed to SPIONPs in the presence of a magnetic force.

METHODS

Nanoparticle Characterization

Dextran SPIONPs with a biotin coating had an iron core of 7 to 10 nanometers (5 mg/mL in H2O), and were purchased from Micromod (Rostock, Germany) at a size of 50 nanometers. Particle size and colloidal stability were measured at a concentration of 17 μg/mL in H2O with dynamic light scattering via a Brookhaven particle size analyzer (Holtsville, New York) and accompanying Zeta Potential software. The SPIONPs were vortexed before testing to minimize clumping.

Cell Culture

A primary line of bovine CECs was purchased from Astarte Biologics (Bothell, Washington). The CECs were seeded in 48-well plates at 21,000 cells per a well. Cells were cultured in Dulbecco’s Modified Eagles Medium (DMEM; Life Technologies, Grand Island, New York) supplemented with 10% Fetal Bovine Serum (Thermo Scientific, Logan, Utah) and 1X Antibiotic-Antimycotic (Life Technologies). The CECs were maintained in culture for 48 hours in a sterile incubator at 37° C, humidified 5% CO₂/95% air environment, before SPIONP loading. All experiments were performed at cell passage number 6 to 8.

Nanoparticle Exposure

Bovine CECs were grown in a monoculture in order to directly target and expose the cells of interest to SPIONPs for endocytosis. After 48 hours of culture, cells were exposed to SPIONPs at the following SPIONPs/Cell concentrations: 0, 1×10⁶, 10×10⁶, and 100×10⁶. Fresh media was placed into each well and respective nanoparticle quantities were added. SPIONPs remained in culture with cells until evaluation at t = 1, 2, 3 days.

Magnetic Exposure

CECs were introduced to a constant magnetic force achieved by placing 35N-neodymium disc magnets (Magnet Shop, Culver City, California) directly beneath the tissue culture plate at a distance of approximately 0.05 inches. A compass was used to ensure uniform direction of magnetic pull throughout the plate. The CECs were exposed to an external magnetic force for up to 3 days and evaluated at t = 1, 3 days. Cells not exposed to SPIONPs were used as a control for all studies.

Superparamagnetic Iron Oxide Nanoparticle Uptake

The internalization of SPIONPs by CECs was confirmed with iron sensitive Prussian blue staining (Sigma Aldrich, St. Louis, Missouri). After incubation with SPIONPs, cells were washed with phosphate buffer saline (PBS-Life Technologies, Grand Island, New York) to remove any free nanoparticles. Cells were fixed in 4% paraformaldehyde (Electron Microscopy Sciences, Hatfield, Pennsylvania) for 10 minutes and then washed with PBS two times. A solution made of equal volumes potassium ferrocyanide (4% w/v) and hydrochloric acid (1.2 mM; Sigma Aldrich) was placed on the cells for 10 minutes. They were then rinsed with deionized water two times and counterstained with a 1% w/v paraarosaniline hydrochloride solution (Sigma Aldrich) for 5 minutes. The cells were then rinsed with deionized water and imaged using an objective of 10× with bright field mode on an Olympus 1X71 microscope (Olympus, Center Valley, Pennsylvania). All solutions were prepared fresh before experiments for each time point.

Cell Viability and Actin Filament Structure

Viability was assessed at t = 1, 2, 3 days using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT; Sigma Aldrich) and Live/Dead staining (Life Technologies). Fresh media was placed into each well and 20 μL of MTT solution (5 mg/mL) was added. Cells were incubated at 37° C for 2 hours to allow for the formation of formazan crystals. The media was then removed and the formazan crystals were dissolved in 200 μL dimethyl sulfoxide (DMSO; Sigma Aldrich) for 10 minutes at 37° C. The absorbance, or
optical density (OD), was read with a Synergy Mx Multi-Mode microplate reader (37°C) at 540 nm using 25 point spectral area scanning. For the Live/Dead assay, media was removed and 150 μL of a 1μM calcein AM and 2μM EthD-1 solution was added. The cells were then incubated for 45 minutes at room temperature and imaged with phase contrast with an immunofluorescent microscope. Actin filament structure was examined using rhodamine phalloidin (Life Technologies). Cells were washed with PBS two times and then fixed in 4% paraformaldehyde for 10 minutes at room temperature. The cells were then washed two more times with PBS and soaked in a 0.1% solution of Triton X-100 (Sigma Aldrich) in PBS for 5 minutes. They were again washed twice with PBS followed by incubation with rhodamine phalloidin (6.6 μM in Methanol) for 20 minutes at room temperature. Cells were rinsed two more times with PBS and the nuclei counterstained with 4',6-diamidino-2-phenylindole (DAPI; Life Technologies) before being imaged. All images were acquired using an Olympus 1X71 microscope under fluorescent imaging modes with Live/Dead photos taken using an objective of 10x and rhodamine phalloidin being taken with a 20x objective in place.

Statistical Analysis
Statistical analysis of numerical results were performed using the GraphPad Prism Statistical Software package (GraphPad Software, San Diego, California). Numerical data are expressed as mean ± standard error of the mean. Comparisons between groups were made using either a two tailed t-test with value \( p \leq 0.05 \) considered statistically significant, or a one-way analysis of variance (ANOVA) with values of \( p \leq 0.05 \) considered statistically significant. Post-hoc analyses were performed using the Tukey’s multiple comparison test.

RESULTS
Nanoparticle Characterization
The mean particle size of the SPIONPs was 76.5 ± 0.69 nm with a polydispersion coefficient of 0.194. A multimodal size distribution occurred with modes at 50 and 100 nm. The surface charge of the particles was determined to be −25.09 mV.

Cellular Uptake
Particle uptake was evidenced by the presence of blue-stained vesicles in the cytoplasm of the CECs (Fig. 1). An increased number of blue-stained vesicles can be visualized with increasing nanoparticle concentration (0 to 100x10⁶ SPIONPs/cell) and more notably when the SPIONP-exposed cells were introduced to magnetic force at 12,300 Gauss. As expected, there was no evidence of blue deposits in the control cells in the absence of SPIONP exposure.

Cell Viability under Nanoparticle Exposure Conditions
In order to evaluate the cellular interaction upon exposure of varying nanoparticle concentrations, Live/Dead staining was performed to ascertain viability of the CECs. The staining suggests that viability as a cellular function is maintained throughout all concentrations of SPIONPs loaded (Fig. 2). There is a larger quantity of dead cells in the highest concentration of the magnetically treated group. These findings were further evaluated using an MTT assay.

![FIGURE 1. Prussian blue staining of corneal epithelial cells after 3 days of culture with superparamagnetic iron oxide nanoparticles(SPIONPs). There is an increase in intracellular SPIONP uptake with increasing nanoparticle concentration and with magnetic exposure as indicated by the darker stained regions. The scale bar is at 500 μm.](image-url)
The viability of CECs subjected to SPIONP exposure with and without magnetic force was quantified using the MTT assay via comparison of mean OD values. The OD value will increase, due to increased metabolic activity, exhibited by the release of formazan crystals into solution. In addition, the effect of magnetic force on cell viability was also determined. A comparison was made of all concentrations of SPIONPs within respective time points (t = Day 1, Day 2, Day 3) to determine if exposure concentration impacted viability on that day. The concentration of nanoparticles exposed to the cells did not have an effect on cell viability during the first 2 days of culture when comparing all days for the non-magnetically exposed group. By day 3, the highest concentration of SPIONP exposure, 100x10^6 SPIONPs/cell, showed a significant decrease (p < 0.05) as compared to all other concentrations on day 3 with a drop of 7.06% for the control and 9.17% and 6.82% for the next two doses respectively (Fig. 3A). Despite this drop in metabolic activity, the mean OD value for 100x10^6 SPIONPs/Cell on day 3 (1.078) is still higher than 100x10^6 SPIONPs/cell day 1 values (1.025) showing a 5.17% increase in metabolic activity at the same concentration. CECs exposed to constant magnetic force showed no significant changes in viability for day 1. However, for days 2 and 3, a significant decrease in viability was found for the highest concentration, 100x10^6 SPIONPs/Cell, against all other concentrations for their respective time points (Fig. 3B). For day 2 as there was a drop in viability by 9.5% as compared to the control and by 7.2% and 7.5% for the following doses respectively. Day 3 showed significant differences in viability for 100x10^6 SPIONPs/cell as compared to all other concentrations within that day showing a drop in metabolic activity of 11.4% for the control and 10.7% and 10.37% for the next two doses respectively. The day 3 OD value (1.038) at the 100x10^6 SPIONPs/cell is lower than it’s day 1 value (1.068) by 2.8%, however despite the drop in metabolic activity on day 2 (1.069) as compared to all other concentrations, the viability at the 100x10^6 SPIONPs/Cell concentration is actually .09% higher on day 2 than day 1.

The previous figures examined the impact of concentration and time on cellular viability, however we wanted to understand the roll that magnetic force played in these results. To isolate
the effect of magnetic force on mean OD (viability) of CECs, a direct comparison was done for each SPIONP concentration examining the difference between the lack of and the application of magnetic force for each time point. This was done to examine the impact of magnetic force on each time point and concentration, thus determining if it’s application had any cellular consequences. Magnetic force alone had an effect on cell viability during the first 2 days with a significant increase in mean OD value being observed for all concentrations of SPIONPs loaded. By day 3, cells that were not treated with magnetic force reached similar viability levels and, at the highest concentration of SPIONP loading (100×10^6 SPIONPs/cell), a significant drop of 3.73% in viability occurred as seen in Figure 4. Thus, this figure shows indications that the application of magnetic force may increase the viability of CECs and increase metabolic activity within the first 48 hours exposure of cells to SPIONPs.

Cytoskeleton Structure

Cytoskeletal staining was performed using rhodamine phalloidin to determine the organization of actin filaments in CECs exposed to SPIONPs with and without exposure to magnetic force (Fig. 5). There were no observed differences in the arrangement of actin filaments within the CECs cytoskeletal structure, even when CECs were exposed to the magnetic field for up to 3 days. Thickness, orientation, and staining intensity of the actin filaments did not appear different between control and experimental groups.

FIGURE 3. 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) analysis showed that there were no significant differences on the effect of superparamagnetic iron oxide nanoparticles (SPIONP) concentration on cell viability throughout days 1 and 2 for the nonmagnetically treated group. By day 3 cells loaded with the highest concentration of SPIONPs, 100×10^6 SPIONPs/Cell, showed a significant decrease in viability as compared to all other concentrations of SPIONPs loaded at that time point (A). Similarly, MTT analysis of magnetically treated cells showed no significant difference in viability relative to SPIONP concentration on day 1, however on days 2 and 3 100×10^6 SPIONPs/Cell showed a significant decrease compared to all other concentrations per their respective time points (B).

FIGURE 4. 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) data showed significant differences for days 1 and 2 over all groups. The corneal endothelial cells (CECs) exposed to a continuous external magnetic force had higher mean optical density (OD) value, indicating increased viability, compared to CECs not exposed to magnetic force. By day 3 CECs loaded with the highest concentration of SPIONPs (100×10^6 SPIONPs/Cell), not exposed to magnetic force had higher mean OD value, indicating increased viability, compared to CECs loaded with the highest concentration of SPIONPs (100×10^6 SPIONPs/Cell), not exposed to magnetic force. SPIONPs coated with dextran in the first 48 hours exposure of cells with SPIONPs showed cellular uptake efficiency of the particles was dependent until an exposure concentration 100×10^6 SPIONPs/Cell is reached; furthermore, we were able to show that the cytoskeletal structure remained unaltered throughout the duration of the study. The results of this study may help to understand not only the interaction of SPIONPs on CECs, but also to elucidate the loading limits if induced cellular movement is to be achieved.

Previous studies have shown that size and coating material of SPIONPs play a critical role in cell endocytosis of the nanoparticles as well as cytotoxicity in mammalian cell lines. A strategy used to enhance the cellular uptake efficiency of SPIONPs is to modify their surface coating. In one study, researchers compared the uptake of aminosilane and dextran coated SPIONPs in six different mammalian cell lines and showed cellular uptake efficiency of the particles was dependent on both the cell type and SPIONP surface characteristics. The results of this study also showed that dextran coated nanoparticles did not adversely impact cell viability of any of the cell lines tested.

Superparamagnetic iron oxide agents coated with dextran are currently used for many purposes in the medical field, including two clinically approved SPIONPs as MRI contrasting agents for the liver and research to transform nanoparticles into target probes. After intracellular uptake, lysosomes

DISCUSSION

The focus of this work was to investigate whether exposing bovine CEC, which hold similar characteristics to human CECs in relation to their pump function and active transport, to SPIONPs (dextran with biotin coating) would adversely impact viability or physical morphology in relation to particle concentration. We have successfully demonstrated that CEC viability is not adversely influenced until an exposure concentration 100×10^6 SPIONPs/cell is reached; furthermore, we were able to show that the cytoskeletal structure remained unaltered throughout the duration of the study. The results of this study may help to understand not only the interaction of SPIONPs on CECs, but also to elucidate the loading limits if induced cellular movement is to be achieved.

The results of this research may help to understand not only the interaction of SPIONPs on CECs, but also to elucidate the loading limits if induced cellular movement is to be achieved.
metabolize the SPIONPs into a nonsuperparamagnetic, soluble form of iron that joins the normal iron pool.\textsuperscript{26,27} Despite the common use of SPIONPs in medical applications, many aspects of nanoparticle-induced cytotoxicity remain unclear. Due to its current role in research, dextran was our coating of choice for these studies with the addition of biotin as a factor to aid surface charge of our particles.

In the process of evaluating the effect of magnetic stimulation on the CECs, exposure of magnetic force had a significant impact on cell-SPIONP loading (visually confirmed via Prussian blue staining), metabolic activity, and viability. It has been previously shown that magnetic fields can stimulate mitosis in the corneal epithelium.\textsuperscript{28} Therefore, this same phenomenon could be presenting itself with the CECs in these experiments. Overall, our findings of maintained cellular function with exposure to SPIONPs alone in the area of viability are consistent with other SPIONP data up to 48 hours.\textsuperscript{29} There was a significant drop in viability beginning on day 2 for the magnetically exposed group at the highest concentration (100×10\textsuperscript{6} SPIONPs/Cell) and for the non-magnetically exposed on day 3. This may be due to the additional loading of SPIONPs when exposed to magnetic force. However, this outcome was not evident with any other highly loaded concentration (1×10\textsuperscript{6} SPIONPS/Cell and 10×10\textsuperscript{6} SPIONPs/Cell) and metabolic activity was still higher for nonmagnetically exposed cells on day 3 than day 1 at 100×10\textsuperscript{6} SPIONPS/cell concentrations. This suggests that in order to maintain cellular function and health, there is a maximum loading rate between 10×10\textsuperscript{6} and 100×10\textsuperscript{6} SPIONPs/cell. The doses required to induce cell death are usually high, therefore these experiments show that we have discovered an appropriate dosage to induce an apoptotic effect and to understand the limitations of SPIONP loading within CECs.\textsuperscript{30}

We sought to understand any impact of the loading of SPIONPs would have on the cytoskeletal structure of CECs. Recent work on actin filaments in fibroblast cells has suggested that they may play a key role for the process of migration and proliferation of cells.\textsuperscript{31} Therefore, understanding how CECs are influenced by loading of SPIONPs is important to the overall goal of this work. Due to the varying cellular effects based upon nanoparticle size and structure, Clyne et al compared various coatings and sizes of iron oxide nanoparticles in increasing concentration to porcine aortic endothelial cells. Results showed that bare nanoparticles caused elongation by 40 to 60\% in cells and that dextran and polyethylene glycol coated particles had no effect on cell length and did not show actin cytoskeleton disruption. However this study was only carried out for 24 hours.\textsuperscript{32} These results further justify FIGURE 5. Rhodamine Phalloidin staining confirmed that cytoskeletal structure of corneal epithelial cells was maintained at all concentrations of superparamagnetic iron oxide nanoparticles loaded in both the non-magnetic and magnetically exposed groups. The cell nuclei were counterstained with 4’,6-diamidino-2-phenylindole. Phase contrast images showed evidence of nanoparticle incorporation for cells viewed. The scale bar is 50\(\mu\)m.
(1) the use of this particular SPIONP design as an effective choice to maintain cellular behavior and (2) the future application of this work as a method for guiding endothelial cells to a targeted area will not disrupt cell structure.

Even seemingly small injuries to the ocular surface resulting from aberrations that occur to the outer and inner cornea can be incapacitating, disfiguring, painful, and blinding. Moreover, they can pose serious psychological implications for both combat soldiers and civilians. Results from our current line of study will assist in the development of human CEC research where we will investigate the effects of SPIONP concentration on cell function and morphology as well as determine if any biochemical or gene expression changes are occurring. If successful, a new treatment strategy that can deliver therapeutic cells to the inner cornea without the need for surgical intervention will be made feasible. Such novel technology will bypass current complications that arise from surgical treatments while maximizing return-to-duty rates for soldiers.

CONCLUSIONS

Prussian blue staining indicated that there was an interaction between the applied magnetic force and nanoparticle accumulation as more SPIONPs were visually apparent in magnetically treated cells. Exposure of nanoparticles to CECs had minimal impact on the viability and structure of the cells. The initial magnetic exposure had a significantly positive effect on viability which could aid in cellular repair and be a potential alternative to surgery. Moreover, the cells responded well to exposure of up to $10 \times 10^6$ SPIONPs/cell. Higher concentrations showed evidence of decreased viability, but did not influence cytoskeletal structure. Overall, there was evidence of an interaction between SPIONPs and the externally applied magnetic force. This work showed that CECs can be exposed to dextran SPIONPs with a biotin surface with little impact on viability and without disturbing cytoskeleton structure. Therefore, the feasibility of using these SPIONPs, in combination with a magnetic force at an exposure rate of up to $10 \times 10^6$ SPIONPs/cell, is accomplished and new steps may be taken to advance this as a viable treatment for ocular injuries. Thus, as we progress forward into human cell studies, the same SPIONP makeup will be utilized. In order to understand the interaction of these SPIONPs on the human corneal endothelium, human CECs will be harvested from a donor creating a primary cell line on which further cell behavior and SPIONP interactions will be investigated with in-depth loading analysis.

ACKNOWLEDGMENTS

This research was funded by the U.S. Army Medical Research and Material Command. This study was conducted under a protocol reviewed and approved by the U.S. Army Medical Research and Material Command Institutional Review Board in accordance with the approved protocol.

REFERENCES


Evaluation of a Bovine Vascular Graft in Sheep

Joseph Knapp, MD, FACC†; Amy Rizzo, DVM*; Matt Maxwell, MD, FACS†; Carlos Duran, MD, PhD*; David Cheung, PhD*

ABSTRACT The study objective was to determine safety and efficacy of a treated bovine vascular xenograft, in two Good Laboratory Practice compliant studies in sheep following carotid graft implantation. In one study, a 3- to 5-mm diameter xenograft was implanted into the right carotid artery of male sheep and compared to autologous jugular vein and a polymeric grafts similarly implanted. In a second study, a 9.5- to 14-mm diameter xenograft similarly implanted into the right carotid artery was compared to an autologous saphenous vein. Monthly Doppler ultrasound evaluation of implant patency and flow in implants and contralateral control carotid arteries was performed. The small vessel cohort 6 month xenograft patency was equivalent (or better) than animals with polymeric vascular graft or autologous vein implants; the aneurysm incidence was less than that of autologous vein grafts. In the large vessel cohort, all 15 xenografts and 12/15 saphenous vein implants were patent at 6 month follow-up. Tissue histology showed mild inflammatory responses in the xenografts that was slightly greater than suture material. In summary, treated bovine xenograft performance in this small study suggests it may be superior to polymeric autologous vein grafts, and may have a similar failure rate as autologous vein grafts after implantation.

INTRODUCTION

Timely reconstruction of battlefield vascular injuries is a significant challenge. Extremity injuries comprise 50 to 88% of U.S. armed combat-related injuries.1 Absent timely intervention, life-threatening hemorrhage, or severe limb ischemia can and often does occur.2 During Operation Iraqi Freedom and Operation Enduring Freedom, exsanguination from extremity wounds deemed potentially survivable accounted for 13.5% of all deaths.3 Local control of blood loss utilizing tourniquets has reduced hemorrhagic deaths by 85%, and has bought time for patients to receive potentially limb salvaging therapies. Standard of care for arterial injuries has been direct arterial repair/reconstruction, or when necessary, vascular interposition grafting.

Autologous veins/arteries have been the preferred conduits for extremity vascular reconstruction, with the saphenous vein from an uninjured limb the most frequently used vessel.4,5 When employed as an arterial conduit, autologous vein grafts have shown acute failure modes of infection or tissue dissolution,6,7 and long-term failure modes that vary from aneurysm formation to intimal hyperplasia, to atherosclerosis, and to thrombosis.8 In a murine model, endothelial to mesenchymal transformation leading to enhanced neointimal formation was seen in veins used as vascular grafts.9 Also important, multiply injured combat casualties frequently lack an uninjured autologous vein. Presently available synthetic conduits have not proven reliable for peripheral vessel reconstruction. In these circumstances, a reliable synthetic graft or xenograft conduit for the urgent initial grafting procedure would be of great value. Prosthetic grafts were evaluated for limb salvage in complex vascular trauma,10 and they were useful in stabilizing patients for subsequent autologous graft revascularization. Short-term patency was 79%, and complications included thrombosis, stenosis, and exposure with presumed infection. In some cases, reconstruction was not performed following prosthetic graft removal due to adequate collateral circulation.

In this article, we report a comparison of L + D-Hydro-treated xenografts with autologous vein and expanded 4-mm diameter polytetrafluoroethylene (ePTFE) grafts after vascular implantation in sheep for 6 months. Patency was assessed using serial hemodynamic interrogation with Doppler ultrasound (DUS). Histopathologic evaluation was correlated with DUS findings at the end of each study. The ultimate goal was to develop a biocompatible xenogeneic vascular graft that (1) can be stored at room temperature, (2) can be rapidly prepared for surgical deployment, (3) is of a size compatible with peripheral extremity vessels, and (4) can be easily implanted.

MATERIALS AND METHODS

Vascular Grafts

Bovine vascular grafts of 25 cm length were prepared by the L + D-Hydro process described in U.S. Patent 7,008,763. In the small vessel study series, xenografts of 3 to 5 mm inner diameter were implanted in the sheep carotid artery. Comparison was made with ePTFE grafts (Gore Medical Products, Inc., Newark, Delaware), and autologous sheep jugular veins similarly implanted. In a separate, larger vessel study series, 9.5- to 14-mm diameter xenografts implanted in the carotid artery were compared with autologous saphenous vein grafts similarly implanted. The contralateral carotid artery was utilized as control for all subject animals. The implant surgeries were directed by the same board-certified cardiovascular surgeon.

*The International Heart Institute of Montana Foundation, 500 West Broadway Street, Suite 350, Missoula, MT 59802.
†Providence St. Patrick Hospital, 500 West Broadway Street, Suite 320, Missoula, MT 59802.
‡Portions of this work were presented at Advanced Technology Applications to Combat Casualty Care, St. Pete Beach, Florida, August 13–14, 2007 and August 11–12, 2008. This article represents data collected from 2006 to 2011. doi: 10.7205/MILMED-D-15-00154
Evaluation of a Bovine Vascular Graft in Sheep

Animals
Castrated male Targhee sheep weighing on average 65 to 76 kg were used for both studies. All animal protocols were reviewed and approved by the University of Montana Institutional Animal Care and Use Committee and the Animal Care and Use Review Office. Animals were randomly assigned to treatment groups in both series of studies. In the first (small vessel) study, a power analysis was used to calculate group size; in the second (large-bore) study, additional animals were assigned to each test group.

Vein Harvest
A 4 to 5 cm segment of the left jugular vein was isolated and kept in heparinized saline (200 U heparin/mL) until implanted. A 4.5 cm section of the lateral saphenous vein was excised and placed in heparinized saline at room temperature until implanted.

Xenograft Preparation
A xenogeneic vascular graft was developed for revascularization surgery, utilizing a two-step L + D-Hydro method to reduce antigenicity and aid in tissue preservation (US Patents 7,008,763 and 6,277,555). The first step combines extraction of antigens (without detergents or digestive enzymes), employing masking (polyglycol), or chemical oxidation under physical conditions that protect the extracellular components. A second step incorporates an anti-inflammatory agent, an antithrombotic agent and cryoprotectants. The tissue conduits are then lyophilized, packaged, and sterilized with vaporized hydrogen peroxide. These xenografts can be stored at ambient temperature and rapidly hydrated in the operating theater for implantation.

The xenograft was rehydrated in 0.9% heparinized saline (200 U heparin/mL) at room temperature for 25.8 ± 7.0 minutes (range 16–41 minutes) in the small vessel study, and 22.8 ± 6.8 minutes (range 16–38 minutes) in the larger vessel study. Graft material was fully submerged in saline to ensure that no air bubbles were entrapped in the vessel lumen.

Implant Procedure
For the small vessel study, the left carotid artery was occluded and a segment excised. Before implant, animals were given heparin intravenously, and activated clotting time was measured to ensure 250 seconds or greater was achieved; Ceftriaxone was also administered. A 4.0 to 4.9 cm segment of bovine xenograft, 4.5 cm length segment of ePTFE graft, or 4.5 cm segment of autologous reverse jugular vein were anastomosed end-to-end with 6.0 Prolene. Just before tying the final suture of each distal anastomosis, the proximal vascular clamp was removed, backfilling each graft to remove air. Hemostasis was subsequently attained with final suture tying. The same process was undertaken in the larger vessel study, implanting 4.6 ± 0.7 cm length large-bore xenograft segments and 4.5 cm length autologous sheep saphenous vein segments. Animal characteristics are shown in Table I. With the exception of days on study in the autologous vein group in the small-bore study, there was little difference between treatment groups.

Graft and Carotid Artery Imaging
A Vivid-e 8L RS linear array transducer attached to a Vivid-I Ultrasound (GE Healthcare, Pittsburgh, Pennsylvania) was employed. Two dimensional imaging and DUS recordings were performed in the proximal, mid, and distal segments of all implanted grafts; similar imaging was performed on the contralateral native right carotid arteries. Images and tracings were utilized to determine patency, internal lumen diameter, peak systolic velocity (PSV), and end-diastolic velocity at each location. In the small vessel study, a baseline image was obtained preimplantation, with subsequent imaging performed immediately postimplantation, at 1 week, 2 weeks, and monthly for 6 months. For the larger vessel study, Doppler imaging was obtained immediately pre and postimplantation, at week 1, and then monthly for 6 months. The DUS imaging studies were directed by a Board Certified Echocardiologist, who trained the operators, reviewed the results and provided interpretation. Imaging recordings were limited to 1 operator for each study for a total of two operators.

Explant Procedure
All sheep were euthanized with an overdose of potassium chloride while under deep anesthesia. In the small vessel study animals, a catheter was inserted into the proximal native carotid artery, and the vessel/graft complex was rinsed with 0.9% saline, pressure-fixed with Histo-choice MB fixative for 10 minutes, photographed, excised, and placed in Histo-choice fixative for histological processing. In the larger vessel study, the xenografts and autologous saphenous vein grafts were snap frozen in liquid nitrogen after being photographed.

Histopathology
In the small vessel study, grafts were stained with hematoxylin and eosin, as well as Movat’s Pentachrome. In the larger vessel study, tissues were snap frozen in liquid nitrogen, and then sectioned in a cryotome before staining with hematoxylin and eosin or Movat’s Pentachrome. Immunohistochemistry

| TABLE I. Characteristics of Animals Receiving Vascular Grafts or Autologous Veins |
|--------------------------------|--------------------------------|-----------------|-----------------|
| Small-Bore Study |
|                  | Weight at Implanted (kg) | Weight at Explanted (kg) | Days on Study | Surgery Time (minutes) |
| ePTFE             | 64.1 ± 9.0               | 70.2 ± 7.9               | 177.2 ± 3.7   | 143.6 ± 19.5          |
| Vein              | 69.4 ± 6.2               | 74.2 ± 5.9               | 143.3 ± 45.2  | 159.5 ± 17.6          |
| Xenograft         | 66.1 ± 13.1              | 73.0 ± 15.8              | 178.6 ± 4.2   | 132.3 ± 17.1*         |
|                  | 67.4 ± 5.5               | 92.2 ± 7.7               | 182.0 ± 0.5   | 73.3 ± 15.2**         |
| Large-Bore Study |
| Vein              | 76.2 ± 4.3               | 93.6 ± 6.1               | 182.1 ± 0.4   | 116.1 ± 26.1          |
| Xenograft         | 76.4 ± 4.3               | 92.2 ± 7.7               | 182.0 ± 0.5   | 73.3 ± 15.2**         |

*p < 0.05. **p < 0.001.
staining was performed in both studies for alpha-smooth muscle actin and extrinsic nitric oxide synthase.

**Humoral Response to the Grafts**

The immunological (IgG) response to the xenograft was measured by enzyme-linked immunosorbent assay using a modification of the method of DeLustro et al. Serum samples from implant, 2 weeks, 1 month, and 6 months were evaluated for the small-bore xenografts, as well as from one autologous vein group animal.

**RESULTS**

**Small-Bore Graft Operative Times and Survival Times**

For the small vessel study, eight sheep received autologous jugular vein implants, nine received ePTFE grafts (one excluded from analysis due to infection), and 10 received xenografts (two excluded due to early postoperative graft hemorrhage). Twenty-four animals were subsequently followed. Operative times for xenograft sheep were 132.3 ± 17.1 minutes (range 102–149 minutes), for autologous vein animals 159.5 ± 17.6 minutes (range 138–177 minutes), and for ePTFE animals 143.6 ± 19.5 minutes (range 119–181 minutes). Operative times for xenograft recipient animals were significantly shorter than for the autologous jugular vein recipients (p < 0.05). Mean survival for xenograft recipients was 178.6 ± 4.2 days, 177.9 ± 3.4 days for the ePTFE graft recipients, and 143.3 ± 45.2 days for the autologous saphenous vein recipients. Early deaths (two euthanized, and one spontaneous) among the jugular vein recipients were secondary to graft aneurysms (Fig. 1).

**Small-Bore Graft DUS Imaging**

DUS imaging showed distinctive patterns among the three cohorts. The autologous vein recipients showed gradual graft dilatation in the mid-graft section, with three animals developing graft aneurysms (Fig. 2 mean ± SEM). Mean graft diameters were relatively stable at the proximal and distal anastomoses, although the ePTFE graft diameter was consistently lower than the xenograft or autologous graft. The mid-graft diameter of the autologous vein became significantly greater than the ePTFE and xenografts. Using repeated measures analysis of variance (ANOVA), a statistically significant difference (p < 0.05) between the xenograft and ePTFE grafts was observed at the mid and distal regions, and between the vein and ePTFE mid regions. At 6 months, 4/5 surviving sheep had patent grafts, with all four patent grafts showing subtly reduced flow velocities and (qualitatively) increases in flow turbulence. Histology of the patent vein grafts revealed a thickened layer of smooth muscle cells and luminal endothelialization. The ePTFE cohort grafts were all occluded (8/8) at 6 months with a fibrocollagenous plug.

**Small-Bore Graft Arterial Blood Flow**

The xenograft behaved differently than the other grafts, and 5/8 xenografts remained patent throughout the study. Of the three failed grafts, all appeared to fail early, within 2 weeks. Two xenografts appeared to be technical failures, damaged by side branch hemoclips placed during graft processing. For the five surviving and patent grafts, hemodynamic flows were similar to those seen in the control contra-lateral carotid arteries. Figure 3 illustrates PSV from sequential DUS measurements. Flows in the vein and xenograft were similar, but by 3 months, flow had ceased in the ePTFE group. Although significant inter- and intra-animal variation was observed for both patent jugular vein recipients and xenograft recipients, both groups showed a subtle trend to declining peak Doppler flow velocities over time in the mid and distal graft regions. Statistically, significant differences in PSV were seen from repeated measures ANOVA between the xenograft and ePTFE grafts in the mid and distal regions, the vein and ePTFE grafts in the proximal and distal regions, and within the xenografts at the proximal and mid and proximal and distal regions.

**Small-Bore Graft Histopathology and Immune Response**

Histological analysis found smooth muscle layered with endothelium in the xenografts, with lesser degrees of intimal hyperplasia than was seen in the autologous vein graft implants. A statistically significant decrease in flow was seen in the ePTFE
group, compared to the autologous vein and xenograft groups \((p < 0.05)\) by repeated measures ANOVA.

A transient rise in humoral markers was seen by enzyme-linked immunosorbent assay early with the small-bore xenografts that normalized by 6 months. Animals that experienced elevated IgG within 2 months of implant had no correlation with extent of chronic inflammation in the xenograft animals seen after 6 months by histopathology (scores ranged from 0 to 2: none to mild) that was similar to that seen with the ePTFE graft (scores ranged from 0 to 1: none to minimal).

**Large-Bore Graft Operative Times and Survival Times**

For the larger vessel study, 15 animals were enrolled in each group, one receiving treated bovine xenografts, the other autologous saphenous vein implants. Operative time was significantly \((p < 0.001)\) longer for the autologous saphenous vein animals \((116.1 \pm 26.1\) minutes\) than for the xenograft recipients \((73.3 \pm 15.2\) minutes\) as shown in Table I. One of the vein recipients died 1 day postimplant. At time of implant, xenografts measured 4.6 \(\pm 0.7\) cm in length, with an outer diameter of 9.2 \(\pm 1.3\) mm. The implanted saphenous vein grafts measured 4.5 \(\pm 0.8\) cm in length, and 5.9 \(\pm 0.8\) mm in diameter. All remaining animals survived to necropsy at 6 months (Table I). At explant (Fig. 4), both xenografts and saphenous veins had dilated (xenografts 12.8 \(\pm 2.3\) mm and saphenous veins 11.7 \(\pm 3.9\) mm); none had become aneurysmal. All xenografts remained patent throughout the 6-month study, with one of the surviving saphenous vein grafts occluding immediately postimplantation.
Both xenografts and saphenous vein graft implants showed parallel patterns of gradual graft dilatation over the 6-month study (Fig. 5), particularly in the mid-graft regions. Internal luminal diameters (xenografts 11.4 ± 2.1 mm and saphenous veins 11.4 ± 3.2 mm) were effectively the same at explant. Statistically (repeated measures ANOVA), significant differences \( (p < 0.05) \) were observed in both xenograft and vein mid-graft regions, compared to their corresponding proximal regions.

**Large-Bore Graft DUS Interrogation**

Similar to the patterns seen in the small vessel study, significant inter and intra-animal variability in peak Doppler flow peak systolic velocities occurred over the course of the study (Fig. 6). Fourteen out of fifteen xenograft recipients maintained physiologic flow throughout the course of the study, with one animal (988) showing minimal flow after 6 months. Twelve out of fifteen saphenous vein recipients maintained physiologic flow throughout the course of study (one animal was occluded and two animals had minimal flow). Statistically (by repeated measures ANOVA), a significant difference \( (p < 0.05) \) was seen only between the distal and proximal vein grafts.

**Large-Bore Graft Histopathology**

Histopathologic evaluation revealed a greater degree of inflammatory cell infiltration into the larger vessel xenografts than saphenous vein autografts. This was primarily because of lymphocytic cell infiltration into the adventitia and outer media. Neutrophils, macrophages, and giant cells were present at anastomotic sites and in side branch suture lines. Minimal to mild mineralization was noted in 11 xenograft anastomosis sections and in three vein graft anastomosis sections in areas of mural trauma associated with sutures. Severe localized mural mineralization was seen in one xenograft section, with a second xenograft evidencing mild, diffuse mineralization of the media. No similar findings were seen in any of the saphenous vein grafts. Fibrous neointima was present in 41 of 45 xenograft sections and in 27 of 34 vein graft sections.

**FIGURE 4.** Representative images of large-bore grafts at explant. A xenograft is shown in panel A and vein in panel B; a right carotid artery from animal 988 is included for comparison in panel C. Grafts were covered by a fibrin coat (dissected away in these images), with the vein grafts firmly encased.

**FIGURE 5.** Autologous saphenous veins and large-bore xenograft mid-graft inner diameters measured by Doppler Ultrasound at time points up to 6 months.
The overall inflammatory response to the xenograft was no greater than that of biocompatible sutures used for anastomoses.

**DISCUSSION**

Commercially available synthetic grafts are available in a variety of sizes, lengths, and configurations, and can be tailored to address a specific vascular injury. Unfortunately, synthetic grafts are prone to occlusive pseudo-intimal hyperplasia and infection, with small diameter synthetic grafts additionally compromised by a high likelihood of thrombosis. For this reason, Fox et al concluded that use of synthetic grafts for combat-related vascular injuries of the extremities “is uniformly associated with poor outcomes and their use is to be discouraged.” More recent data suggest that synthetic grafts might function for transient limb salvage following vascular trauma, as a bridge to autologous vein graft insertion.

Xenograft vessels prepared by various methodologies have proven of limited clinical value due to a propensity to thrombosis, aneurysm formation, and vascular tissue degradation. The xenogeneic matrix promotes inflammatory responses that result in foreign body reaction, tissue degeneration, and graft failure.

In summary, vascular grafts represent a commonly implanted medical device, for which no universally reliable graft conduit presently exists. All commercially available grafts, including autologous venous grafts, are prone to early graft failure. In the absence of a truly reliable product, alternative options (i.e., heterologous, synthetic, and xenogeneic) should be explored. This study provides evidence that an L + D-Hydro-treated bovine graft should be considered as an alternative to autologous vein grafts or polymeric grafts.

**Hemodynamics**

Bovine xenografts processed with the L + D-Hydro technique remained patent and performed as well as, or better than, autologous veins or ePTFE graft materials when implanted in the carotid artery. These results contrast with results from prior studies utilizing bovine arteries treated with glutaraldehyde fixation. The lyophilized L + D-Hydro-processed bovine xenografts allow these grafts to be stored dry at ambient temperature, and rehydrated after 15 minutes of immersion in room temperature saline. Time to implant with the xenografts was significantly less than that for autologous veins, even when the time for vein harvest is excluded. Although the smaller-bore bovine xenografts were more prone to occlusion than larger-bore xenografts, both proved as reliable or more reliable than the alternative conduits of similar size that were studied, with no tendency toward the aneurysmal dilatation seen with autologous vein implants. The major problem seen in this study with implanted xenografts appeared to be related to a processing flaw, i.e., tissue damage from misapplied side branch clips. This flaw should prove remediable with alternative processing techniques. Both small- and large-bore graft cohorts evidence gradual dilatation over time, without the aneurysmal dilatation seen in the smaller vessel groups.

**Histology**

The xenografts were recellularized by host cells, with evidence for both smooth muscle and endothelial cell ingrowth. These bovine xenografts, had a transient IgG humoral response that returned to baseline by 6 months. These data lead one to conclude that the L + D-Hydro process is effective at reducing the host immune response, making the bovine graft both biocompatible and durable. Inflammatory cell infiltration was more common among the xenograft implants, as was tissue mineralization, albeit mild. The degree of inflammation was similar to that seen with suture materials.
Surgical Implantation

Surgical implant times for the xenografts were significantly shorter than for other materials. Besides excellent biocompatibility, bovine xenograft conduits have advantages of off-the-shelf availability, room temperature storage, simplified shipping, and shortened implant times. The graft material is easy to handle and use, can be tailored to the particular vascular application, and is suturable without tearing. These characteristics make L + D-Hydro-treated bovine xenografts a potentially valuable resource for military field hospitals, with expanded applications in noncombat-related vascular injuries as well, military and civilian.

It remains unclear as to whether L + D-Hydro-treated vascular grafts will prove valuable as a long-term vascular conduit, or as a bridge therapy to definitive repair. Alternative tissue-engineered grafts are undergoing clinical trials, but the time required to prepare such grafts limits their use in acute vascular injury. Future studies may include seeding of xenografts with autologous stem cells, which might well improve endothelialization without restenosis. Unfortunately, such products would likely not be able to be held at room temperature, and certainly would not prove useful in emergent, urgent clinical scenarios. Such circumstances might better be served by L + D-Hydro-treated xenografts, at least as interim therapy.

CONCLUSIONS

After 6 months, the surviving small vessel bovine xenografts showed overall hemodynamic performance that was superior to autologous jugular vein as measured by absence of aneurysm formation and ePTFE grafts in patency that was similar to the contralateral carotid controls. For larger bore vessels, patency and flow dynamics were similar in the xenograft and autologous saphenous vein cohorts. In summary, performance of the treated bovine xenograft in this small study suggests it may be superior to polymeric autologous vein grafts, and may have a similar failure rate as autologous vein grafts after implantation. Additional studies are needed to evaluate the statistical and clinical significance of the new graft material.

ACKNOWLEDGMENTS

The authors would like to acknowledge the writing and editing assistance of Milton Marshall, PhD, International Heart Institute Montana Foundation and review of article by Timothy Descamps, BS, CRA, International Heart Institute Montana Foundation. No additional compensation was received for such contributions. United States Army Medical Research and Material Command contracts W81XWH-06-1-0569 and W81XWH-10-1-0808.

REFERENCES

Exploratory Laparotomy for Proximal Vascular Control in Combat-Related Injuries

MAJ Emily Hathaway, MC USA*; LCDR Jacob Glaser, MC USN†; CPT Cassandra Cardarelli, MC USA*; James Dunne, MD*; CAPT Eric Elster, MC USN‡; CDR Shawn Safford, MC USN‡; CAPT Carlos Rodriguez, MC USN*

ABSTRACT Background: Combat casualties have endured an increase in traumatic lower extremity amputations secondary to improvised explosive devices. Often surgical control of the proximal vasculature is required. We evaluate the safety profile of exploratory laparotomy (EXLAP) for proximal control (PC) in combat-injured patients. Methods: Records of 845 combat casualties from June 2009 to December 2011 were reviewed. Patients undergoing EXLAP were divided by indication into PC and non-PC groups. Demographics, Injury Severity Score, mechanism of injury, transfusion requirements, EXLAP findings, reoperation rates, and abdominal-related complications were recorded. Results: 44 patients were identified as PC and 91 as non-PC. Age was similar (23.7 ± 4.1 vs. 24.0 ± 4.6, p = 0.7138), and all were male. improvised explosive devices blast was the most common mechanism of injury. Injury Severity Score (25.8 ± 8.2 vs. 21.4 ± 9.1, p = 0.0075), lower extremity amputation (93.1% vs. 28.6%, p = 0.0001), and transfusion requirements were different. Days to fascial closure (1.8 ± 1.9 vs. 1.7 ± 2.8, p = 0.8308) and number of EXLAPs were similar (2.4 ± 1.3 vs. 2.1 ± 1.5, p = 0.2581). PC had higher complications (43.1% vs. 24.2%, p = 0.0292). Conclusion: PC demonstrated an increase in abdominal complications. The reason for this remains unclear. Alternative approaches of achieving proximal vascular control may avoid the morbidity associated with laparotomy, and predeployment training of such procedures should be considered for the general surgeon. Further studies are warranted to determine best practices for these patients.

INTRODUCTION

Since 2010, combat casualties related to Operation Enduring Freedom (OEF) occurring in Afghanistan have endured a marked increase in traumatic lower extremity amputations secondary to increased use of improvised explosive devices (IED). After injury, patients are rapidly evacuated via medical air support teams to nearby facilities with surgical capability where they undergo damage control surgery with the primary goal of hemorrhage control. This may occur at outlying care facilities where resources are limited, or casualties may be brought directly to more robust combat hospitals where advanced damage control measures may be implemented.

The Joint Trauma System’s Clinical Practice Guideline entitled, “Management of High Bilateral Amputations,” recommends achieving vascular control as distal as possible. However, the powerful blast mechanism of modern IEDs often results in extensive proximal bony and soft tissue damage precluding the effective use of combat or pneumatic tourniquets for immediate hemostasis, adequate wound exposure, and surgical debridement. In these cases, surgical approaches to the femoral vessels are necessary through the groin or to the iliac vessels by either extraperitoneal or transabdominal means.

These approaches are not uncommon, as suprainguinal control has been necessary in approximately 20% of combat-related traumatic amputations. Poon et al. reported an acceptable intraoperative safety profile for surgically achieving suprainguinal, transabdominal proximal control (PC).

In trauma, laparotomies are routinely performed and their associated morbidity has been well documented. Rates can be significant and have been reported between 15% and 50%. Technological developments, such as resuscitative endovascular balloon occlusion the aorta (REBOA), have made it possible to obtain suprainguinal PC without trans-abdominal. In deciding whether alternatives such as REBOA are advantageous, we believe it is prudent to study the morbidity profile of the transabdominal approach for PC in the postoperative period.

METHODS

All combat casualties evacuated to a single U.S. military facility from June 2009 through December 2011 were identified, and records were retrospectively reviewed. Patients who underwent exploratory laparotomy (EXLAP) downrange

*Walter Reed National Military Medical Center, 8901 Rockville Pike, Bethesda, MD 20889.
†Division of Trauma and Surgical Critical Care, Department of Surgery, School of Medicine, R Adams Cowley Shock Trauma Center, University of Maryland, 22 South Greene Street, Baltimore, MD 21201.
‡Uniformed Services University of Health Sciences, 4301 Jones Bridge Road, Bethesda, MD 20814.

Portions of this work were presented at the 72nd annual meeting of the American Association for the Surgery of Trauma, September 18, 2013, the U.S. Navy Committee on Trauma research competition, May 28, 2014, and the Military Health System Research Symposium, Fort Lauderdale, FL, August 19, 2014.

The views expressed are those of the author and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, or the U.S. Government. I am a military service member (or employee of the U.S. Government). This work was prepared as part of my official duties.

Title 17, USC, §105 provides that “Copyright protection under this title is not available for any work of the U.S. Government.” Title 17, USC, §101 defines a U.S. Government work as a work prepared by a military service member or employee of the U.S. Government as part of that person’s official duties.

doi: 10.7205/MILMED-D-15-00155
EXLAP for Proximal Vascular Control in Combat-Related Injuries

Patients were then divided into subsets based on primary indication for initial laparotomy: PC and those who underwent laparotomy for an alternative indication (non-PC [nPC]). Patients who were foreign nationals and those who died before discharge from index hospitalization were excluded.

Documentation from all care levels was reviewed. Data collected included patient age, injury severity score (ISS), mechanism of injury, associated injuries, number and type of blood products transfused, indication for laparotomy, operative findings, and days to discharge. Also collected were total number of exploratory laparotomies required, total days to fascial closure, and those who required a reoperation after the initial fascial closure. Abdominal complications were considered if they were directly attributable to the laparotomy and only if they occurred after the initial fascial closure. These included ileus, small bowel obstruction, abdominal wall dehiscence, and intra-abdominal abscess formation. Ileus was defined as nausea and vomiting requiring nasogastric decompression.

To determine significance and odds ratios, data were compared with Fisher exact and two-tailed Student t tests where applicable. Kuskal–Wallis chi-square test and binary logistic regression were used where applicable. All data were collected under an existing OEF/Operation Iraqi Freedom Institutional Review Board protocol, “Caring for Casualties,” a comprehensive database designed to capture casualty data in theatre and across the casualty evacuation system to the United States.

RESULTS

Demographics

Of the 845 combat casualties evacuated to our U.S. military hospital, 145 were identified to have undergone laparotomy downrange. A total of 10 were excluded from analysis for missing operative reports (5), death before discharge (4), and extensive injury resulting in loss of abdominal wall precluding fascial closure (1), leaving 135 (16.0%) for evaluation.

Forty-four patients underwent laparotomy for PC, and 91 patients underwent laparotomy for another indication (nPC). Characteristics of the study cohort are listed in Table I. Age was similar between the two groups (23.7 ± 4.1 vs. 24 ± 4.6, p = 0.7138), and all casualties were male. The most common mechanism of injury was IED blast, injuring 66.4% of patients overall. In the PC group, 89.1% of patients were injured by IED blast, 2.17% by gunshot wound (GSW), 2.17% by mortar/rocket propelled grenade (RPG), and 6.5% were injured by some other mechanism. The nPC group’s distribution of injury patterns were different (p < 0.05 for all), 56.0% of patients were injured by IED blast, 24.4% by GSW, 14.4% by RPG blast, and 5.6% by some other mechanism. When comparing PC to nPC, ISS (25.8 ± 8.2 vs. 21.4 ± 9.1) and presence of lower extremity amputation (93.1% vs. 28.6%) were significantly different (p = 0.0075 and 0.0001, respectively).

Indication/Operation Performed

All patients in the PC group underwent laparotomy for vascular control; there were no incidental findings or occult injuries identified. After the initial operation, 72.7% (32/44) of these patients were left open. Vascular control was obtained at the common iliac vessels or below, with 75.0%, 65.9%, and 59.1% of patients achieving control at the level of the right common iliac, left common iliac, or bilateral common iliac arteries. Within the nPC group, the most common indication for laparotomy was positive computed tomography findings, followed by penetrating mechanism, and hypotension on presentation. Eight patients were explored for a positive focused assessment with sonography in trauma (FAST) examination and five for a positive diagnostic peritoneal lavage (DPL) (Table II). Findings at exploration fell within four categories: “intra-abdominal injury” (solid organ or small bowel injury), 56%; colon injuries, 16%; rectal injuries, 12%; and negative explorations, 35% (Table III). Interestingly, of the five patients who were explored for a positive DPL, four had negative explorations.

Blood Transfusion

In the PC group, 97.8% of patients underwent massive transfusion. They received a mean of 37.5 ± 30.3 units of packed red blood cells (pRBCs), 35.6 ± 26.7 units of fresh frozen plasma (FFP), 4.9 ± 5.9 units of cryoprecipitate, 5.2 ± 4.1 units of platelets, and 1.5 ± 3.3 units of whole blood. The nPC group received significantly less blood products (Table IV), 44.4% of patients underwent massive transfusion, and received a

<table>
<thead>
<tr>
<th>TABLE I. Patient Demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
</tr>
<tr>
<td>Number</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>ISS</td>
</tr>
<tr>
<td>MOI</td>
</tr>
<tr>
<td>GSW</td>
</tr>
<tr>
<td>RPG/Mortar</td>
</tr>
<tr>
<td>Other</td>
</tr>
<tr>
<td>Lower Extremity</td>
</tr>
</tbody>
</table>

ISS, injury Severity Score; MOI, mechanism of injury; IED, improvised explosive device; GSW, gunshot wound; RPG, rocket propelled grenade.
Operations

Both the PC and nPC groups underwent a similar number of total explorations 2.5 ± 1.2 vs. 2.1 ± 1.5, respectively (p = 0.1247). There was no difference in the numbers of patients with open abdomens either on arrival to Landstuhl Regional Medical Center (41.3% vs. 31.1%, p = 0.2524) or on arrival to the United States (8.7% vs. 8.9%, p = 1.0000). Time to fascial closure was also similar between the PC and nPC, 1.8 ± 1.9 days vs. 1.6 ± 2.8 days, (p = 0.2000). All patients were able to be closed in a primary or delayed fashion. Patients requiring component separation were similar in each group (2.2% vs. 1.1%, p = 1.0000). None required mesh (Table V). Of the 44 patients in the PC group, 24 (54.5%) received colostomy formation.

Intra-Abdominal Complications

Intra-abdominal complications were significantly higher for the PC group (43.1% vs. 26.4%, p = 0.0292), Chi-Square odds ratio 0.45 (95% confidence interval = 0.21, 0.95). Similarly, more patients suffered two or more complications in the PC group 11.4% vs. 1.1% (p = 0.0141). Removing ileus as a complication still revealed a 22.7% complication rate in the PC group vs. a 9.9% rate in the nPC group, though this was not statistically significant (p = 0.0622). Abdominally related reoperation rates were similar for the PC group (29.5% vs. 19.8%, p = 0.2747). Table VI illustrates the distribution of complications between the groups.

A subgroup analysis of the PC group, those with colostomy vs. those without, did not demonstrate a significant difference in the complication rates (41.7% vs. 45%, p = 1.000), respectively. Similarly, removing ileus did not change complication rate significantly (25.0% vs. 20.8%, p = 1.000).

DISCUSSION

Recent reports regarding OEF casualties suggest that 1 in 5 traumatic lower extremity amputations require suprainguinal vascular control and the transabdominal approach is an option associated with little morbidity. However, our data demonstrate a postoperative abdominal-related complication rate of

### TABLE II. Indication for Laparotomy

<table>
<thead>
<tr>
<th>Injury Type</th>
<th>Proximal Control (n = 91)</th>
<th>Nonproximal Control</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Abdomen</td>
<td>44 (100)</td>
<td>0</td>
<td>0.0001</td>
</tr>
<tr>
<td>Colon Injury</td>
<td>26 (28.9)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rectal Injury</td>
<td>15 (16.6)</td>
<td>0</td>
<td>0.0001</td>
</tr>
<tr>
<td>Negative Laparotomy</td>
<td>12 (13.3)</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### TABLE III. Findings at Exploration

<table>
<thead>
<tr>
<th>Injury Type</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra-abdominal Injury</td>
<td>56</td>
</tr>
<tr>
<td>Colon Injury</td>
<td>16</td>
</tr>
<tr>
<td>Rectal Injury</td>
<td>12</td>
</tr>
<tr>
<td>Negative Laparotomy</td>
<td>35</td>
</tr>
</tbody>
</table>

### TABLE IV. Average Initial Transfusion Rates Over 24 Hours Following Injury. Massive Transfusion Defined As Having Received Over 10 Units of pRBC

<table>
<thead>
<tr>
<th>Component</th>
<th>Proximal Control</th>
<th>Nonproximal Control</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>pRBC (Units)</td>
<td>37.5 ± 30.3</td>
<td>12.2 ± 13.2</td>
<td>0.0001</td>
</tr>
<tr>
<td>FFP (Units)</td>
<td>35.3 ± 26.7</td>
<td>11.6 ± 14.0</td>
<td>0.0001</td>
</tr>
<tr>
<td>PLT (6 Packs)</td>
<td>5.2 ± 4.1</td>
<td>1.6 ± 2.5</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cryoprecipitate (Units)</td>
<td>4.9 ± 5.9</td>
<td>1.8 ± 4.0</td>
<td>0.0004</td>
</tr>
<tr>
<td>Whole Blood (Units)</td>
<td>1.5 ± 3.3</td>
<td>1.4 ± 4.1</td>
<td>0.8842</td>
</tr>
<tr>
<td>Massive Transfusion</td>
<td>43/44 (97.83%)</td>
<td>40/491 (44.07%)</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

### TABLE V. Patient Surgical Outcomes. This Table Compares Time to Closure, Number of Exploratory Laparotomies (EXLAP) Needed to Approximate Abdominal Fascia, and Need for delayed primary closure (DPC), Component Separation, and Colostomy Formation

<table>
<thead>
<tr>
<th>Component</th>
<th>Proximal Control</th>
<th>Nonproximal Control</th>
<th>p Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXLAP to Fascial Closure</td>
<td>2.5 ± 1.2</td>
<td>2.1 ± 1.5</td>
<td>0.1247</td>
</tr>
<tr>
<td>Open Abdomen (n, % of total) on Arrival to Landstuhl Regional Medical Center</td>
<td>18 (41.3%)</td>
<td>28 (31.1%)</td>
<td>0.2524</td>
</tr>
<tr>
<td>Time to Fascial Closure (Average No. of Days)</td>
<td>1.8 ± 1.9</td>
<td>1.6 ± 2.8</td>
<td>0.2000</td>
</tr>
<tr>
<td>Primary Fascial Closure (n, % of Total)</td>
<td>12 (27.2%)</td>
<td>42 (46.2%)</td>
<td>0.0406</td>
</tr>
<tr>
<td>DPC (n, % of Total)</td>
<td>31 (70.5%)</td>
<td>48 (53.7%)</td>
<td>0.0238</td>
</tr>
<tr>
<td>Component Separation</td>
<td>1 (2.3%)</td>
<td>1 (1.1%)</td>
<td>1.0000</td>
</tr>
<tr>
<td>Colostomy Formation</td>
<td>24 (54.5%)</td>
<td>26 (28.6%)</td>
<td>0.0044</td>
</tr>
</tbody>
</table>

pRBC, packed Red Blood Cells; FFP, fresh frozen plasma; PLT, platelets.

mean of 12.2 ± 13.2 units pRBCs, 11.6 ± 14.0 units FFP, 1.8 ± 4.0 units cryoprecipitate, 1.6 ± 2.5 units platelets, and 1.4 ± 4.1 units of whole blood. However, the mean number of units of whole blood did not significantly differ (1.5 ± 3.3 vs. 1.4 ± 4.1, p = 0.8842).
such an approach of over 43%. Notably, the complication rate in the PC group was nearly twice that of the nPC group (Table VI), and multiple complications (2 or more in one patient) were 10 times more likely. Fear of a missed injury may prompt some surgeons to prefer laparotomy, but our PC had no incidental intra-abdominal injuries noted on exploration. Although it is critical to minimize missed injuries, if other adjuncts (FAST, mechanism, DPL, etc.) rule out injury, it may be considered safe to perform an extraperitoneal approach. Our data suggest that avoiding a laparotomy in these patients, in favor of alternate methods of vascular control and colostomy formation, may help to avoid these complications.

Bograd et al. studied similar combat injured patients undergoing damage control laparotomy, finding they were at higher risk for abdominal complications versus those undergoing nondamage control laparotomy. In addition, they found the rate of complications in all patients undergoing damage control surgery was 49%. This was found to be consistent with reported civilian equivalents (ranging from 15% to 63%) complication rates. Previously reported studies have complications that can often be attributed to concomitant intra-abdominal injury (perforation of viscera, pancreatic injury, prolonged time to abdominal closure, etc.), but this is not the case in our PC group.

There are two main reasons we believe the overall complication rate in the PC group is higher. First, ISS is known to under predict the severity of injury in cases of traumatic amputation. The PC group, in our study, had an amputation rate of 93.1% vs. 28.6% in the nPC group, suggesting that in addition to the higher ISS in the PC group, these patients may have sustained an even more significant total body injury burden than implied by ISS alone. This may be the primary reason for the increased complications, but should also serve as an argument against a transabdominal approach for vascular control.

Survivability has been optimized with advances in prehospital care, tourniquet use, sophisticated transfusion strategies, and damage control techniques, despite increasing ISSs. If we accept that a laparotomy puts these patients at even higher risk of complications, which our data suggest, then we should avoid the laparotomy if not absolutely necessary. It is reasonable to assume that avoidance of a laparotomy will avoid the postoperative complications of bowel obstruction, ileus, and midline wound dehiscence.

The second plausible explanation for the higher complication rates in the PC group is greater amount of blood transfusions required for resuscitation and its associated immunomodulating effects. In 2004, Dunne et al. studied transfusion in trauma patients and documented a two-fold to nearly six-fold increase in systemic inflammatory response syndrome and a significantly increased risk of postoperative infection in patients who were transfused during their hospital course. Further, in 2009 Dunne et al. demonstrated large volume transfusion to be associated with increased infection rates and impaired overall wound healing in combat trauma. Additional studies have shown multiple transfusions carry an increased, independent, dose-related risk of infectious processes and contribute to multiorgan failure, prolonged intensive care unit admission, and increased length of stay. Although this may be unavoidable given the more extensive patterns of injury in this patient group subset, efforts should be made to limit the number of blood transfusions to what is absolutely necessary to maintain homeostasis, hemostasis, and perfusion.

In these patients, fecal diversion was often indicated to minimize contamination of lower extremity and pelvic soft tissue wounds and was frequently associated with transabdominal proximal vascular control. In fact, a recent study performed on soldiers returning to Walter Reed National Military Medical Center indicated fecal diversion remains the preferred standard for combat-related rectal injuries. This practice has shown to be associated with lower mortality rates among the injured with fecal diversion when compared to those without. Our subgroup analysis of the PC group, those with colostomy vs. those without, did not demonstrate a significant difference in the complication rates (41.7% vs. 45%, p = 1.000), and supports its safety in combat-related injuries. However,
if extraperitoneal approaches can be successfully implemented for vascular control, alternative approaches to colostomy formation can be considered. Laparoscopic methods or left lower quadrant incisions may be a viable alternative and should be further studied.\textsuperscript{18,19}

Regardless of the underlying etiology, a procedure with an associated complication rate of 43\% is high, and alternative approaches should be sought. Techniques for achieving vascular control without directly violating the peritoneum are available and should be considered. This would avoid bowel manipulation and introduction of foreign material thus eliminating the opportunity for intra-abdominal adhesions to develop and reducing the likelihood of ileus and small bowel obstruction.\textsuperscript{20} These techniques would also eliminate the risk for midline wound infection, abdominal wall dehiscence, and potential evisceration.

Although circumstances and injury patterns can occur in combat that call for laparotomy where expedient, effective control of hemorrhagic shock is critical to survivability of the injured, we suggest the use of a transabdominal approach, specifically for PC, to be a calculated judgment with consideration of its postoperative safety profile. Effective alternative approaches exist, and have robust training associated with them. The American College of Surgeons offers the Advanced Surgical Skills for Exposure in Trauma (ASSET) course, a key component of which is extraperitoneal approaches to the iliac vessels.\textsuperscript{21}

Intuitively, an extraperitoneal approach would be associated with less abdominal morbidity; however, its safety profile has not been fully studied in combat trauma. Nonetheless, this approach can be performed expeditiously and the authors recommend universal adoption of ASSET as part of predeployment training for surgeons.

The recent development and evaluation of potential combat uses of REBOA should be logistically evaluated in the deployed environment.\textsuperscript{8} While data are preliminary, and the logistics of placement in the setting of traumatic amputation are still being evaluated, it is a promising alternative to a transabdominal approach. This technique may also be considered for inclusion in the predeployment training algorithm, by inclusion of the Basic Endovascular Skills for Trauma (BEST) course.\textsuperscript{22}

CONCLUSION

Combat trauma patients requiring laparotomy for proximal vascular control of hemorrhage are at higher risk for complications than patients who require laparotomy for other surgical indications. The deployed general surgeon should have familiarity with retroperitoneal vascular access, as well as endovascular approaches, to provide optimal care for the combat-injured patient. Our study demonstrates that these skills can significantly improve the safety and recovery of our injured troops and training in such should be implemented and required in the predeployment training. With courses such as ASSET and BEST being universally available, efforts should be made to evaluate the feasibility of these courses to the military population.

JG, EH, and CR designed the study. JG, CC, and EH searched the literature. EH, CC, JD, SS, EE, and CR collected and analyzed the data. JG, CC, JD, EH, SS, and CR participated in the manuscript writing and critical revision.

REFERENCES


Intra-Arterial Perimortem Resuscitation Using a Micellar Colloid

Cuthbert Simpkins, MD*; Krishna Talluri, MD*; Mallory Williams, MD†

ABSTRACT  Objective: The following were studied in a perimortem mouse model of rapid blood loss: (a) efficacy of a prototypical micellar colloid, Intralipid 20%, (IL20), compared to albumin (b) comparison of intra-arterial and intravenous resuscitation, (c) efficacy of IL20 at a volume 2 x the volume of blood removed, and (d) efficacy of oxygenated IL20 after clinical death (CD). Methods: CD, the absence of breathing and zero blood pressure (BP), was produced by removing 55% of the blood volume within 3 minutes. After CD, the chest was opened to observe ventricular contraction. IL20, Ringer’s lactate (RL), or albumin was infused perimortem. Results: Without resuscitation CD occurred in 2.85 ± 0.40 minutes. Ventricular contraction persisted 20.50 ± 1.11 minutes after CD. RL infused immediately after CD restored breathing if given intra-arterially but not intravenously. IL20 was superior to the prototypical colloid, albumin in maintaining the BP. Increasing the volume of IL20 further increased BP. Delayed RL infusion after CD failed to restore breathing. Delayed resuscitation after CD with oxygenated IL20 restored breathing and BP. Conclusions: Micellar colloid is superior to the prototypical colloid albumin and can possibly be of use when signs of life are no longer present. In extremis, intra-arterial infusion is superior to intravenous infusion.

INTRODUCTION

Exsanguination is the cause of 91% of potentially survivable battlefield deaths.1 Fifty-one percent of soldiers who die on the battlefield are encountered in cardiorespiratory extremis.2 Most battlefield deaths occur within 10 minutes of injury (http://www.nhs.uk/Livewell/Militarymedicine/Pages/Surviving battlefield.aspx). There is an unmet need for effective interventions in the 10 minute perimortem time frame sometimes referred to as “the platinum 10 minutes” as opposed to the “golden hour” (http://www.army.mil/article/55508/Battlefield_medicine_and_the_urgency_to_save_Soldiers/).

To address this, we developed a new mouse model of perimortem state by rapid and massive blood removal, because current animal models are not representative of the soldier who has massive bleeding on the battlefield, which is when most of the losses occur. In current animal models, the animals are intubated to control respiration, and body temperature is controlled. This model is analogous to a wounded soldier in the operating room (OR) where these factors can be controlled. This is not the case on the battlefield. The soldiers who have rapid and large-volume blood loss due to major organ or vessel injury often do not survive to get to the OR. Or the evacuation times are longer because of hostile surroundings, which make slower bleeding critical. For this reason in our experiment, the animals were not intubated. Also neither the animals nor the resuscitation fluids were warmed. These experiments were designed to determine how to salvage soldiers who do not survive with the current technology.

In general, for resuscitation, colloids are favored over crystalloids. This is because colloids raise and maintain blood pressure (BP) with a smaller volume and for a longer time. This enables paramedics to carry less weight as fluid onto the battlefield. For a soldier on the battlefield who has a weak pulse due to blood loss, U.S. Military Tactical Combat Casualty Care guidelines call for the rapid intraosseous or intravenous infusion of 500 mL of a colloid, 6% hetastarch (Hextend; Hospira, Lake Forest, Illinois), when blood products are not available. If there is no improvement in the pulse, an additional 500 mL bolus of Hextend is infused (Military Health System and Defense Health Agency Website: http://www.health.mil/tccc).

However, concerns have been raised about the safety of hetastarch-based (HES) products. A Cochrane review of the effect of hetastarch on renal function concluded that:

“The current evidence suggests that all HES products increase the risk in AKI (acute kidney injury) and RRT (renal replacement therapy) in all patient populations and a safe volume of any HES solution has yet to be determined. In most clinical situations it is likely that these risks outweigh any benefits, and alternate volume replacement therapies should be used in place of HES products”.

Increased mortality,4,5 and increased bleeding because of dysregulation of clotting factors,6,7 are the concerns with hetastarches. In order to reduce the possibility of bleeding, the military protocol limits the volume of Hextend administration to 1,000 mL.7 The validity of this upper dose limit in trauma patients is not established and the limit is possibly

*Vivacelle Bio, Inc., Enterprise Works Chicago, 2242 West Harrison Street, Chicago, IL 60612.
†Division of Trauma and Critical Care, Department of Surgery, Howard University, 2041 Georgia Avenue NW, Suite 4502, Washington, DC 20060.
(Dr. Williams is also a member of the scientific advisory board of Vivacelle Bio, Inc.)

Portions of this work were presented at the Advanced Technology Applications for Combat Casualty Care (August 13–16, 2012) and the Military Health System Research Symposium (August 18–21, 2014), both held in Ft. Lauderdale, FL.

Animal procedures were approved by The Animal Care and Use Committee of the LSU Health Sciences Center at Shreveport. The authors are officers of the company, Vivacelle Bio, Inc., (VBI). VBI was formed in order to develop products that are based on the findings of this article. These findings were made before the formation of VBI or any other commercialization effort.

doi: 10.7205/MILMED-D-15-00206
even lower given the increased bleeding tendency associated with trauma. The Food and Drug Administration (FDA) has issued a warning against the use of HES products in critically ill patients due to safety concerns.8

The 1,000-mL limit also creates a significant logistical challenge, which threatens the military medical mission. Evacuation times in future wars could be much longer than has been the case in recent conflicts. This could require multiple boluses with volumes well beyond 1,000 mL to maintain pulse and mentation.

Therefore, the military has an unmet need for a more flexible and effective resuscitation fluid with broader applications. Stated otherwise, a safer and more effective resuscitation fluid for saving the lives of our military service men and women is urgently needed.

In order to meet this need, we are developing a new class of colloid for resuscitation that consists of components that have been shown to be safe and effective when introduced into the bloodstream. The colloid of this fluid is a micelle that contains highly purified nonallergenic soybean oil. Highly refined soybean oil is not considered allergenic by the FDA.9 A micelle is comprised of molecules in which one segment is soluble in water and the other segment is soluble in oil. When such molecules are placed into water and mixed they spontaneously form an aggregate known as a micelle as illustrated in Figure 1. The water-soluble part of the molecule (black) faces outward, whereas the oil soluble tail (dark green) faces inward forming a hydrophobic core. When soybean oil (light green) is added to the mixture, it localizes to the oily inner part of the micelle, which is the hydrophobic core.

The properties of this colloid are consistent with a safer and more effective resuscitation fluid. The FDA approved this colloid for intravenous use in total parenteral nutrition in 1972. Examples of these commercially available products are Intralipid (Baxter International, Deerfield, Illinois) and Liposyn (Hospira, Lake Forest, Illinois). Intralipid at a soybean oil concentration of 20 g/100 mL Intralipid 20% (IL20) has also been rapidly infused for local anesthetic toxicity10 and toxicity associated with overdose of verapamil.11 This supports the hypothesis that the micelles would be safe in hypotension due to blood loss. In these initial experiments, we did not add balancing electrolytes because as a baseline we wanted to determine the effect of unaltered emulsion. Isotonicity is provided by 2.25% w/v glycerol in Intralipid so that the osmolality of 260 is approximately that of Ringer’s lactate (RL). Adding electrolytes would have created a hypertonic infusate.

We previously showed in mice that IL20 was superior to RL, the resuscitation fluid recommended by the American College of Surgeons, in raising and maintaining the BP after the removal of 55% of the blood volume in mice. In addition, we showed that there was a linear relationship between the oxygen content of the emulsion and soybean oil concentration. When exposed to 100% oxygen, the oxygen content of Intralipid at 30% soybean oil was equivalent to the oxygen available from blood at a hemoglobin level of 12 g/dL. At 20% soybean oil, the oxygen content was equivalent to that of blood at a hemoglobin level of 7 g/dL. We found that the oxygen was rapidly released from the emulsion. In another study, blood was removed from mice and a mean BP of 30 mm Hg was maintained for 90 minutes. Resuscitation was then carried out using either shed blood or IL20. At 24 hours after resuscitation with IL20, there was no evidence of fat emboli in the lungs (hematoxylin/eosin stain).12 In these current pilot studies, we proceeded to determine how the micellar colloid would be deployed in the perimortem period.

**METHODS**

**Animals and Animal Procedures**

The Animal Care and Use Committee of the LSU Health Sciences Center at Shreveport, Louisiana, approved the animal procedures. Male and female mice weighing 27 to 47 g were utilized. Mice were anesthetized using ketamine/xylazine administered intraperitoneally. After making sure the mice were well anesthetized, the internal jugular vein and/or the carotid artery were cannulated and 55% of blood volume was removed over 3 minutes. At various time points after blood removal, test fluids in volumes equal to one or two times the volume of blood removed were infused over 2 minutes. BP was measured at the carotid artery using a BP-2 monitor made by Columbus Instruments (Columbus, Ohio), which measures the BP as a voltage. A standard curve was prepared. Measured voltages were converted to BP using the following formula:

\[ BP = \frac{(Voltage - 0.1006)}{0.0107} \]
In order to mimic battlefield conditions, warming measures and respiratory support were not applied to the mice. In spite of the absence of a controlled physiological environment, there was little variability in the effects of severe rapid blood loss and anesthesia. This was evidenced through a study of the pH, PO2, and PCO2 after the induction of anesthesia and before the removal of blood. These data were obtained from 50 mice. In these mice, we found mean ± SE values of pH = 7.15 ± 0.01, PO2 = 115.68 ± 2.49, PCO2 = 51.86 ± 1.25, and base excess = −9.822 ± 0.4744. In each mouse, we noted the time of clinical death (CD) as defined by loss of respirations and BP. After CD, the chest was opened and the time it took for ventricular contraction to cease was observed. In current animal models, the respirations would be controlled as if in an OR. Our intention in these experiments was to observe cardiac contraction in the uncontrolled environment that exists on or near the battlefield. The effect of intravenous and intra-arterial RL infusion was compared in preliminary experiments. In these experiments, 55% of the blood volume was removed and the mice were observed until breathing ceased. Immediately after the cessation of breathing, the mice were given RL either intravenously or intra-arterially. None of the five mice given RL intravenously resumed respirations. In contrast, six of six mice given RL intra-arterially resumed respirations for various times after infusion. These experiments revealed that in extremis intra-arterial infusion was more effective than intravenous infusion. Therefore, a carotid artery catheter only was placed and infusions were given intra-arterially in all subsequent experiments.

**MATERIALS**

NaCl, human albumin, and heparin were obtained from SIGMA-Aldrich (St. Louis, Missouri). IL20 was obtained from SIGMA-Aldrich (or Baxter) RL was prepared in the laboratory with the salts obtained from SIGMA-Aldrich. It was comprised of deionized distilled water, 109 mM NaCl, 28 mM, Na (L) lactate, 4 mM KCl, and 1.5 mM CaCl2.

**Statistical Analysis**

Unpaired two-tailed Student’s t test was used to compare the difference between two mean BP s. Two-way analysis of variance was used to evaluate differences in BP response curves between two different volumes of IL20. Nonparametric data were analyzed using the Mann–Whitney U test. Significance was set at p < 0.05.

**RESULTS**

There was no significant difference in the initial BP achieved after resuscitation with either IL20 or 5% albumin. However, 1 hour after the initial increase, the mean BP of the mice that received 5% albumin was 59.15 ± 2.95% (n = 6) of the prehemorrhage pressure. In contrast, the mean BP of mice 1 hour after the infusion of IL20 was 73.2 ± 2.78% (n = 6) of the prehemorrhage pressure. These mean values were significantly different with \( p = 0.0060 \). The difference in pressures may be translatable to the clinical situation in which a mean pressure of 65 mm Hg is considered adequate while a pressure below 60 mm Hg is not. These findings are illustrated in Figure 2.

Figure 2 shows representative changes in BP observed after replacing 55% of the blood volume with either 5% albumin or IL20 immediately after the removal of blood. The volume of replacement fluid was equal to the volume of blood removed.

From left to right, there are three phases. In the first phase, BP is recorded before removal of blood. In the second phase, the result of removing 55% of blood volume over 3 minutes is seen. This resulted in a mean pressure ±SE of 10.488 ± 1.474 in mice that were later given 5% albumin and shows a high initial pressure, and decline to below prehemorrhage pressure. (B) Comparison of BP response after infusion of 5% albumin in normal saline to infusion of IL20. The graph depicts BP vs. time after infusion of 5% albumin and shows a high initial increase, and decline to pressure within the range of prehemorrhage pressure.

![FIGURE 2.](image)
hemorrhage BP as shown in Figure 2B. The mice that received albumin and the mice that received IL20 were observed for an average of 2.41 ± 0.008 hours and 3.23 ± 0.669 hours after the infusion, respectively. There was no significant difference in these times.

Next, the BP attained after giving IL20 in a volume double the volume of removed blood was determined and compared to the BP obtained after the infusion of a volume of IL20 equal to the removed blood volume. The rationale of this experiment was to determine the effect of giving an excess of the emulsion. Our prediction was that it would raise the BP higher than lesser volumes of the emulsion. However, because there were no previous reports of the emulsion being given after severe and rapid blood loss, we were concerned about the possibility of cardiodepressant effects and decrease in BP such as that observed when perfluorocarbons are rapidly infused.14,15

Systolic pressures are shown in Figure 3A and diastolic pressures are shown in Figure 3B. The difference between the curves for 2 × vs. 1 × the removed blood volume was significant to the level of p < 0.001. As predicted, the BP increased with the administration of a higher volume of the emulsion.

**Infusion After CD**

We found that CD, as indicated by the loss of respiration and measurable BP, occurred within 2.85 ± 0.40 minutes (n = 75 mice) of removal of blood. After CD, the chest was opened to observe cardiac ventricular contraction. Ventricular contraction persisted for a mean duration of 20.50 ± 1.11 minutes (n = 93 mice) after CD. We compared the effect of administering RL intra-arterial or intravenously immediately after CD. The chest was not opened in these experiments. Intra-arterial resuscitation restored breathing in 6 of 6 mice, albeit with a high degree of variability in post resuscitation survival time. In contrast, intravenous resuscitation after CD failed to restore breathing in five of five mice. Using the Man–Whitney U test, the p value was 0.0055 for the difference in the median values. This suggests intra-arterial resuscitation was superior to intravenous resuscitation. Survival times are listed in Table I.

Next, instead of infusing fluid immediately after removal of blood we delayed intra-arterial infusion of RL for 1 (n = 2) and 3 minutes (n = 2) after CD. RL infusion failed to restore respiration 1 and 3 minutes after CD. However, intra-arterial infusion of oxygenated IL20 5 minutes after CD resulted in the restoration of spontaneous respiration in six of six mice. Oxygenation was achieved by bubbling 100% oxygen through

**TABLE I.** Length of Breathing Time After Infusion (Minutes)

<table>
<thead>
<tr>
<th></th>
<th>Intra-Arterial</th>
<th>Intravenous</th>
</tr>
</thead>
<tbody>
<tr>
<td>165.00</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>6.78</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>4.00</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>20.00</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>5.00</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3.00</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Mean ± SE</td>
<td>33.96 ± 26.33</td>
<td>0</td>
</tr>
</tbody>
</table>

**TABLE II.** Data Showing Results After 55% of Mouse Blood Volume was Removed Over 3 Minutes

<table>
<thead>
<tr>
<th>Mouse Number</th>
<th>Survival Time After Clinical Death (Minutes)</th>
<th>1st Peak Blood Pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.5</td>
<td>58/13</td>
</tr>
<tr>
<td>2</td>
<td>77.1</td>
<td>81/69</td>
</tr>
<tr>
<td>3</td>
<td>97.8</td>
<td>35/28</td>
</tr>
<tr>
<td>4</td>
<td>103.0</td>
<td>88/61</td>
</tr>
<tr>
<td>5</td>
<td>111.1</td>
<td>134/83</td>
</tr>
<tr>
<td>6</td>
<td>114.8</td>
<td>125/115</td>
</tr>
<tr>
<td>Mean ± SE</td>
<td>84.2 ± 17.41</td>
<td>86.83 ± 15.53 (Systolic)</td>
</tr>
</tbody>
</table>

Mice were observed until they stopped breathing. 5 minutes after the cessation of respiration, oxygenated IL20 was infused over 2 minutes. In six of six trials, mice infused with oxygenated IL20 had restored respiration and BP. Mouse numbers 1 and 3 survived up to 1.5 and 97.8 minutes, respectively. Mouse numbers 2, 4, 5, and 6 were euthanized at 77.1, 103, 111.1, and 114.8 minutes, respectively. Mouse number 2 was euthanized after awakening completely.
IL20 for 2 minutes. The survival times of the 6 mice are shown below in Table II.

Mouse number 1 and 3 survived up to 1.5 and 97.8 minutes, respectively. Mouse numbers 2, 4, 5, and 6 were euthanized at 77.1, 103, 111.1, and 114.8 minutes, respectively. Mouse number 2 was euthanized after awaking completely.

DISCUSSION
Over the past 2 decades, there have been multiple efforts to develop effective and safe resuscitation fluids and “blood substitutes.” The following are the results of those efforts and their associated problems.

— Hetastarch-based products reported to cause renal failure and increased mortality. FDA warns against use in critically ill patients. Increased bleeding tendency compels the use of no more than 1,000 mL. This number is not established in trauma patients.
— Perfluorocarbons did not fare well in clinical trials. Produces hypotension and an intense inflammatory response when rapidly infused. These products are not FDA approved.
— Modified hemoglobin such as Hemopure increased mortality in clinical trials, very expensive to produce. This product is not FDA approved.
— Plasma requires blood typing. Is a cause of the transfusion-related acute lung injury; increases mortality and can transmit disease. Anaphylactic reaction is possible.
— Whole blood requires blood typing may use O negative initially. Can cause transfusion-related lung injury; can transmit disease. Supply is very limited and blood is expensive. Anaphylactic reaction possible.
— Hypertonic saline causes kidney failure and increased mortality.
— Albumin increases bleeding; expensive; supply is limited, can transmit disease.

An ideal colloidal resuscitation fluid would be effective and safe; would not promote bleeding and could be given in large volumes. It would maintain a viable BP and mentation in the event of a delay in evacuation without the fear of disease transmission. Also, there would not be any inherent supply limitations. The fluid would be inexpensive and have an extended shelf life. We now report experiments that demonstrate a prototypical micelle preparation, IL20, can be the basis of an effective resuscitation fluid after the loss of blood. Evidence in support of the safety of micellar colloids is found in our previous report in which mice survived 24 hours after resuscitation and there was no evidence of lung injury or fat embolism. In addition, there are numerous reports of rapid infusion of IL20 as an antidote to the overdose of local anesthetics and other medications. In contrast to the hetastarch products and albumin that have been shown to enhance bleeding, IL20 has been shown to preserve clotting. Other benefits are that the formulation based on IL20 will not transmit disease, has a long shelf life of ≥18 months, is available in unlimited supply and is fairly inexpensive to manufacture. Other properties of IL20 that enhance its attractiveness as the basis of a resuscitation fluid are that it has been shown to inhibit reperfusion injury, activate prosurvival genes and can be metabolized to produce energy. These and other properties could reduce the incidence of multiple organ dysfunction syndrome post resuscitation.

The greater potency relative to RL and albumin in raising and/or maintaining the BP suggests that IL20 would be a very effective low-volume resuscitation fluid that would minimize the weight of fluid a paramedic needs to carry onto the battlefield. Higher concentrations of soybean oil in the formulation would potentially have even greater efficacy.

These pilot experiments also provide insight into the transition between life and death.

One of the earliest reports of the advantage of intra-arterial over intravenous resuscitation was that of Seeley in 1951. Dr. Sam F. Seeley was a Brigadier General in the Medical Corps and Chief of Surgery at the Walter Reed Army Hospital. We have shown that after CD with the absence of pulse and respirations, there is a period of 20 minutes during which the heart continues to contract, though these contractions are inadequate to produce a measurable BP. Restoration of breathing and BP with intra-arterial infusion is superior to the intravenous route possibly because arterial delivery produces better perfusion of the coronary and bronchial arteries than venous delivery. This approach could be useful for nontraumatic as well as traumatic cardiac arrest. In Advanced Cardiovascular Life Support, the immediate goal is to improve coronary artery perfusion. The rapid intraarterial infusion of micellar colloids could markedly improve the attainment of this goal.

Access to the arterial system could be gained via the femoral vessels using the pulse or if there is no pulse using a handheld ultrasound.

IL20 has worked well in these and previous experiments as a prototypical micellar colloid containing soybean oil. Although our results have shown it to be effective, it is not ideal in its present form for use as a resuscitation fluid. One reason is that it lacks electrolytes. Replacement of a significant portion of the blood volume with a fluid that does not contain sodium could result in hyponatremia and related complications. In experiments, now in progress, we have added electrolytes and in addition reduced the diameter of the micelle thereby increasing its stability. “We do not consider our micellar preparation to be a blood substitute. In our view, our emulsion is a resuscitation fluid with many properties that are potentially useful for rescuing patients in shock. Two properties, which we explored in an earlier publication, are the ability of the emulsion to reversibly absorb oxygen and nitric oxide. We found that a 20% soybean oil emulsion can be loaded with sufficient oxygen to support life. Moreover, the reversible absorption of nitric oxide would enable the emulsion to reverse the drop in BP caused by excessive nitric oxide that is released in shock after blood
loss. In this experiment, we established the necessity of the intra-arterial infusion of resuscitation fluid in order for soldiers with extreme blood loss to survive. In addition, we have begun the development of a new resuscitation fluid that is superior to the standard colloid albumin in maintaining the BP and that does not have the adverse effects of albumin and other colloids in use today.”

REFERENCES


Screening of Nanoemulsion Formulations and Identification of NB-201 as an Effective Topical Antimicrobial for *Staphylococcus aureus* in a Mouse Model of Infected Wounds

**Yongyi Fan, MS*; Susan Ciotti, PhD†; Zhengyi Cao, MD*; Rone Eisma, MS‡; James Baker Jr., MD*;‡; Su He Wang, MD*‡**

**ABSTRACT**

Despite advances in antimicrobial therapies, wound infection remains a global public health concern. We aimed to formulate and assess various nanoemulsions (NEs) for potential effectiveness as stable antimicrobial agents suitable for topical application. A total of 106 NEs were developed that varied with respect to nonionic and cationic surfactants. Stability testing demonstrated that the NEs tested are broadly stable, with 97/106 formulations passing 2-week stability tests. Two NEs, NB-201 and NB-402, were selected to test antimicrobial activity in a wound model in mice. Skin abrasion wounds were infected with *Staphylococcus aureus* followed by NE treatment. Infected skin was then evaluated by measuring colony forming units. NB-201 reduced median bacterial counts by 4 to 5 log compared to animals treated with saline, whereas NB-402 reduced bacterial counts by 2 to 3 log. Additional stability tests on NB-201 demonstrated that NB-201 is stable in the presence of human serum, and is stable for at least 6 months at 5°C, 25°C, and 40°C. Finally, in vitro studies, NB-201 was found to be effective against *S. aureus* at a higher dilution than the commercially available silver sulfadiazine. Altogether these results demonstrate that NB-201 is a stable and effective topical antimicrobial for the treatment of *S. aureus*.

**INTRODUCTION**

Wound infections are a significant concern for the U.S. military. Wounds resulting from high-energy gunshot, blast, and burn injuries can affect large areas of skin, which is then susceptible to bacterial or fungal infection. In fact, the majority of mortality following battlefield wounds is related to infection rather than osmotic shock and hypovolemia, and the risk of sepsis following local wound infection is increased by delays in receiving front line emergent medical or surgical intervention. In addition, critically wounded soldiers often undergo multiple invasive surgical procedures, long-term intensive care and prolonged hospitalization, increasing the risk of nosocomial infection.

The standard of care for acute wounds includes wound cleaning and may include the application of topical antimicrobials with or without an occlusive dressing. Moist wound care aimed at maximizing retention of water with aqueous products and occlusive dressings is often used for chronic wounds and is increasingly applied to acute wounds including minor lacerations, abrasions, and superficial burns. The development of an antimicrobial to be sprayed onto acute wounds to provide a moist environment, prevent infection, and treat developing infections would greatly enhance combat-related wound care.

Nanoemulsions (NEs) are novel oil-in-water emulsions formulated from nontoxic materials that have been demonstrated to have broad antimicrobial properties following topical application. The emulsions are made using shear conditions to produce high-energy droplets that are thermodynamically driven to fuse with lipids in microbial outer membranes, leading to membrane destabilization and lysis of pathogens, while leaving healthy skin intact. This selective cytotoxicity arises from the structure, size, and mechanism of action of the NE particles. When optimal ratios of oil, water, solvent, and surfactant are emulsified under a water interface, the resultant particles have high potential energy by virtue of their small size and large surface area. The small size of the particles allows them to pass through pores in the skin and mucous membranes, but not through intercellular junctions that surround epithelial cells; as a result, droplets applied to the skin do not coalesce on the surface, but rather are quickly delivered into the skin. Particles concentrate at the site of an infection and surround infectious organisms, where they are thermodynamically driven to fuse with the outer membrane of the pathogen, leading to membrane destabilization and death.

The natural toxicity of NEs to infectious agents, lack of irritation to healthy skin and mucous membranes, retention of activity at high temperatures, and ease of portability make them an ideal candidate for use in the combat setting. Here we summarize the synthesis, biophysical characterization, and short-term stability of more than 100 NE compounds. One formulation, NB-201, has previously been shown to be effective in a rat burn injury model. Here we further test NB-201...
and a second formulation, NB-402, against *staphylococcus aureus* in a mouse model of infected wounds.

**MATERIALS AND METHODS**

**Reagents**

Unless otherwise indicated, all reagents were purchased from Sigma-Aldrich Corp. (St. Louis, Missouri).

**NE Formulations**

A total of 106 propriety NE formulations were manufactured at a 500 g scale by varying five different cationic surfactants and/or the ratio of cationic to nonionic surfactants (surfactant blend) in the NE formulation within a particular surfactant family. All compounds contained 20 mM ethylenediaminetetraacetic acid in the external aqueous phase. NB-201 utilizes Benzalkonium chloride (BZK) (Stepan Chemical Company, Northfield, Illinois) as the cationic surfactant with Tween 20 (Croda, Mill Hall, Pennsylvania) as the nonionic surfactant. For NB-402, Cetylpyridium chloride (CPC) (Vertellus, Zeeland, Michigan) is the cationic surfactant and Poloxamer 407 (P407) (Spectrum Chemical Manufacturing Corporation, New Brunswick, New Jersey) as the nonionic surfactant. Placebo formulations lack the cationic surfactant. All excipients are considered “Generally Recognized as Safe” by the U.S. Food and Drug Administration (FDA).

**Short-Term Stability Testing**

An aliquot of each NE compound was stored in a glass vial undisturbed at 22°C and 40°C for 2 weeks. Dynamic light scattering using the Malvern Zetasizer (Malvern Instruments, Worcestershire, United Kingdom) was used to determine particle size, particle size distribution profiles, and polydispersity index (PdI). Acceptance criteria was a change in mean particle size of no more than 30%, PdI of less than 0.25, and appearance of no visible phase separation.

**Long-Term Stability of NB-201**

Long-term stability of NB-201 was assessed under various temperature and relative humidity (RH) conditions (5°C, 25°C/60% RH, 30°C/65% RH, and 40°C/75% RH) over 6 months using the criteria described above. In addition, high-performance liquid chromatography was used to measure the potency of BZK, with acceptance criteria of 90 to 110% BZK label claim.

**Stability Testing of NB-201 in the Presence of Human Serum**

A mixture of human serum and Mueller Hinton broth (1:1) was mixed with NB-201 to produce a 1% final concentration of NE in the mixture and allowed to incubate for 24 hours at 37°C. The particle size, pH, and zeta potential were taken before and after incubation at 37°C.

**Bacterial Strains and Culture Conditions**

Minimal inhibitory concentrations (MICs) were determined using a modification of the Clinical and Laboratory Standards Institute approved serial dilution method as described by LiPuma et al. Formulations were tested in a dilution range of 1:1 to 1:2,048. Subsequent two-fold NE dilutions were made using Mueller Hinton broth such that the final volume in all wells was 150 μL. For experiments to test the effect of serum, 50 μL of heat-inactivated human serum was added followed by 10 μL of diluted bacteria. The reaction was incubated and treated with resazurin (R&D Systems, Minneapolis, Minnesota), and MICs were recorded as the least concentrated dilution in which the wells remained blue.

**Mouse Skin Abrasion Model**

CD-1 mice were housed in specific pathogen-free conditions, and all experiments were approved by the University Committee on the Use and Care of Animals at the University of Michigan. Under anesthesia, the dorsal fur on mice weighing approximately 25 to 30 g was shaved and exposed skin was washed with 70% ethanol. The disinfected skin was then abraded using a sterile 28-gauge needle to create a 6 × 6 grid within a defined 1.5 × 1.5 cm² area. After 5 minutes, a 50 μL suspension of 1 × 10⁷ colony forming units (CFU) of *S. aureus* in sterile saline was inoculated on to the abraded area and immediately covered with a 1.5 × 1.5 cm² Telfa pad and then further covered with a Tegaderm occlusive dressing. After 3 hours, dressings were removed and wounds on three animals were sprayed each with 600 μL of NB-201, NB-402, NB-201 placebo lacking BZK, NB-402 placebo lacking CPC, or sterile saline. Wounds were redressed with Telfa and Tegaderm, and NE treatments were repeated 16 hours later. Animals were euthanized 48 hours after the first NE treatment, and tissue at the abraded site was immediately excised for analysis.

**Bacterial Quantification**

Excised tissues (100 mg) were mechanically homogenized with 0.9% NaCl, serially diluted, and plated in duplicate on Edge nutrient agar plates. Plates were incubated for 24 hours at 37°C and CFUs were counted.

**Statistical Analysis**

Data were analyzed by student’s *t* test or one-way analysis of variance followed by student’s *t* test. *p* values of less than 0.05 were considered statistically significant.

**RESULTS**

**Stability of NE Formulations**

A total of 106 stable NE formulations were created at NanoBio Corporation (Ann Arbor, Michigan) by varying the surfactant blend ratio of six different cationic surfactants in...
In combination with three different types of nonionic surfactants (Tweens, Spans, and Poloxamers). These NE compositions varied by mean particle-size profiles and other biophysical properties. Each formulation was evaluated for stability based on physical appearance and particle size over a range of times and temperatures. Altogether 97 of the formulations tested passed all stability criteria after 2 weeks at both 22°C and 40°C (Table I).

### Antimicrobial Activity of NB-201 and NB-402 in Mouse Skin Abrasion Wounds Infected With S. aureus

To assess the antimicrobial activity of NB-201 and NB-402, a mouse wound model was developed in which a superficial wound was created on the dorsal skin of the mouse by light abrasion. These wounds were infected with *S. aureus* and then treated twice with sterile saline, NB-201, NB-402, or NE placebo controls. Two days later, the animals were sacrificed and CFU from the tested skin tissue was counted. Figure 1 shows that the NB-201-treated mice showed a 4 to 5 log reduction in CFU as compared to saline-treated mice and a 3 to 4 log reduction as compared to placebo-treated mice. NB-402 was also effective against *S. aureus*, and showed a 3 to 4 log and 1 to 2 log reduction in CFU, as compared with saline-treated and placebo-treated mice, respectively. CFUs of NB-201-treated mice were also much lower than that of NB-402-treated mice. We note that the NB-201 formulation containing BZK has a larger polar head group but a shorter tail than the CPC used in NB-402 (Table II). The structural differences may account for differences in the observed antimicrobial activity of the two compounds.

### NB-201 is Effective at a Higher Dilution than Silver Sulfadiazine In Vitro

The antimicrobial activity of NB-201 against *S. aureus* was compared to that of silver sulfadiazine, a topical antimicrobial agent that is widely used for the prophylaxis and treatment of skin wound infections.13,14 MICs were determined for both NB-201 and silver sulfadiazine using the serial dilution method described by LiPuma.12 These studies showed that 2 different batches of NB-201 were effective at inhibiting *S. aureus* at a 1:512 dilution, whereas silver sulfadiazine required a much higher dose (1:16 dilution) to achieve a comparable effect (Table III). When 25% serum was included in the assay, the effectiveness of both products was reduced, but NB-201 remained effective at concentrations of 1:16 to 1:32, whereas silver sulfadiazine required a concentration of 1:8 to exhibit the same antimicrobial activity. These results indicate that at the concentrations tested, NB-201 may be much more effective in reducing *S. aureus* than the commonly used silver sulfadiazine.

### NB-201 is Stable in the Presence of Human Serum

To determine whether the physical characteristics of NB-201 are stable in the presence of serum, the PdI, particle size, and zeta potential were measured in the presence of different concentrations of human serum, which is known to have adverse effects on the efficacy of antimicrobials.15 Table IV shows that NB-201 was stable with respect to mean particle size and PdI in the presence of 6.25, 12.50, and 25% human serum. As the percentage of human serum was increased to 25%, the zeta potential of the formulation was reduced and the mean particle size became larger, but the changes were not significant as compared to broth only. In contrast, a reduction of both the zeta potential and mean particle size and upregulation of PdI was observed when the percentage of serum was 50%. These data suggest that NB-201 retains its physical characteristics in the presence of at least 25% serum.

### NB-201 Stable for at Least 6 Months at Different Temperatures

Experiments demonstrating that NB-201 is effective in the treatment of skin wounds infected with *S. aureus* in mice,
and that NB-201 retains its physical characteristics in the presence of serum, suggest that NB-201 may be a highly effective agent for the prevention and/or treatment of skin wounds in both battlefield and civil settings. To determine whether NB-201 additionally meets requirements of product stability over a wide range of temperature and humidity conditions, NB-201 was tested for long-term stability. In these experiments, NB-201 maintained physical properties and potency for 6 months under refrigerated (5°C), room temperature (25°C), and accelerated temperature (40°C) conditions (Table V).

**DISCUSSION**

In this study, a panel of 106 NE formulations was derived by variation of the cationic surfactant, nonionic surfactant, and surfactant blend. This panel of NEs was characterized for biophysical properties and stability over a range of human serum concentrations, and the vast majority (97/106) were found to be stable.

NB-201 has previously been shown to be highly effective in decreasing bacterial growth in infected partial-thickness burn wounds. To test whether NE formulations are similarly effective in wounds in which the skin is damaged but not burned, we developed a skin abrasion wound model in mice in which the skin is damaged by scratching to damage the stratum corneum and upper layer of the epidermis, but not the dermis. These wounds were then infected with *S. aureus*, followed by treatment with NB-201, NB-402, placebo formulations, or saline. Our experiments demonstrated that in this model of infected wounds, treatment with either NB-201 or NB-402 significantly reduces bacterial counts as compared to placebo controls or saline control.

In these experiments, the antimicrobial ability of NB-201 was more effective than NB-402. This difference in antimicrobial effect may be attributable to different nano-components in the formulations, as the former contains BZK and Tween 20 whereas the latter contains CPC and P407. BZK and CPC are cationic surfactants and Tween 20 and P407 are nonionic surfactants, resulting in a larger polar head group and a shorter tail for NB-201 as compared with NB-402 (Table II). How

### TABLE II. Structural Differences in Cationic and Nonionic Surfactants in NB-201 and NB-402

<table>
<thead>
<tr>
<th>Surfactant</th>
<th>Surfactant Type</th>
<th>Molecular Weight (g/mol)</th>
<th>Hydrophilic Lypophilic Balance</th>
<th>Hydrophobic Chain Length</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>NB 201</td>
<td>Cationic</td>
<td>354</td>
<td>24.0</td>
<td>12, 14, 16, 18 mixture</td>
<td></td>
</tr>
<tr>
<td>Poloxamer 407</td>
<td>Nonionic</td>
<td>12,600</td>
<td>22.0</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>Polyoxyethylene Sorbitan Monolaurate (Tween 20)</td>
<td>Nonionic</td>
<td>1,226</td>
<td>16.7</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Benznalkonium Chloride</td>
<td>Cationic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cetylpyridinium Chloride</td>
<td>Cationic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE III. Comparison of NB-201 to Silver Sulfadiazine In Vitro

<table>
<thead>
<tr>
<th>Species</th>
<th>NB-201</th>
<th>Silver Sulfadiazine</th>
<th>NE Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Staphylococcus aureus</em></td>
<td>1:512</td>
<td>1:16</td>
<td>1:1</td>
</tr>
<tr>
<td><em>Staphylococcus aureus + 25% Serum</em></td>
<td>1:16</td>
<td>1:32</td>
<td>1:1</td>
</tr>
</tbody>
</table>

### TABLE IV. NB-201 Stability in Presence of Human Serum

<table>
<thead>
<tr>
<th></th>
<th>Zeta Potential (mV)</th>
<th>Mean Particle Size (nm)</th>
<th>Polydispersity Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>14</td>
<td>277.2</td>
<td>0.126</td>
</tr>
<tr>
<td>Broth</td>
<td>13.4</td>
<td>288.3</td>
<td>0.125</td>
</tr>
<tr>
<td>Broth + 6.25% Serum</td>
<td>11.8</td>
<td>288.4</td>
<td>0.109</td>
</tr>
<tr>
<td>Broth + 12.50% Serum</td>
<td>10.2</td>
<td>300.5</td>
<td>0.156</td>
</tr>
<tr>
<td>Broth + 25.00% Serum</td>
<td>8.7</td>
<td>363.7</td>
<td>0.163</td>
</tr>
<tr>
<td>Broth + 50.00% Serum</td>
<td>5.8</td>
<td>132.2</td>
<td>0.517</td>
</tr>
</tbody>
</table>

**Screening of Nanoemulsion Formulations and Identification of NB-201**
these differences in components and structures affect the antimicrobial ability of these two nano-compounds will require further investigation.

In a previous study, NB-201 was shown to be effective against the gram-negative *Pseudomonas aeruginosa* in infected burn wounds. These results suggest that NB-201 is a potential pan-antimicrobial agent for both gram-positive and gram-negative bacteria. This is important since skin wounds are frequently associated with multi-bacterial infections of both gram-positive and gram-negative species.\(^{16,17}\)

To compare the antimicrobial effects of NB-201 with the FDA-approved product silver sulfadiazine, MIC assays were conducted to determine the lowest concentration at which each product remained effective at inhibiting growth of *S. aureus*. These experiments were conducted both in the presence of human serum to mimic a wound setting and in the absence of human serum. In both cases, NB-201 inhibited microbial activity at a higher dilution than the commercially available silver sulfadiazine.

In addition to acting as a highly effective antimicrobial, NB-201 has several other advantages as a potential agent for the treatment of topical wounds in both battleground and civil settings. The components of NB-201 include purified water, soybean oil, ethylenediaminetetraacetic acid, glycerol, Tween 20, and BZK. Each of these components is inexpensive and recognized as Generally Recognized as Safe described by FDA.\(^{19}\) Thus, the production of NB-201 for use as a topical antimicrobial will not likely be constrained by cost or safety concerns.

Another important requirement for an antimicrobial in an acute wound setting is stability. We demonstrate here that NB-201 is stable for at least 6 months across a range of temperatures, including 5, 25, and 40°C. In contrast, the unbuffered sodium hypochlorite previously developed for wound infection has a shelf life of 6 days.\(^{19}\) Furthermore, we have shown here that NB-201 is stable in the presence of 25% human serum, which suggests that it will retain its antimicrobial properties even after contact with the blood and debris that are common in a wound setting.

Altogether the results of this study suggest that NB-201 has the appropriate characteristics for the topical treatment of acute wounds. NB-201 is a highly effective antimicrobial in a *S. aureus*-infected skin abrasion wound in mice, is inexpensive to produce, and is stable for at least 6 months across a wide temperature range and in the presence of human serum. Furthermore, in vitro studies suggest that it is effective at a higher dilution than the commercially available silver sulfadiazine. Further studies will determine whether NB-201 is similarly effective against a wider range of relevant bac teria, including methicillin-resistant *S. aureus* infections in full-thickness burn and traumatic wounds. The logistical requirements for the product to be effective/stable for a long (>6 months) periods of time as well as cost per use of the product will also need to be determined.

### ACKNOWLEDGMENTS

Wendy Lockwood Banka contributed to the writing of this manuscript. This project has been funded in whole or in part with Federal funds from the Department of Defense Congressionally Directed Medical Research Programs Award W81XWH-11-2-0005 (DM102308).

### REFERENCES


---

**TABLE V.** NB-201 Stability at Different Temperatures and Lengths of the Storage

<table>
<thead>
<tr>
<th>Storage Condition</th>
<th>Storage Interval (Months)</th>
<th>Appearance</th>
<th>Benzalkonium Chloride Potency (% Label Claim)</th>
<th>pH</th>
<th>Mean Particle Size (nm)</th>
<th>Polydispersity Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>—</td>
<td>Initial</td>
<td>Pass</td>
<td>101.5</td>
<td>4.7</td>
<td>313.7 ± 3.6</td>
<td>0.127 ± 0.005</td>
</tr>
<tr>
<td>5°C</td>
<td>1</td>
<td>Pass</td>
<td>99.8</td>
<td>4.7</td>
<td>317.2 ± 4.3</td>
<td>0.131 ± 0.018</td>
</tr>
<tr>
<td>5°C</td>
<td>2</td>
<td>Pass</td>
<td>103.2</td>
<td>4.6</td>
<td>314.7 ± 6.9</td>
<td>0.119 ± 0.006</td>
</tr>
<tr>
<td>5°C</td>
<td>3</td>
<td>Pass</td>
<td>100.2</td>
<td>4.7</td>
<td>315.5 ± 2.7</td>
<td>0.116 ± 0.037</td>
</tr>
<tr>
<td>25°C/60 %RH</td>
<td>1</td>
<td>Pass</td>
<td>100.1</td>
<td>4.7</td>
<td>315.1 ± 0.7</td>
<td>0.108 ± 0.020</td>
</tr>
<tr>
<td>25°C/60 %RH</td>
<td>2</td>
<td>Pass</td>
<td>102.1</td>
<td>4.6</td>
<td>317.2 ± 4.5</td>
<td>0.140 ± 0.008</td>
</tr>
<tr>
<td>25°C/60 %RH</td>
<td>3</td>
<td>Pass</td>
<td>100.2</td>
<td>4.6</td>
<td>313.6 ± 4.7</td>
<td>0.172 ± 0.017</td>
</tr>
<tr>
<td>40°C/75 %RH</td>
<td>1</td>
<td>Pass</td>
<td>99.7</td>
<td>4.6</td>
<td>316.7 ± 5.1</td>
<td>0.128 ± 0.004</td>
</tr>
<tr>
<td>40°C/75 %RH</td>
<td>3</td>
<td>Pass</td>
<td>102.4</td>
<td>4.5</td>
<td>312.6 ± 6.1</td>
<td>0.114 ± 0.029</td>
</tr>
</tbody>
</table>

BZK, Benzalkonium Chloride; RH, Relative Humidity.
The 2014 Military Health System Research Symposium Awards: Recognizing Those Pushing the Science Envelope

Col Patricia A. Reilly, USAF BSC (Ret.)

INTRODUCTION

The 2014 Military Health System Research Symposium (MHSRS) broke all previous records for both abstract submission and attendance (Table I)!
It comes as no surprise, then, that the number of awards submissions followed suit.
While awards were given out at previous Advanced Technology Applications to Combat Casualty Care and MHSRS events, 2014 put in place a structured process of award solicitation, as well as defined scoring and selection criteria.
The response was overwhelming! Thirty-one individual and team nomination packages spread over three categories were received.
Nominees were submitted from all three Services as well as academia.
All packages were reviewed by a tri-Service panel assisted by a representative from the conference sponsor, Health Affairs.
In addition, the conference sponsored two new oral competitions focused on recognizing up-and-coming young investigators: the American College of Surgeons Region 13 (Military Region) Resident Presentations and the Young Investigator Competition.
These on-site competitions joined the traditional poster awards, which cited 10 outstanding posters out of the 387 displayed.
The individuals cited in the following pages survived stiff competition to emerge as the best of the best in 2014.
Congratulations to all of them!

DISTINGUISHED SERVICE AWARD

Designed to recognize individuals who, over the years, have contributed significantly to the success of military health system research (Fig. 1).

2014 Winner
John Parrish, MD, Massachusetts General Hospital

Citation
In recognition of an exceptionally meritorious career in military health system research.
Dr. Parrish, largely motivated by his experience as a Navy doctor serving in the battlefield in Vietnam, sought to utilize technology to improve the care available to warfighters and society.
Based on his prolific experience as a pioneer of numerous technology-enabled treatments himself, he established the Consortia for Improving Medicine through Innovation & Technology, and demonstrated its effectiveness as a model to enable clinicians, researchers, and engineers to engage in and contribute to collaborative research for military medicine.

Submitted by Steven Schachter and Dr. John Collins

OUTSTANDING RESEARCH ACCOMPLISHMENT/INDIVIDUAL

Designed to recognize outstanding research contributions by an individual research scientist with the focus on significant accomplishment(s) of high impact achieved during the past year (Fig. 1).

2014 Winner
LTC Brett Freedman, MC USA, Landstuhl Regional Medical Center

Citation
LTC Freedman and his Near-Infrared Spectroscopy (NIRS) Team are commended for their exhaustive translational approach to bringing a clinically relevant, useful, and non-invasive device to market for the detection and monitoring of acute compartment syndrome (ACS) following extremity trauma.
Over the course of two completed clinical trials and three animal studies, LTC Freedman and his Team have convincingly and repeatedly demonstrated the potential utility of NIRS techniques, as well as the essential need to monitor uninjured control compartments.
Additionally, as a noninvasive modality for the diagnosis of ACS, NIRS can offer continuous monitoring of injured compartments at risk that both improves upon existing “spot check” techniques and permits the monitoring of combat casualties in transit.
With continued diligent efforts and building upon their work to date, this team will collectively revolutionize the diagnosis, monitoring and, management of ACS to the benefit of hundreds of wounded warriors.

Submitted by: Dr. Michael Shuler and LTC Benjamin Potter

OUTSTANDING RESEARCH ACCOMPLISHMENT/TEAM

Designed to recognize outstanding research contributions by a team of research scientists with focus on significant accomplishment of high impact achieved during the past year (Fig. 2).
2014 Winner/Intramural Team
Naval Medical Research Center, Bethesda, MD

Team Lead
CDR Jonathan Forsberg, MC USN

Team Members
Jonathan Forsberg, Benjamin Kyle Potter, Eric Elster, Tom Davis, Nicole Crane, Dane Hope, Elizabeth Polfer, Keith Alfieri, Korboi Evans, Fred O’Brien

<table>
<thead>
<tr>
<th>TABLE I. Statistics on the 2014 MHSRS</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstracts Submitted</td>
<td>520</td>
<td>1,135</td>
</tr>
<tr>
<td>Abstracts Accepted</td>
<td>516</td>
<td>624</td>
</tr>
<tr>
<td>Number of Review Panels</td>
<td>20</td>
<td>23</td>
</tr>
<tr>
<td>Posters Presented</td>
<td>252</td>
<td>387</td>
</tr>
<tr>
<td>Young Investigator Submissions</td>
<td>N/A</td>
<td>131</td>
</tr>
<tr>
<td>Number of Registrants</td>
<td>1,175</td>
<td>1,524</td>
</tr>
<tr>
<td>Number of Exhibitors</td>
<td>80</td>
<td>76</td>
</tr>
<tr>
<td>International Attendees</td>
<td>Not Counted</td>
<td>84/16 Countries</td>
</tr>
</tbody>
</table>

FIGURE 1. The 2014 MHSRS individual award winners pictured with RADM Doll. The picture on the left shows the Distinguished Service Award winner, John Parrish, MD. The picture on the right shows the Outstanding Research Accomplishment/Individual Award winner, LTC Brett Freedman, USA.

FIGURE 2. The 2014 MHSRS Outstanding Research Accomplished/Team Award winners with RADM Doll. The picture on the left shows Intramural Team lead CDR Jonathan Forsberg, MC USN. The picture on the right shows the Academic Team lead, Dr. M. David Rudd (center) and team member Dr. Craig Bryan (left).
Commander Jonathan Forsberg has led a collaborative team of orthopedic surgeons and basic scientists over nearly the past decade in combat-related trauma research. While the research team contributions have been many, the culmination of the team’s research is illustrated by a recently published prospective study of 200 combat-wounded active duty service members who sustained high-energy extremity injuries between 2008 and 2012. This prospective study demonstrated that careful analysis of inflammatory mediators could potentially correlate complications such as wound failure and heterotopic ossification with distinct systemic and local inflammatory profiles. This study applies advanced science to the assessment of combat wounds that unfortunately has not changed substantively during the last century, until now.

Submitted by LCDR Scott Tintle

2014 Winner/Academic Team
University of Memphis, Memphis, TN and University of Utah, Salt Lake City, UT

Team Lead
M. David Rudd, PhD

Team Members
M. David Rudd, PhD (University of Memphis) and Craig J. Bryan, PhD (University of Utah)

FIGURE 3. The 2014 Young Investigator Competition Award winners shown with RADM Doll and Col Rasmussen. Upper left - First Place: Lt Joseph D. Roderique, USN, MD, Naval Medical Center, San Diego, CA. Title: Evaluation of High-Dose Hydroxocobalamin and Ascorbic Acid as an Injectable Antidote for Carbon Monoxide Poisoning. Bottom left - Second Place: Ms. Somayyeh Mossadeq, BM, Royal Air Force, Royal Centre for Defence Medicine, United Kingdom. Title: A Novel Anatomical and Physiological Scoring System for Military Pelvic and Perineal Blast Injuries Shows Significantly Superior Predictive Outcomes than Current Scoring Systems; Injury Severity Score-ISS, New ISS, Trauma ISS, Revised Trauma Score. Upper right - Third Place: Ernesto Lopez, MD, University of Texas Medical Branch, Galveston, Texas. Title: Nebulized Epinephrine Limits Pulmonary Vascular Hyperpermeability to Water and Proteins in Ovine Burn and Smoke Inhalation Injury.


Citation
For exceptionally meritorious accomplishments as a research team from September 2009 to August 2014. M. David Rudd, Craig Bryan, Evelyn Wertenberger, Sean Williams, and Kim Arne of the Army Suicide Prevention and Intervention Research at Evans (ASPIRE) team have distinguished themselves as leaders in military suicide research by establishing and leading a cutting-edge research program that has led to improved methods for assessing and treating suicide risk among members of the U.S. Armed Forces. Their innovative brief cognitive behavioral therapy contributed to 60% reductions in suicide attempts and 50% gains in retention among active duty soldiers while dramatically reducing inpatient hospitalization days by more than half. The pioneering achievements of the ASPIRE team serve as an example to researchers and clinicians and reflect great credit upon them,

<table>
<thead>
<tr>
<th>Award</th>
<th>Title</th>
<th>Authors</th>
<th>Institutional Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gold/First Place</td>
<td>Stem Cell-Derived Inner Ear Organoids for High-Content Drug Screening</td>
<td>Karl Koehler* Xiao-Ping Liu† Andrew Mikosz* Jeffrey Holt† Eri Hashino*</td>
<td>*Indiana University School of Medicine †Harvard Medical School</td>
</tr>
<tr>
<td>Silver/Seond Place</td>
<td>Can Mathematical Models Help Diagnose Lower-Airway Lung Diseases?</td>
<td>Bora Sul* Anders Walkvist* Michael Morris† Jaques Reifman* Vineet Rakesh*</td>
<td>†Telemedicine and Advanced Technology Research Center Fort Detrick, MD ⁜San Antonio Military Medical Center, San Antonio, TX</td>
</tr>
<tr>
<td>Bronze/Third Place</td>
<td>Inter-Observer Reliability of Reporting of Injury Severity Scoring Following Combat Trauma; Different Perspectives, Different Values?</td>
<td>Iain M Smith†‡ David N Naumann* Paul Guyer§ Jonathan Bishop‡ Jonathan B Lundy∥ Douglas M Bowley*</td>
<td>⁯Royal Centre for Defence Medicine, United Kingdom</td>
</tr>
<tr>
<td>Honor Mention</td>
<td>Investigation of the Concussion Mechanism: An End-to-End Model That Translates External Measures to Internal Neurologic Injury for Head Impact</td>
<td>Dr. Laurel Ng Melissa Gibbons, Pi Phohomsiri Vladislav Volman, Jianxia Cui Darrell Swenson David Shelley Jessica Wong Kiran Mathews James Stuhmiller</td>
<td>‡202 (Midlands) Field Hospital, United Kingdom ⁯NIHR Surgical Reconstruction and Microbiology Research Centre ⁯Institute of Naval Medicine ⁯United States Army Institute of Surgical Research, San Antonio, TX L-3/ATI</td>
</tr>
<tr>
<td></td>
<td>Changes in Barometric Pressure Alter Susceptibility to Central Nervous System Oxygen Toxicity</td>
<td>Heather E. Held* Raffaele Pilla* Csilla Ari† Tina Fiorelli‡ Carol S. Landon* Jay B. Dean*</td>
<td>*Morsani College of Medicine, University of South Florida Tampa, FL †Byrd Alzheimer’s Institute University of South Florida, Tampa, FL</td>
</tr>
<tr>
<td></td>
<td>Evaluation of Prehospital Pelvic Binder Usage in U.S. Military Casualties With Pelvic Fractures</td>
<td>Chetan U. Kharod* Creighton C. Tubb† Kirby Gross‡ John F. Kragh, Jr‡</td>
<td>*San Antonio Uniformed Services Health Education Consortium, San Antonio, TX †San Antonio Military Medical Center, San Antonio, TX ⁯U.S. Army Institute of Surgical Research, San Antonio, TX</td>
</tr>
<tr>
<td></td>
<td>Lumbar Spine Postures in Marines During Simulated Operational Conditions</td>
<td>David Berry† Karen Kelly* Ana Rodrigues-Soto† Sara Gombatto‡ Rebecca Jaworski* Sam Ward†</td>
<td>*Naval Health Research Center, San Diego, CA †University of California, San Diego, La Jolla, CA ⁯San Diego State University, San Diego, CA</td>
</tr>
</tbody>
</table>
the National Center for Veterans Studies, their universities, and the Department of Defense.

Submitted by Dr. Craig Bryan

YOUNG INVESTIGATOR COMPETITION
This category was designed to highlight and promote the research accomplishments of residents, fellows, postdocs within 5 years of graduation from a terminal degree, and service academy cadets. The 2014 response to the Young Investigator Category was overwhelming! A total of 130 abstracts were received for consideration by a panel of 10 tri-Service subject matter experts. Of those 130, 12 abstracts were selected for oral presentation at a special plenary session at the 2014 MHSRS. The competition winners are pictured in Figure 3.

AMERICAN COLLEGE OF SURGEONS COMMITTEE ON TRAUMA REGION 13 RESIDENT PRESENTATIONS
For the first time, the American College of Surgeons asked the MHSRS to host their annual Region 13 (Military Region) resident trauma presentations. The session featured the six

FIGURE 4.  The 2014 MHSRS Poster first, second and third place, and honorable mention awards were accepted by the individuals shown above, pictured with RADM Doll. The picture on the upper left shows Dr. Karl Koehler, Indiana University, accepting the first place award. The picture on the upper right shows Dr. Vineet Rakesh, Fort Detrick, MD, accepting the second place award. The picture on the lower left shows Maj. Iain M Smith, Royal Centre for Defence Medicine, United Kingdom accepting the third place award. The picture on the bottom right shows the honorable mention award winners.
best trauma papers of the year. In the Clinical Science Category, the winner was Capt Dylan Pannell, Canadian Forces for his presentation on Factors Associated with Post Traumatic Stress Disorder Amongst Canadian Armed Forces Personnel Deployed to Afghanistan. In the Basic Science Category, the winner was CPT Jonathan J. Sexton, USA for his presentation on Administration of FTY720 During Tourniquet-Induced Hind Limb Ischemia-Reperfusion Injury Attenuates Morbidity and Mortality.

**2014 MHSRS POSTER COMPETITION**

A total of 387 posters were presented over two sessions, a 52% increase over 2013. Ten posters received special recognition (Table II and Fig. 4).
ACKNOWLEDGEMENTS

Special thanks to the following individuals for their contribution to the 2014 Military Health System Research Symposium (MHSRS) and this supplement.

Editorial Support
Ms. Teresa L. Hendrickson, Combat Casualty Care Research Program: Capitol IT Solutions
Mr. Ramin A. Khalili, Combat Casualty Care Research Program: Isom Global Strategies

MHSRS Website and IT Support
Ms. Kelly A. Conboy, Universal Consulting Services, Inc.
Ms. Evelyn K. Stephenson, US Army Medical Research Acquisition Activity

MHSRS Planning
Mr. Robert Sarvaida, Col USAF (Ret)
Marriott Harbor Beach Resort & Spa, Ft Lauderdale, FL

Supplement Reviewers
Col Nicole H. Armitage, USAF
Khalid W. Barazanji, PhD
Surgeon Captain Stephen Bree, Royal Navy
LTC Teresa L. Brininger, USA
Sylvain Cardin, PhD
Lt Col David T. Carey, USAF
LTC Robert Carter, USA
LTC Kristen L. Casto, USA
Wayne L. Chappelle, PhD
Lt Col Paul F. Crawford, USAF
Alicia T. Crowder, PhD
LTC Jurandir J. Dallelucca, USA
Lt Col Michael R. Davis, USAF
Lt Col Susan F. Dukes, USAF
CDR Elliot M. Jessie, USN
CDR Dennis J. Faix, USN
Col Raymond Fang, AF
Douglas B. Gibson, PhD
CDR Carl W. Goforth, USN
Tony D. Gover, PhD
Lt Col Jennifer J. Hatzfeld, USAF
MAJ Jeremy R. Hershfield, USA
COL Sidney R. Hinds, USA
Jeffrey T. Howard, PhD
MAJ H. Dean Hudson, USA
Saber M. Hussain, PhD
Judson C. Janak, PhD
Col Cheten U. Kharod, USAF
Bijan S. Kheirabadi, PhD
Michael Kinsky, MD
Maj Kevin R. Kupferer, USAF
Frank J. Lebeda, PhD
Lori A. Loan, PhD
Victor W. MacDonald, PhD
Carlos Maldonado, PhD
Donald W. Marion, MD
Stephanie N. Maxfield-Panker, PhD, DPT
Maj Jonathan J. Morrison, RAMC (V), British Army
Darcy L. Mortimer, MSN
LTC John S. Oh, USA
LTC Kyle N. Remick, USA
Col Stacey Shackelford, USAF
Maj Brandon M. Tourtillot, USAF
Heuy-Ching H. Wang, PhD
Lt Col Brian A. Young, USAF